ГОДИШНИК
НА БЪЛГАРСКАТА НАРОДНА БАНКА

2015-2018

ANNUAL
OF THE BULGARIAN NATIONAL BANK

ТОМ/VOLUME 5

БЪЛГАРСКА НАРОДНА БАНКА
BULGARIAN NATIONAL BANK

ГОДИШНИК
на Българската народна банка

ANNUAL
of the Bulgarian National Bank

Издателски съвет на Българската народна банка

Председател:
Калин Христов, подуправител и член на Управителния съвет на БНБ

Зам.-председател:
Виктор Илиев

Членове:
Елица Николова, член на Управителния съвет на БНБ
Николай Неновски, проф. д. ик. н, член на Управителния съвет на БНБ
Ивайло Николов, д-р
Даниела Минкова, д-р
Зорница Владова

Секретар:
Людмила Димова

Заместник-секретар:
Христо Яновски

Bulgarian National Bank Publications Council

Chairman:
Kalin Hristov, Deputy Governor and Member of the BNB Governing Council

Vice Chairman:
Viktor Iliev

Members:
Elitsa Nikolova, Member of the BNB Governing Council
Nikolay Nenovsky, Prof., Ph. D. Member of the BNB Governing Council
Ivaylo Nikolov, Ph. D.
Daniela Minkova, Ph. D.
Zornitsa Vladova

Secretary:
Lyudmila Dimova

Assistant Secretary:
Christo Yanovsky

Включените изследвания са анонимно двойно рецензирани.
Авторското право върху изследванията е на техните автори.
Материалите отразяват гледищата на своите автори и не ангажират позицията на БНБ.
Евентуалните грешки и пропуски са изцяло отговорност на авторите.

Studies included in this publication are double-blind peer reviewed.
The studies copyright reserved by the authors.
Views expressed in these studies are those of the authors and do not necessarily reflect the BNB policy.
Responsibility for the non-conformities, errors and misstatements lies entirely with the authors.

Мненията си изпращайте до:
Издателски съвет на БНБ

Send your comments and opinions to:
Bulgarian National Bank Publications Council

e-mail: BNB_Publications@bnbank.org
ISSN 2683-0728 (online)
www.bnb.bg

ГОДИШНИК
на Българската народна банка

2015–2018

ANNUAL
of the Bulgarian National Bank

Том/Volume 5

БЪЛГАРСКА НАРОДНА БАНКА
BULGARIAN NATIONAL BANK

© Peter Stoyanov, DP/98/2015
© Petar Peshev, DP/99/2015
© Zornitsa Vladova, Mihail Yanchev, DP/100/2015
© Tania Karamisheva, DP/102/2016
© Petra Kotseva, Mihail Yanchev, DP/103/2017
© Daniel Kasabov, Petra Kotseva, Andrey Vassilev, Mihail Yanchev, DP/104/2017
© Neli Ivanova, Evgeni Ivanov, DP/105/2017
© Mariya Mitkova, DP/106/2017
© Михаил Михайлов, DP/107/2017
© Ventsislav Hristev, DP/108/2017
© Evgeni Ivanov, DP/109/2018
© Evgeni Ivanov, Neli Ivanova, Daniel Kasabov, Kristina Karagyozova DP/110/2018

© Българска народна банка, 2020 г., ГОДИШНИК
© Bulgarian National Bank, 2020, ANNUAL

Contents of Volume 5

DP/98/2015
INTERNATIONAL MERCHANDISE TRADE FLOWS: DEFINING SAMPLES AND IDENTIFYING DISCOVERIES 687
Peter Stoyanov

DP/99/2015
DETERMINANTS OF INTEREST RATE SPREADS IN BULGARIA 731
Petar Peshev

DP/100/2015
EMPIRICAL EVIDENCE ON THE RELATIONSHIP BETWEEN MONEY SUPPLY DYNAMICS AND PRICES IN BULGARIA 771
Zornitsa Vladova, Mihail Yanchev

DP/102/2016
ANALYSIS OF THE BANK LENDING SURVEY RESULTS FOR BULGARIA (FOR THE 2003–2014 PERIOD) 819
Tania Karamisheva

DP/103/2017
ANALYSIS OF THE HOUSING MARKET DEVELOPMENTS AND THE UNDERLYING MACROECONOMIC FUNDAMENTALS IN BULGARIA 869
Petra Kotseva, Mihail Yanchev

DP/104/2017
RELATIONSHIP BETWEEN INFLATION, POTENTIAL OUTPUT AND STRUCTURAL UNEMPLOYMENT IN BULGARIA 919
Daniel Kasabov, Petra Kotseva, Andrey Vassilev, Mihail Yanchev

DP/105/2017
THE ROLE OF BULGARIA IN GLOBAL VALUE CHAINS 951
Neli Ivanova, Evgeni Ivanov

DP/106/2017
LABOUR MARKET MODEL WITH ON-THE-JOB SEARCH, AN AGENT-BASED APPROACH 987
Mariya Mitkova

DP/107/2017
DISEQUILIBRIUM MODEL OF THE CREDIT MARKET IN BULGARIA 1021
Mihail Mihaylov

DP/108/2017
MACROPRUDENTIAL BANK STRESS-TESTING – THEORETICAL FUNDAMENTALS, FRAMEWORK, AND CHALLENGES 1071
Ventsislav Hristev

DP/109/2018
CONSTRUCTING AN UNCERTAINTY INDICATOR FOR BULGARIA 1111
Evgeni Ivanov

DP/110/2018
DETERMINANTS OF INVESTMENT IN BULGARIA 1151
Evgeni Ivanov, Neli Ivanova, Daniel Kasabov, Kristina Karagyozova

Съдържание на том 5

DP/98/2015
МЕЖДУНАРОДНИ ТЪРГОВСКИ ПОТОЦИ: ИДЕНТИФИКАЦИЯ НА ТЪРГОВСКИ ОБРАЗЦИ И НОВОВЪЗНИКНАЛИ ВРЪЗКИ 687
Петър Стоянов

DP/99/2015
ДЕТЕРМИНАНТИ НА ЛИХВЕНИТЕ СПРЕДОВЕ В БЪЛГАРИЯ 731
Петър Пешев

DP/100/2015
ЕМПИРИЧНИ ДОКАЗАТЕЛСТВА ЗА ЗАВИСИМОСТТА МЕЖДУ ДИНАМИКАТА НА ПАРИЧНОТО ПРЕДЛАГАНЕ И ЦЕНИТЕ В БЪЛГАРИЯ 771
Зорница Владова, Михаил Янчев

DP/102/2016
АНАЛИЗ НА РЕЗУЛТАТИТЕ ОТ АНКЕТАТА ЗА КРЕДИТНАТА АКТИВНОСТ НА БАНКИТЕ В БЪЛГАРИЯ (ЗА ПЕРИОДА 2003–2014 г.) 819
Таня Карамишева

DP/103/2017
АНАЛИЗ НА ДИНАМИКАТА НА ЖИЛИЩНИЯ ПАЗАР И ОСНОВНИТЕ МАКРОИКОНОМИЧЕСКИ ПОКАЗАТЕЛИ В БЪЛГАРИЯ..... 869
Петра Коцева, Михаил Янчев

DP/104/2017
ЗАВИСИМОСТ МЕЖДУ ИНФЛАЦИЯТА, ПОТЕНЦИАЛНОТО ПРОИЗВОДСТВО И СТРУКТУРНАТА БЕЗРАБОТИЦА В БЪЛГАРИЯ 919
Даниел Касабов, Петра Коцева, Андрей Василев, Михаил Янчев

DP/105/2017
РОЛЯТА НА БЪЛГАРИЯ В СВЕТОВНИТЕ ВЕРИГИ ЗА СЪЗДАВАНЕ НА СТОЙНОСТ 951
Нели Иванова, Евгени Иванов

DP/106/2017
МОДЕЛ ЗА ТРУДОВИЯ ПАЗАР ПРИ ТЪРСЕНЕ НА РАБОТА БЕЗ НАПУСКАНЕ НА СЪЩЕСТВУВАЩАТА, АГЕНТНО ОРИЕНТИРАН ПОДХОД 987
Мария Миткова

DP/107/2017
НЕРАВНОВЕСЕН МОДЕЛ НА КРЕДИТНИЯ ПАЗАР В БЪЛГАРИЯ 1021
Михаил Михайлов

DP/108/2017
ПРОВЕЖДАНЕТО НА МАКРОПРУДЕНЦИАЛНИ СТРЕС ТЕСТОВЕ НА БАНКИТЕ – ТЕОРЕТИЧНИ ОСНОВИ, РАМКА И ПРЕДИЗВИКАТЕЛСТВА 1071
Венцислав Христев

DP/109/2018
КОНСТРУИРАНЕ НА ИНДИКАТОР НА НЕСИГУРНОСТТА ЗА БЪЛГАРИЯ 1111
Евгени Иванов

DP/110/2018
ДЕТЕРМИНАНТИ НА ИНВЕСТИЦИИТЕ В БЪЛГАРИЯ 1151
Евгени Иванов, Нели Иванова, Даниел Касабов, Кристина Карагьозова

DP/98/2015

International Merchandise Trade Flows: Defining Samples and Identifying Discoveries

Peter Stoyanov

Abstract. This paper aims to contribute to the study of export and import product discoveries by examining the definitions of trading samples and discoveries. It first looks at the definition of what constitutes a 'trading sample' by contrasting the results of applying the traditional dollar threshold (at $0, $10,000 and $100,000), four relative criteria (based on value and quantity) and two composite relative criteria (using both value and quantity). The nine examined definitions differ substantially both in the number of flows they tag as samples and in the individual flows they tag as samples. Second, a definition of 'product discovery' – similar to that used in Klinger & Lederman (2011) and Cadot et al. (2011) but allowing for a flexible middle window (the sending samples phase) – is presented and some of its potential benefits are discussed. The three examined definitions of discovery differ substantially in the flows they tag as discoveries. Despite the embedded flexibility in the proposed definition, in the majority of discovery episodes, for both exports and imports, products jump straight to established product status within less than a year. This result is robust across the nine examined definitions of trading samples, and could be linked to models such as Rauch & Watson (2003) and the empirical findings in Besede & Prusa (2006). The analysis is conducted at the bilateral level, using UN Comtrade data covering 1996–2012, at the 6-digit HS'1996 product level.

Резюме. Настоящата статия има за цел да улесни идентификацията на нововъзникващите търговски връзки (внос или износ) чрез прецизиране на разграничаването им от пробните търговски потоци. Разгледани са различни дефиниции на пробен търговски поток на база традиционното разграничение въз основа на стойността (три гранични стойности – 0 долара, 10 000 долара и 100 000 долара), четири варианта на база относителни размери (въз основа на стойността или физическите мерни единици) и два съставни критерия (едновременно стойност и физическо количество). Общо деветте различни дефиниции водят до значително различаващи се резултати както по отношение на общия брой пробни потоци, така и по отношение на конкретните потоци, идентифицирани като пробни.
Разгледани са и три дефиниции на нововъзникнала търговска връзка, в смисъла на *Klinger & Lederman* (2011) и *Cadot et al.* (2011), но с възможност за различна по дължина втора фаза (тестване на пазара с пробни потоци). Различните дефиниции отново водят както до различен общ брой нововъзникнали връзки, така и до различни потоци, които са идентифицирани като такива. Независимо от гъвкавостта на използваната дефиниция, в преобладаващия случай фазата на тестване на пазара е по-кратка от година. Този резултат се наблюдава и при деветте използвани дефиниции, като може да бъде свързан с модели от класа на представените от *Rauch & Watson* (2003) и емпиричните резултати на *Besedeš & Prusa* (2006).
Анализът е извършен на база данни от системата *Comtrade* на ООН за периода 1996–2012 г. на ниво 6 знака (HS'1996) и двустранни връзки.

Contents

Introduction .....

1. Data .....

2. What Constitutes a 'Trading Sample' .....
   - 2.1. Alternative Definitions of 'Trading Sample' .....
   - 2.2. Transition between States (no trade, sample and established product) .....
   - 2.3. Defining Samples: an Look at Bulgaria .....

3. Identifying Export Discoveries .....
   - 3.1. Extensive and Intensive Margins in International Trade ..
   - 3.2. Identification of Export Discoveries: Overview .....
   - 3.3. Identification of Product Discoveries with a Flexible Middle Window .....
   - 3.4. Defining the Year When the Discovery is Made .....
   - 3.5. Comparing the Results of the Three Definitions .....
   - 3.6. The Length of Window 2 .....

4. Discussion and Conclusion .....

References .....

Appendix A. Comparing the Definitions of 'Sample' .....

Appendix B. Transition Matrices by Definition of 'Sample': Exports and Imports .....

Appendix C. Comparing the Three Definitions of Discovery .....

Appendix D. Length of the Samples Phase in the PS Definition of Discovery .....

Acknowledgment: I would like to thank Andrey Vassilev from the Economic Research and Forecasting Directorate of the Bulgarian National Bank, for helpful discussions and detailed comments on earlier drafts of this paper.

Peter Stoyanov
Sofia University "St. Kliment Ohridski"
Faculty of Economics and Business Administration
peter.stoyanov@feb.uni-sofia.bg

Introduction

A strand of recent literature analyzes international trade via its extensive and intensive margins, *i.e.* whether growth happens within the same products or through the discovery (and the subsequent survival or disappearance) of new products. This paper is linked to the empirical work aimed at establishing stylized facts that are, at a later stage, examined and explained by a formal theoretical model. The aim is to contribute by (i) underlining the importance of how we define small trade flows, or 'trading samples', and (ii) by offering a definition of discovery that allows for more flexibility than the ones currently employed in the literature and thus helping enhance the understanding the dynamics of innovation implied by the emergence of new products in a country's export basket. Both issues are directly relevant to policymakers contemplating diversification opportunities. Considering the substantially different results produced by applying the approaches currently available in the literature — such as Klinger & Lederman (2011) and Cadot et al. (2011) — and the more general method described in this paper, there are potentially important implications in distilling the vast amounts of trade data into stylized facts and the subsequent analysis and policy formulation.

The analysis of product discoveries typically employs a base-line period (window 1), a second window during which some magic happens and a discovery is made, and a third window, during which it is verified whether the product has emerged as an established product or has failed to. While the literature often does not explicitly discuss the technical details on the construction of windows, it seems that the common practice is to employ fixed windows along at least one of the three main dimensions (starting point, ending point, and duration). This is unnecessarily restrictive, as it implies that discoveries in all possible commodities, and among all possible trading country pairs, follow the same temporal paths — something that is difficult to justify. This paper offers a more flexible approach, which (a) allows the middle window to be of variable length, within user-defined limits, and (b) allows for an interleaved discovery processes, instead of opting for waves of a fixed window structure whenever data is available over a longer period of time. The flexibility is applied at the lowest level of data availability (reporter-partner-commodity), making it possible for a single commodity to be 'discovered' several times within the same trading country pair, and to have different length of window 2 for each of the discovery episodes.

Some of the definitions of export discovery in the literature allow for the product to be exported as 'samples' before emerging as an established product but to the best of the author's knowledge there has been no systematic effort in defining what constitutes a trading sample. The current practice in the literature is to set a threshold dollar value, usually without formal justification, and to label all trade flows below this threshold value 'samples'. Similarly to the fixed windows structure, this is overly restrictive — the thresholds may well be different for different products within the same analysis. For example, consider the model developed in Rauch & Watson (2003). They model the decision of a developed-country buyer considering buying from unknown less-developed-country suppliers — the buyer may choose to place the full order (and face a potential loss if the supplier does not perform to expectations), or start with placing small orders until sufficient knowledge of the partner is accumulated. Since the size of the small orders is linked to the cost of training the supplier and the size of the full order, it is to be expected that what is considered a small order (or a sample in the terminology of the export discoveries literature) is dependent on both the trading partner and the product.

The discussion in the paper is not linked to any of the many theoretical models but most of the arguments are applicable in cases where the model allows for small vs large flows and existing vs new products. These include both models where the dynamics are generated on the production side (heterogeneous firms, multi-product firms) or on the demand side (searching for a dependable partner). In both cases the proper formulation of 'small flows', or samples, at the stylized facts phase of the analysis is important.

The proposed definition of discovery could be applied in establishing and classifying a set of 'patterns' in the discovery episodes — whether certain products for example require more years of sending samples, whether the length of the samples phase is dependent on the source and the destination of the flow (*i.e.* do north-north, north-south, south-south, and south-north flows require systematically different samples phases), whether the length of the samples phase is linked to the subsequent 'stability' of the flow, *etc.* One potential use of such set of patterns is illustrated in section 3.6, where the results generated by the approach employed in this paper relate to the well-known results in Besedeš & Prusa (2006) on the relatively short median duration of trade flows. Another possible use would be to combine the set of frequently-observed patterns with the 'product space' approach offered by Hidalgo et al. (2007). The product space reflects the typical connections between products, allowing the construction of a set of prod-

ucts that are close to the country's capabilities. This information can be combined with the set of typical discovery patterns (since moving to a new product will constitute a discovery), *i.e.* will allow to estimate the expected speed of the new product becoming an economically important flow.

The paper is structured as follows. Section 1 describes the dataset used in the analysis and the necessary adjustments that must be made prior to the analysis. Section 2 discusses the issue of defining what would constitute a trading sample, and provides quantitative comparison of the results applied to both exports and imports flows. Section 3 describes the need for a more flexible definition of what constitutes a new product discovery, presents the proposed definition with a flexible middle window, and provides a comparison to two of the approaches used in the literature. Section 4 motivates the need for further discussion on the topic and concludes.

1. Data

This paper uses UN Comtrade data on annual export and import flows over 1996–2012. The data is at the 6-digit commodity level, using the 1996 version of the HS nomenclature. Commodity aggregates and "UN special" codes have been removed, leaving a total of 5,111 commodities. Economy aggregates ('n.e.s.' entries and similar) have also been removed, as they do not represent economic entities and cannot be matched to country statistics. The data contains dollar value and quantity information, both in kilograms and in supplementary quantity units. A single quantity indicator was constructed, giving preference to the supplementary quantity units. Where supplementary units were not available in the original data (not reported or reported as zero), net weight in kg was used instead. About 1 percent of the total number of flows have no quantity information reported. In about 1.6 percent of the flows the value for quantity is reported as zero, which is considered to be unrealistic and has been treated as 'not available'.

An important aspect of UN Comtrade data is that different countries use different rules regarding the minimum size of flows that are reported. In 1996 and 2006 about 37 percent of national customs organizations used a threshold value below which the flow was not reported.[1] This issue is very important because it leads to bias in the discovery counts both across countries and across time, regardless of the definition of trading sample or product discovery used. Given the available information, there seems to

[1] See question 32 (1996 survey) and questions 9.01 and 9.02 (2006 survey) in the *International Merchandise Trade Statistics National Compilation and Reporting Practices Survey Results 2006 and 1996*.

be no objective way to distinguish between a flow which is genuinely new (*i.e.* did not exist before) from a flow which had previously existed but was not reported, or even a consistent source listing which countries used what thresholds in which years.

The issue is probably most visible in the case of Poland. Prior to 2004, there are no flows below $50,000 in the dataset. Starting in 2004, the threshold was apparently dropped, and there are flows with reported value as low as $1. Between 2003 and 2004 Poland's exports in the dataset grew by 39.8 percent in terms of dollar value and by 273 percent in terms of number of flows. The sum total of export flows which were below $50,000 in 2004 was just 0.78 percent of total export value but accounted for 67.8 percent of the number of flows. For Poland 2004 was not a 'normal' year — it joined the European Union on 1 May 2004 — and a boost to exports and export discovery activity is to be expected but likely not in such extreme amounts in terms of number of new flows.

In an attempt to mitigate the spurious emergence of false discoveries, flows with dollar value of $2500 and below have been removed from the dataset. This threshold is the one apparently applied by the USA, and is the second-highest after Poland.[2] This decision is costly in terms of number of data points but not so much in terms of total value of flows — between 10 and 25 percent of the number of export flows are eliminated each year but they account for less than 0.2 percent of the respective year's world trade in terms of total dollar value. Ultimately we're left with a dataset of slightly over 68 million exports and 69 million imports flows, covering 5,111 products, 187 reporter and 234 partner economies, over 1996–2012. Quantity data is missing in about 1.7 million of both the exports and imports flows, incl. cases where the zero quantity entries were replaced by missing data.

2. What Constitutes a 'Trading Sample'

For the purposes of research similar to this paper, 'trading samples' are implicitly taken to mean any flows that test a new market, *i.e.* flows before a product can be treated as established.

The traditional approach in the empirical literature seems to be to choose an arbitrary (dollar) value that, in the respective author's opinion, corresponds to the research question they pose, and to treat all flows with a value below that threshold as samples. Influential papers like Helpman

[2] Poland still remains an issue in 2004 — this threshold removes only about half of the sub-$50,000 flows in 2004, meaning there remain about 34 thousand flows (of value between $2500 and $50,000) which are reported in 2004 but could not have been reported in 2003.

et al. (2008) and Hummels & Klenow (2005) do not use thresholds at all. Evenett & Venables (2002) use five threshold levels between 0 and 500,000 US dollars, with the intention to "reduce the likelihood of misclassified imports or economically unimportant levels of imports distorting the analysis",[3] and most often report for a threshold of 50,000 US dollars. Klinger & Lederman (2011) use a threshold of 10,000 US dollars per year but provide no details on how the value was chosen. Cadot et al. (2011) use no threshold at all and explicitly discuss that using a threshold depends on whether the researcher is interested in the overall searching process of companies attempting to enter new markets (no threshold, so that all attempts, including subsequent failures, are captured), or just the successful discovery attempts (flows that eventually matured, *i.e.* ended up above the threshold). Agosin & Bravo-Ortega (2009) examine a more focused question (success stories in Chile), and as such impose a high threshold — one million dollars in 2000 constant prices.

While such definitions may seem intuitive in the concrete cases, as long as data is available,[4] there is little reason to support (a) choosing dollar value as the relevant metric, as opposed to physical quantity, unit value or some other characteristic, and (b) fixing the same value for all products and trading economies. There is no clear alternative, and there are many approaches that could prove useful in certain scenarios. For example, one could argue that, with respect to samples, a quantity-based measure would be better than dollar value, as different companies (or customs administrations) may have different accounting policies to determine the reported dollar value of the samples, but the physical quantity units should be more uniform.[5] Another possibility is to look, at the individual flow level, for step-like dynamics in dollar value, quantity or unit value, with jumping to a plateau signifying the transition from samples to established product. Or one may look for clustering of data points. In short, the possibilities are numerous.

[3] Evenett & Venables (2002), p. 7.

[4] It should be recognized that detailed trade data availability is a fairly recent phenomenon in economics, and a substantial portion of the literature originated before that.

5 Of course, this assumes the available quantity information is of good quality, which is often not the case. An example of present but unrealistic quantity information is the exports of product 870210 "Motor vehicles for the transport of ten or more persons, including the driver; With compression-ignition internal combustion piston engine (diesel or semi-diesel)" by Jordan to Iraq. In 2001 the dollar value was $17m, but quantity was reported as 284kg — a clearly unrealistic weight for even a single passenger motor vehicle carrying ten or more persons.

2.1. Alternative Definitions of 'Trading Sample'

This section of the paper tests several definitions of what constitutes a trading sample.[6] First is the standard approach, with three dollar values for the fixed threshold — no explicit threshold (labeled **s0**), which is equivalent to a threshold of $2500, since flows below $2500 were removed, as discussed in section 1, and thresholds of $10,000 and $100,000, labeled **s10** and **s100** respectively.[7,8] Second, a set of four relative criteria are used, separately based on value and on quantity. Definitions labeled **q1** and **v1** consider a flow as a sample if it falls, in terms of quantity or dollar value respectively, in the lowest percentile of all flows of the same commodity,[9] across all years and all trading pairs. Definitions **q5** and **v5** are the same but consider the bottom five percentiles as flows of samples. Lastly, two composite criteria, **c1** and **c5**, are constructed which require that both **q1** and **v1** (for **c1**) and **q5** and **v5** (for **c5**) tag a flow as a sample to consider it as such.

Table A.1 and Table A.2 in Appendix A provide information how the results of the application of the different definitions of 'sample' stack against each other, for exports and imports respectively. The alternative definitions of 'sample' were applied to all flows in the database, and each flow was classified as a trading sample or as an established product. Each 2x2 block in the table compares the result of two criteria — in how many cases they agree on the classification (both produce a 'sample' or both produce an 'established product'), and in how many they produce opposite results (the

---

[6] Outside of those intended to mimic the approaches in the literature, the examined definitions and the respective threshold values were largely selected only considering what type of information is available in the dataset (value and physical quantity), and should be seen as proof-of-concept examples rather than something with firm theoretic background.

[7] Again, because all flows below $2500 have been removed, a flow is defined as a sample if its value is between $2500 and $10,000, respectively $100,000. This consideration applies to all value-based definitions.

[8] Note that the s10 threshold is not equivalent to the threshold used by Klinger & Lederman (2011) — the dollar value is the same but is being applied at the bilateral level here, as opposed to aggregate exports in their paper.

[9] Comtrade provides quantity measures in net weight (kg) for all products, as well as a supplementary quantity unit which differs among products, and often even within products. As described in the data section, supplementary quantities were used where available, as they are more 'natural'. Percentiles were calculated on a (commodity, supplementary unit) pair basis. Each pair comprises all trade flows for a given product, across all years and all partners, whenever the same supplementary quantity was used in the reporting. A commodity reported in different supplementary units by different countries is taken to be different products for the sake of tagging samples. Thus, there are 5,111 individual commodities in the HS'1996 classification, but in the calculation of the relative thresholds there are 13,531 (commodity, supplementary quantity unit) pairs. This may introduce some bias as the decisions of different customs administrations to use different supplementary units are unlikely to be random in nature, but the alternative — using only net weight in kg — seems even more biased.

flow is classified as a sample by one of the criteria, and as an established product by the other).

There are two important results. First, the two sets of relative criteria (**q1** and **q5**, and **v1** and **v5**) produce substantially fewer samples than using a fixed threshold, at the chosen thresholds. For example, the definition **s10** tags 14,499,631 flows as samples, while **v5** only tags 3,407,133. Most of the flows tagged as a sample by **v5** are also tagged by **s10** though, indicating that flows that are in the bottom 5 percentiles tend not to exceed $10,000 in value. The issue, however, is in the remaining circa 11m flows that are tagged as samples using the $10,000 fixed dollar threshold but are not in the bottom 5 percentiles — these could be low-value flows in industries where low-value flows are the norm rather than an expression of search activity. Under the **s10** definition, such flows would never mature into established-level products. Whether a bilateral-level flow of less than $10,000 per year should be seen as a mature product of sufficiently high economic importance depends on the concrete case and the research question.[10] Among other considerations, the scope of the individual leaf elements in the HS nomenclature varies substantially, with some product definitions being more detailed (*i.e.* restrictive) than others.

Reconciling this issue would require either lowering the dollar amount of the fixed thresholds or increasing the relative cutoff point. To get a number of flows tagged as samples by the value-based relative definition broadly similar to those by **s10** would require defining the bottom 20–25 percentiles of the flows by value as samples. For exports, setting the cutoff at the 20th percentile yields 13,643,609 flows tagged as samples, vs. the 14,499,631 tagged by **s10**. The two definitions are much closer — they agree in 12,264,384 cases and disagree in 3,614,472 — but it is not clear whether a full one-fifth of the value distribution of the flows should be considered to be samples. Conversely, lowering the fixed dollar threshold would negatively impact the economic importance of those flows.

The second important result is that the disagreement between the different definitions is usually substantial, meaning that even if the alternative definitions tag a similar *number* of flows as samples, these are usually *different* flows. This is most visible when comparing the quantity-based to the value-based definitions — for exports, **q5** tags 3,006,594 flows as samples and **v5** tags 3,407,133. The overlap is just 995,451 flows though, and the defini-

[10] While the amount may seem small from the point of view of a developed-country business, the situation in a developing country may well be dramatically different. For example, analyzing Rwanda's export diversification, Chandra et al. (2007) (p. 161) conclude that "Each US$1000 increment of a non-traditional export is a precious achievement."

tions disagree on over 4.3 million flows. One possible explanation is that, if the analysis is carried out on the basis of dollar value alone, one cannot distinguish between flows of differing qualities within the same commodity — a flow may be an established product of low quality, hence low unit value, and even though it is being exported in sufficiently high quantities, it may remain under the value-based definition of 'established product' and never be tagged as one (a false negative). Similarly, it would be possible that sending out just a few samples of an expensive variety causes the flow to erroneously reach maturity status (a false positive).

Definition **s100** is an interesting case — it classifies more flows as samples (about 41-42m, for both exports and imports) than as established products (about 27m). Considering that a large number of small (sub-$2500) flows have already been removed from the dataset as discussed in section 1, this should probably be interpreted that setting a fixed threshold of $100,000 pushes a very high share of international trade in the realm of search activity rather than established product trade — a result that might be too strong for most general research efforts. This serves to underscore a tentative conclusion based on the considerations above — the choice of a definition of what constitutes a sample is important and should be tailored to the concrete research question and the theoretical basis.

2.2. Transition between States (no trade, sample and established product)

Another way to compare the definitions is to look at the probabilities for a given flow to transition from one state to another state in the subsequent period (the transition matrices) for the different definitions. The different sample definitions imply three possible states[11] in which a flow can be — non-existent state (**o**), sample state (**s**), or established product state (**X**).

As highlighted by Helpman et al. (2008), there is no trade between a significant share of all possible pairs of economies, esp. when looking at a very disaggregated product level, but those 'zeroes' carry important information. Generating the full set of relationships covering all possible country pairs for all commodities, however, is technically difficult to handle — 234 economies trading with 233 potential partners in 5,111 products over 17 years yields over 4.7 billion data points. To simplify the problem, this paper assumes that relationships (at the reporter-partner-commodity level) that never see realization in all years in the dataset have no economic meaning when estimating transition probabilities, since their state never

[11] Definition **s0** only has the non-existent flow (state **o**) and established product (state X) states.

changes. Including them in the analysis would just inflate the probability for a non-existent flow to remain non-existent.

Therefore estimates of the transition probabilities was carried out with only those bilateral trade relationships, at the 6-digit product level, where there have been exports, resp. imports, in at least one year during 1996–2012, *i.e.* where there has been at least one transition from one state to another. It is done in a naïve way — by creating a list of the states in all adjacent years and then counting the frequencies with which each possible pair of states occurs. No attempt has been made to include the prior history in the relationship, or to estimate what happens at the truncation points (the start and end of the period covered in the dataset).

The estimates of the transition probabilities for both the exports and imports flows at the level of the whole world[12] and for select countries are presented in Appendix B. Except for **s100**, which is again an outlier, several interesting observations can be made for the world-level flows.

First, both the non-existing-flow state (when a product is not traded) and the established-product state are quite stable, both in terms of high probability to remain in the same state and across the different definitions of what is a sample. The probability for a non-existing flow to remain non-existing in the subsequent period is about 86 percent in all cases, and as expected is constant across the definitions. The probability of an established product to remain so varies in a very narrow range, from 74 to 79 percent.

Second, the trading-sample state **s**, whatever the definition except **s100**, is not stable — a flow that is classified as a sample in the current period has a (usually substantially) higher probability to become extinct (go to state **o**) or to become an established product (go to state **X**) in the subsequent period than to remain a sample. This is to be expected — samples by definition are an expression of search activity and over time should either disappear or mature into an established product — but the speed with which this happens may be a little surprising (more than half of sample flows are expected to disappear each year).

Third, samples are typically about two times more likely to disappear than to become established products. The probability of not being exported in the subsequent period is typically in the range of 0.50–0.65, and the probability to mature to an established product is never higher than 0.37. This result can be linked to the findings in Besedeš & Prusa (2006) — since it is trade flows in samples that disappear quickly, the observed low median

[12] Here 'world' is defined as a list of all countries in the dataset, *i.e.* the transition probabilities have been estimated on all available bilateral flows, not on some sort of world aggregate.

duration and high mortality rates in the early years of trade flows is logical. Once a relationship matures into an established product, it is likely to survive substantially longer. The result is also in line with the theoretical model of Rauch & Watson (2003).

Fourth, when using the relative and the composite definitions of what constitutes a trading sample, flows that have just emerged are typically more likely to jump straight to established-product status than to be traded as a sample. The probability to transition from state **o** to state **X** is usually at least an order of magnitude higher than the probability to go from **o** to **s**. This is not the case for the fixed-dollar-value thresholds.

Definition **s100** again is an outlier. Under it non-existent flows and established products are stable, as in the other definitions. The transition probabilities of samples, however, are markedly different — there is a 0.5313 probability that a sample export flow will remain a sample, a 0.3654 probability that the flow will disappear, and a probability of just 0.1033 that an export sample will emerge as an established product. The probabilities in the imports flows are similar. These estimates are at odds with basic economic logic. Samples are supposed to reflect search activity — to either succeed or fail — not to remain stable. This contradiction is an indication that a fixed threshold of $100,000 is too high when working at the bilateral level.[13]

The observations above are for flows at the world level, where there is little expected difference between exports and imports flows — differences would come from differing reporting times and/or reporting requirements (f.o.b. or c.i.f.). Consequently the estimated probabilities for exports and imports are essentially the same. This does not apply at the economy level, as illustrated in the other tables in Appendix B.

Overall, the differences in the results of applying the alternative definitions of trading sample to the same dataset underlines the need to use a definition appropriate to the research question and its theoretical foundation. Additionally, sending samples to a potential partner is a micro phenomenon — it is firms that trade, not economies. Trade flows observed at the economy level are aggregates and may well depend on the size of the reporter and/or the partner economy or have other aggregation specifics when going from micro to the macro level.

[13] If similar results were obtained in the other definitions of trading sample, another explanation would be that samples take longer than one period to either disappear or mature into established products.

2.3. Defining Samples: an Look at Bulgaria

The implications of using different definitions of what constitutes a trading sample are well illustrated when looking at country-level data. Let's assume a hypothetical policymaker is interested in stimulating export diversification.[14] The economic literature offers a range of options, from export promotion (marketing and facilitating the search for partners), to mitigation of transportation costs and domestic barriers. In a setting of scarce public resources, some indication would be needed where to focus the attention, and the estimated transition probabilities offer a starting point.

One could start with the definition that does not allow for samples at all, **s0**, which says that there is a probability of 0.1172 for a new flow to emerge within the next period (Table B.2). The hypothetical policymaker would then be interested in distinguishing established-product flows from search activity. The distinction is important both from the point of view of value of trade and volatility of trade flows. This is where the differences in the definitions of what constitutes a trading sample become apparent. If we take the **s10** definition, we would conclude that the share of new exporting relationships that are able to find partners quickly (*i.e.* jump straight to established-product exports in the subsequent period) is very close to the share of those that need sending out samples — the probability for a new flow to emerge as an established product is 0.0616 vs. 0.0556 to emerge as a sample. Were we to adopt one of the relative definitions, however, this is no longer the case — for example, under **v1**, and with the rest of the relative definitions providing similar conclusions, the probability of a new flow emerging as an established product is more than an order of magnitude higher (0.1138 vs. 0.0034). This latter set of probabilities indicates that likely local businesses face little constraints in their ability to locate partners and quickly establish a relationship. Thus, as a first approximation, the focus would go to measures lowering domestic trading costs rather than trade promotion abroad.

Similarly, looking at the stability of existing established products — the probability for an established-product flow to remain so is about two thirds (between 0.6447 and 0.6611) for all definitions except **s100**. What happens to the other one third — the ones that drop back to samples or disappear altogether — is quite a different story, though, depending on the definition. According to definition **s10**, 26.08 percent disappear, and

[14] Export diversification may happen at the intensive margin (equalizing shares of existing exports) or the extensive margin (new products). This is further discussed in section 3.1. Also, see for example Dennis & Shepherd (2011), Persson (2010) and Feenstra & Ma (2014) on the link between trade facilitation and diversification.

7.81 percent drop from established-product status to trading-sample status. This dynamic is difficult to explain. At the micro level, it is plausible — an exporter's existing flow may be disrupted for whatever reason, and drop to zero. The exporter would naturally start looking for new partners, and will send out samples, so the observed exports flow would drop from established-product level to samples level. At the aggregate level, however, for a flow to drop from established-product level to samples level would require either that there is a single exporter who has lost a previous partner and is now seeking new partners by sending out samples, or that the cutting off of existing established-product relationships happened to all exporters simultaneously (and there are just a few exporters). Even in the simplest case with only two exporters which initially trade at established-product levels, if one were to suffer a disruption of the trade flow, the other exporter's flow would be sufficient to keep the aggregate export level at established-product levels.

The relative definitions of trading sample offer a more plausible story — the probabilities for an established product flow to drop back to samples level rather than to zero is very small, peaking at 2.2 percent in the case of **v5**.

3. Identifying Export Discoveries

This section outlines the approach to identifying product discoveries used in the literature, and describes a proposed extension. Most of the literature focuses on export products only or treats imports as a source of intermediate inputs but the same approach can be applied to import flows and to intra-industry trade (two-way flows) as well. For ease of presentation, the exposition here talks about export discoveries. Application to imports and two-way flows is straightforward – the former is essentially identical to exports and the latter would just require more possible states, *i.e.* would for example have to include states like 'exports at an established-product level, imports at samples level'.

3.1. Extensive and Intensive Margins in International Trade

There is no single, universally accepted definition of extensive and intensive margins in the literature. Besedeš & Prusa (2011) discuss some of the approaches in the context of getting to a definition of what constitutes a 'trade relationship'. Some authors define it at the product level alone (*e.g.* the emergence of a new product in an economy's exports basket, irrespec-

tive of partners). Others define it at the country level, *e.g.* the emergence of a new partner country. The third approach is to go lower, at the country-product level, *i.e.* discoveries to be defined as any of (a) exporting a 'new' product to a 'new' partner, (b) exporting an 'old' product to a 'new' partner or (c) exporting a 'new' product to an 'old' partner. This reporter-partner-commodity level definition is used in this paper. More formally, using the labels from Comtrade, a trade relationship is a triplet (reporter, partner, commodity), for any given year. This definition assumes that there are at least two groups of factors influencing the emergence of a discovery (and trade in general) — domestic factors, which affect the production side (the ability of the reporting economy to manufacture the product and bring it to market) and market factors at the partner economy, which affect the demand for the product.

For a discovery to emerge, we must have both the ability of the reporter to manufacture the product and the existence of demand at the partner economy for this product. The discovery event may be related to any of the steps in the process, from originally developing the product and/or mastering the production process to discovering or creating foreign demand for it (in each individual partner economy) and actually shipping the product to the destination. Thus exporting the same product (which may have been produced for domestic consumption for some time before being exported) for the first time or to a new partner-economy involves some element of discovery. In other words, as Klinger & Lederman (2006) put it, a 'discovery' may happen within the production possibility frontier and does not necessarily imply pushing the production possibility frontier outwards.

3.2. Identification of Export Discoveries: Overview

The identification of an export product discovery usually involves establishing an identification procedure, possibly with a filter for 'suspicious' cases. Klinger & Lederman (2004) and Klinger & Lederman (2011) look at the introduction of new export products as an aspect of economic development and diversification, using data covering a ten-year period (1994–2003). Their identification of an export product discovery (defined as the successful export of a product by an economy that had not exported it before) requires three periods:[15]

1) A baseline period, during which there are no exports of the product. The length of this window is 3 years (1994–1996).

[15] This description is based on Klinger & Lederman (2011). Their earlier working paper uses a more restrictive but fundamentally similar approach — a product is considered to be a discovery if exports were less than $10,000 at the beginning of the period but above $1,000,000 at the end.

2) A second period, during which something happens and the product is 'discovered'. For the purpose of the identification it is not necessary to know what and how caused the discovery to emerge. The length of this window is 5 years (1997–2001). A product can emerge as a discovery in any year within this period.
3) A third window (2002–2003), in which the product is confirmed to be established, *i.e.* it is being exported in values above a predefined threshold (export value of $10,000). Exports below the threshold are considered trading samples. The length of this window is 2 years (2002–2003).

Let's label this approach KL.[16]

Besedeš & Prusa (2006) look at the survival rates of trade flows (in a number of datasets). Even though they do not formally define and analyze discoveries, in the context of their analysis a discovery is simply a product which has not been traded in the previous period — substantially more relaxed that the approach in KL.

Cadot et al. (2011) provide a different classification method, hereafter labeled as CS. A product is identified as a discovery (in the current year) if it has not been exported for the two preceding years, and has been exported in any value, however small, in each of the two subsequent years.[17] Since a discovery can be flagged using 5 years of data, they employ a "moving 5-year sub-sample" to cover the full period of the available data. Importantly, since their research question is different than that of KL, the CS approach does not distinguish between established products and samples. Due to the lack of a minimum threshold and the short length of the period, the definition captures the small-scale activity of entrepreneurs trying to discover new markets. For example, two years of no exports followed by three years of exports at samples level will be counted as a discovery.

[16] It should also be noted that they apply the identification process on data which is disaggregated only along the commodity dimension, *i.e.* they look at the aggregate exports at the HS 6-digit product level. In addition, they employ a filter to identify 'suspicious' cases. Both of these are outside of the scope of the current discussion.

[17] They define discoveries as "lines that were inactive in the country's export trade in the preceding two years but were exported in the following two years (two-years cutoff)" (p.596-597). This formulation is not explicit about whether there must be exports in the *current* year. The analysis in the present paper assumes that the current year's exports must be positive for a flow to be tagged as a discovery. The alternative — tagging a non-existent flow as a discovery — does not seem desirable. This paper uses only the two-year-cutoff version, which according to the authors "strikes a balance between the very conservative definition used by [Klinger & Lederman (2006)] and the very liberal one used by [Besedeš & Prusa (2006)]".

An important issue that both the CS and KL methods do not address is the possibility that the discovery process may be of different lengths for the different products. And, since both use data aggregated across partners, they implicitly assume away any differences with respect to partners.

The lengths of the first and third periods are more or less fixed along products, as it would be difficult to argue that product X will need a different number of years to be verified as an established product than product Y, or that one product will need fewer years for the baseline period test than another. The lengths of these periods are of course debatable — *e.g.*, does it take two years to verify that a product is not being traded, or should we require five or more years? — but once set, the lengths should apply equally to all products. This leaves the middle period, and it would seem sensible to allow that product X may require a longer period of being exported as a sample than product Y, before it becomes established (or fails to). In other words, the identification procedure should allow for a flexible duration of the middle period while the lengths of the first and third periods remain the same. This is even more valid when the analysis is done at the most disaggregated level. Attempting to export widgets to a developed market economy like the USA or Germany seems likely to take much less time than attempting to export the same product to a less-developed economy with poorly functioning markets and more difficult access to information regarding potential partner companies. This argument can be made for many of the traditional determinants of trade — the easier the trade between two partners, for whatever reasons, the shorter the expected duration of the discovery process should be, and *vice versa*. Finally, over time conditions in the reporting and/or partner economy may change sufficiently to impact the speed of discovery for one and the same product. That is, if a product is first 'discovered' and then 'forgotten', rediscovering it may require a different number of years of sending samples.

Let's illustrate the problem using the KL approach. Consider a product which has not been exported for three years (1994–1996 in their case), which then appears as an established product in 1997 and survives for five years (*i.e.* is exported in above-sample-threshold values for each year in 1997–2001). The product then disappears in 2002. In the KL approach, this product will not be flagged as a discovery, as it is not being exported in sufficiently high value in their control window which has fixed starting and ending points (2002 and 2003 respectively). Should this product have been identified as a discovery? Considering the results in Besedeš & Prusa (2006) — that the duration of most trade flows is relatively short (median duration of exporting a product to the US on the order of two to four

years) — a product which has been exported for five years in large values should indeed have been deemed a discovery. It would have been flagged as a discovery using the CS method but it still uses a similar rigid 2+1+2 structure.

The positioning of the three periods (baseline, emergence, established product) within the available data is a second issue which has not seen sufficient discussion in the literature. The KL paper uses ten years of data, and constructs a single set of windows. An obvious question is how to proceed when data for a longer time period is available. There seem to be several possibilities:

- Use several sub-samples (a.k.a. waves, or moving windows) of a length less than the total number of years for which data is available. For example, the dataset used in this paper covers 17 years (1996–2012), and we could construct eight consecutive waves of ten years each (wave 1 covering 1996–2005, wave 2 covering 1997–2006, through wave 8 covering 2003–2012). This approach is used in CS.
- Use longer lengths of the three periods, so that the whole period is covered. It would seem difficult to defend this approach since the length of the windows should motivated by factors independent of the length of the available data. Recall that the window lengths and start/end points are fixed and are the same for all products.
- Use a more flexible definition of the lengths of at least some of the windows, *i.e.* allow for different *lengths* and different *starting points* of the three periods to be applied for different product-partner flows. As argued above, it would be difficult to motivate different *lengths* of the first and third windows but this does not apply to their *starting points*. It would seem that with respect to the middle window, when the magic happens, window length and window starting point should be as flexible as possible to reflect the different specifics of the partner-commodity relationships (*e.g.* a reporter might succeed in establishing a relationship for product **c** with partner $\mathbf{p_1}$ after only a single year of sending samples, but may need three or four years to establish a relationship for the same product with another partner, $\mathbf{p_2}$).

This paper uses the third option as it is the most flexible. The details are presented in the next subsection.

3.3. Identification of Product Discoveries with a Flexible Middle Window

This subsection outlines the procedure for the identification of product discoveries used in this paper. In general terms, the procedure uses the same three windows (baseline, emergence, established product test) but allows for the length of the middle period to differ among reporter-partner-product triplets, and allows it to start anywhere within the available data.[18]

Before proceeding, let us define a simple rule to convert the quantitative exports[19] data into qualitative data. For commodity **c** exported by reporter **r** to partner **p** during period **t** define:

$$F^{c}_{rpt} = \begin{cases} \mathbf{o} & \text{if no trade is reported} \\ \mathbf{s} & \text{if the flow is determined to be a sample} \\ \mathbf{X} & \text{if the flow is determined to be an established product} \end{cases}$$

We can then represent the available data on every reporter-partner-commodity relationship as a vector $F^{c}_{rpt} = (F^{c}_{rp1}, F^{c}_{rp2}, ..., F^{c}_{rpT})$, where **T** is the total number of years in the sample. From a software implementation point of view, the vector is simply a string made up of **T** characters, each of which corresponds to one period in the data. We can then identify patterns in the trade flows by identifying patterns in the text strings.

Using these text strings, we can construct arbitrary flexible definitions of windows patterns. Let's illustrate the approach by formalizing the following discovery definition (hereafter labeled PS), which is broadly similar to those in the papers cited above but allows for a flexible middle window. Define a discovery episode as a case where a product:

1. (*baseline test*) Has not been exported for at least three consecutive years, then
2. (*emergence*) Has been exported in above-zero levels[20] for zero to five consecutive years, inclusive, then

[18] A recognized deficiency of this approach, common with all explored models, is the lack of accounting for censored and/or truncated data. If a discovery episode has begun before the period of the available data, or ends after it, it will not be identified. A potentially fruitful line of further research would be to borrow from the survival analysis literature to augment the identification procedure, esp. when running explanatory models.

[19] This definition deals with one-way trade flows but can easily be extended to incorporate two-way trade flows as well by introducing more codes. Uni-directional flows (exports or imports) require three codes, as listed in the definition above. Bi-directional flows require nine codes, *e.g.* a code corresponding to 'exports of samples and imports of an established product', *etc.*

[20] 'Above-zero levels' means any combination of **s** and **X** that does not contain two consecutive **X**'s. The emergence period cannot contain two consecutive years with established-product level exports **(XX)**, as this would trigger the established-product criterion from Window 3.

3. (*established product test*) Has been exported as an established product for at least two consecutive years immediately following period 2.

That is, we would be looking for substrings **oooXX** (zero-length of window 2), **ooosXX** (window 2 of length one, with samples-level exports), through **ooossssXX** or **oooXsXssXX** anywhere in the F^{c}_{rp} strings. Since our sample covers 17 years, it is possible to have several discovery episodes for a single **(r, p, c)** relationship — we can (and actually do) observe the shortest discovery episode (the five-character string **oooXX**) up to three times.

This definition of 'product discovery' is flexible in the following sense. First, only windows 1 and 3 have fixed lengths. Window 2 does not, and may vary between zero and five years in length, even within the same **(r, p, c)** triplet. Let's illustrate with an example. In our data, we have the string **oooXXoooooXsXXosX** that represents the trade for a single relationship **(r, p, c)** over the 17 years.[21] There are two discovery episodes: (1) over 1996–2000 inclusive (**oooXX**), with a zero-year length of window 2 and (2) over 2003–2009 (**oooXsXX**), with a two-year window 2, **Xs**, where the established-product exports in 2006 are counted as part of the samples window because they are not followed by another year of above-samples exports.

Defined like this, the combined length of all three windows varies between five and ten years, inclusive. While similar to the KL approach, the new element here is that the established-product test (window 3, two consecutive years of established-level exports) may start at any time, rather than being constrained to two fixed years at the end of the sample (or sub-sample if using waves). Second, the definition is flexible in the sense that it does not have a fixed starting point, *i.e.* the first year of window 1 may be any year within our sample, up to **T – 5**.[22] This is important if we think that discoveries will happen with different speeds (different lengths of the samples phase) for different products and/or trading partners, or may have begun at different points in time. It also takes into account that discovered products may disappear before the end of the sample.

Finally, there is some built-in flexibility in the definition of the length of window 3. It is the minimum length of the window that is fixed, and it will

[21] Exports from Brazil to India of HS'1996 product code 860719, "Bogies, bissel-bogies, axles and wheels, and parts thereof :— Other, including parts". Trading sample was defined as a flow with a value of less than 100,000 dollars per year (**s100**).

[22] Since we are looking for a pattern representing at least five years, it cannot be observed if we have fewer than five years of observations left. Similarly, since we require a three-year no exports in the baseline test, there can be no discoveries in the first three years in the data. Truncated and/or censored data is explicitly not accounted for.

expand as long as exports remain at established-product levels. This is not directly useful to the identification procedure here but may be used for example to analyze the survival rates of newly-established products.

3.4. Defining the Year When the Discovery is Made

Whatever the definition of 'discovery', it is a multi-year process, and once a discovery episode is flagged, it must be assigned to a particular year (or years). There seem to be two possibilities, putting the emphasis on different aspects of the discovery process. One is to count a product as a discovery in the year when it first appears as a positive flow (first year of window 2), *i.e.* to focus on when the first attempt is made irrespective of how long it took to reach maturity. The second is to count a flow as a discovery in the first year when it becomes mature, *i.e.* when it started being exported at above-sample levels (the first year of window 3). The longer the 'samples' phase allowed by the identification methodology, the larger the possible discrepancy between the two points.

KL and CS opt for the first approach. In the case of CS the choice is mostly irrelevant — as there is only a single year in window 2, the difference between the two approaches of assigning a discovery to a year is always exactly one year. In the KL approach, there is some uncertainty — the first positive flow may emerge in any of the five years in the fixed window 2.

Using the KL and CS definitions of a discovery episode but assigning it to the first year when the product becomes mature (first year of window3) is of little use. In the CS case, as already discussed, the difference is always exactly one year, and it is not clear what the benefits of shifting the counting from **t** to **t+1** would be. In the KL approach counting discoveries in the first year of window 3 is counterproductive — window 3 always starts at the same fixed year and all discoveries will be assigned to just a single year.

In the PS definition of discovery proposed in this paper, there is always a variable difference between the first year of window 2 and window 3 — from zero (*i.e.* when there is no window 2 at all, the product jumps straight to established product status), to a maximum of five years. In addition, since the start points of the windows are flexible, they would correspond to different years. To remain closer to the KL and CS approach, here a discovery is assigned to the first year of window 2, or if it is of zero length, to the first year of window 3.

3.5. Comparing the Results of the Three Definitions

This subsection discusses the results of applying the three definitions of 'discovery' (a.k.a. patterns[23] to be found in the text strings representing trade) to the Comtrade 1996–2012 dataset.

Two of the three patterns explored here — KL and CS — are intended to mimic the approach in the respective papers. The similarity is only in the way the identification windows are constructed. The present analysis is done at the reporter-partner-commodity-year level, an important difference compared to the original sources, which are done at the level of aggregate exports (*i.e.* reporter-commodity-year). In the KL case, this paper also uses several waves, since the time period covered in the dataset is longer.

Similarly to the discussion of what constitutes a trading sample, the different approaches to defining a discovery yield markedly different results. This is illustrated using three simple comparisons, out of the many that can be constructed on the basis of the individual trade flows.

The first comparison — Table C.1, Table C.2 and Table C.3 in Appendix C — is a simple count of the number of flows tagged as a discovery by each of the three discovery definitions, using each of the nine definitions of trading sample. Since the CS definition does not use samples, the numbers of discoveries found is the same for all definitions of trading sample. It tags as discoveries about 4 percent of the exports flows and a slightly higher number of the imports flows. As expected, the more stringent definition KL tags a substantially lower number of flows as discoveries than CS does, and the more flexible PS definition tags relatively more flows (except for **s10** and **s100**).

The second comparison is a cross-tab of the discovery counts for the same definition of trading sample and the different definitions of discovery, at the individual flow level (Table C.4, Table C.5 and Table C.6). Even if the total number of flows tagged is similar, *different* flows are being tagged by the different definitions. Since this comparison may be too restrictive — it requires that two definitions of discovery agree on the classification of a flow in the same year — Table C.7 and Table C.8 present a third comparison, a cross-tabulation of how many times a reporter-partner-commodity relationship (*i.e.* stripping away the time dimension) has been tagged as a discovery by the different definitions, even if not in the same year.

[23] In terms of implementation, the three verbal descriptions of what constitutes a discovery were translated into regular expressions, which in turn were applied on the text vectors describing trade flows. The regular expressions are as follows: **o{2}[sX]{3,}** for CS, **o{3}.{5}X{2,}** for KL and **(o{3})([sX]{0,5}?)(X{2,})** for PS.

3.6. The Length of Window 2

Since the main feature of the proposed definition of product discovery is the flexibility of the middle window, it only makes sense to exploit this flexibility. A look at the observed length of window 2 (Table D.1 and Table D.2) reveals a surprising result — most episodes of discovery have no samples phase, they jump from non-existence straight to established-product status. The results vary slightly depending on the definition of trading sample chosen, being stronger in the relative definitions and weaker in the fixed-dollar-threshold ones. Of course, if the thresholds for the sample definitions are raised, the no-samples episodes decline in number.

Considering the data is on an annual basis, a samples phase of length zero does not necessarily mean that there is no samples phase at all. Rather, the flow goes through the samples phase and into maturity within the same calendar year. The result is even stronger considering the second-frequent length of the samples phase is one year — something that may easily happen if the end of the calendar year comes before the flow reaches maturity, even if the total length of the samples phase is less than twelve months. An obvious line of future research is to examine the phenomenon using a higher-frequency dataset, *e.g.* using monthly data.

4. Discussion and Conclusion

This paper draws attention to and examines two technical issues in the analysis of new product discoveries that would benefit from further discussion as the current approach is lacking in terms of theoretic foundations and overly restrictive in some important aspects.

First is the definition of what constitutes a trading sample. Nine alternative definitions are examined — three fixed-dollar-value thresholds (the usual approach in the literature), four relative definitions (defining as samples the flows comprising the bottom 1 and 5 percent of the distribution, by value and by physical quantity), and two composite criteria (a flow is considered a sample if it falls in the bottom 1, resp. 5, percent of the distribution by both value and physical quantity). The alternative definitions produce substantially different results thereby making the choice of definition of trading sample an important step in any analysis. Further research is needed in defining what constitutes a 'trading sample', in particular linking the definition to concrete models of search activity and/or production structure. In this respect, the definitions examined in this paper should be seen as illustration of the importance of the issue and a starting point for further discussion rather than an exhaustive examination of possibilities.

Second, the paper discusses a deficiency in some of the empirical definitions of 'new product discovery' used in the literature, namely the fixed structure of the three windows of the identification procedure, and proposes a procedure which uses a middle window (the sending samples phase) of variable length to account for potential differences in the specifics of trade relationships at the reporter-partner-commodity level. The proposed procedure can be tailored to specific research needs. For example, by observing the length of the third window, one can obtain a rough approximation to a more rigorous duration (survival) analysis of discoveries. Or, the length of the middle window can be used to assess the importance of samples. The dynamics of the lengths of these windows can be especially interesting in cases where we have repeated discoveries of the same product in the same reporter-partner pair; or to contrast the differing experiences of a single reporter exporting the same product to several partners. Of course, as has been done in the literature, this definition of discovery can be inverted to tag episodes where a product has been 'forgotten' or dropped, rather than discovered.

An interesting preliminary result, coming from both lines of analysis in this paper, is that the observed length of the samples phase in the vast majority of discovery episodes is zero or one years, indicating that most product discoveries go through the samples phase and emerge as established products within twelve months. The result is observed for both exports and imports flows, and is qualitatively similar across all nine definitions of trading sample used in the paper. It seems also linked to well-known results in the literature (Besedeš & Prusa (2006)). Further discussion and research are needed, however, to determine whether the observed pattern is due to underlying economic logic or just an interesting-but-random pattern that has appeared in a moderately large dataset, or an improper choice of the thresholds for the different definitions.

The proposed definition of product discovery facilitates the process of better understanding the dynamics of innovation, and can serve as a starting point for analyzing the differences across products, trading economy pairs and time. At the same time, it remains just a proof-of-concept without firm theoretic background.

References

Agosin, M.R. & Bravo-Ortega, C., 2009. *The Emergence of New Successful Export Activities in Latin America: The Case of Chile*, Inter-American Development Bank Research Network Working Papers R-552.

Besedeš, T. & Prusa, T.J., 2006. Ins, outs, and the duration of trade. *Canadian Journal of Economics*, 39(1), pp.266–295.

Besedeš, T. & Prusa, T.J., 2011. The role of extensive and intensive margins and export growth. *Journal of Development Economics*, 96(2), pp.371–379.

Cadot, O., Carrère, C. & Strauss-Kahn, V., 2011. Export Diversification: What's behind the Hump? *Review of Economics and Statistics*, 93(2), pp.590–605.

Chandra, V., Li, Y. & Rodarte, I.O., 2007. Commodity Export Diversification in Rwanda - Many Export Discoveries with Little Scaling-Up. In *Rwanda: Toward Sustained Growth and Competitiveness Volume II, Report No. 37860-RW*.

Dennis, A. & Shepherd, B., 2011. Trade facilitation and export diversification. *World Economy*, 34(1), pp.101–122.

Evenett, S.J. & Venables, A.J., 2002. *Export Growth By Developing Countries: Market Entry and Bilateral Trade*, Mimeo.

Feenstra, R.C. & Ma, H., 2014. Trade Facilitation and the Extensive Margin of Exports. *Japanese Economic Review*, 65(2), pp.158–177.

Helpman, E., Melitz, M.J. & Rubinstein, Y., 2008. Estimating Trade Flows: Trading Partners and Trading Volumes. *Quarterly Journal of Economics*, 123(2), pp.441–487.

Hidalgo, C.A. et al., 2007. The product space conditions the development of nations. *Science*, 317(5837), pp.482–7.

Hummels, D. & Klenow, P.J., 2005. The Variety and Quality of a Nation's Exports. *American Economic Review*, 95(3), pp.704–723.

Klinger, B. & Lederman, D., 2004. *Discovery and Development: An Empirical Exploration of New Products*, Washington D.C. World Bank Policy Research Working Paper 3450.

Klinger, B. & Lederman, D., 2006. *Diversification, Innovation, and Imitation inside the Global Technological Frontier*, World Bank Policy Research Working Paper 3983.

Klinger, B. & Lederman, D., 2011. Export discoveries, diversification and barriers to entry. *Economic Systems*, 35(1), pp.64–83.

Persson, M., 2010. *Trade Facilitation and the Extensive Margin*, Stockholm, Sweden., Research Institute of Industrial Economics Working Paper 828.

Rauch, J.E. & Watson, J., 2003. Starting small in an unfamiliar environment. *International Journal of Industrial Organization*, 21(7), pp.1021–1042.

Appendix A. ***Comparing the Definitions of 'Sample'***

Comparison of the results of applying the different definitions of 'sample' to export flows. Each 2x2 block provides the number of cases where two definitions of 'sample' agree or disagree. For example, the bottom-left block in Table A.1 compares **c5** to **s10** and shows that 52,409,956 of the export flows are classified as established products by both the composite criterion **c5** and the **s10** criterion; 995,067 flows are classified by both criteria as samples; in 13,093,147 cases **s10** classified the flow as a sample while **c5** classified it as an established product, and in 384 cases the opposite happened.

Blocks that compare a definition to itself just provide information how many flows the given definition classifies as established products and how many as samples.

Table A.1. **Comparing the Definitions of 'Sample': Exports**

		q1		q5		v1		v5	
		established	**sample**	**established**	**sample**	**established**	**sample**	**established**	**sample**
q1	**established**	65,948,588	0	63,491,960	2,456,628	65,338,014	610,574	62,843,642	3,104,946
	sample	0	549,966	0	549,966	495,013	54,953	338,336	211,630
q5	**established**	63,491,960	0	63,491,960	0	63,062,214	429,746	61,170,835	2,321,125
	sample	2,456,628	549,966	0	3,006,594	2,770,813	235,781	2,011,143	995,451
v1	**established**	65,338,014	495,013	63,062,214	2,770,813	67,546,846	0	64,822,589	2,724,257
	sample	610,574	54,953	429,746	235,781	0	682,876	0	682,876
v5	**established**	62,843,642	338,336	61,170,835	2,011,143	64,822,589	0	64,822,589	0
	sample	3,104,946	211,630	2,321,125	995,451	2,724,257	682,876	0	3,407,133
s10	**established**	52,323,454	86,886	51,859,193	551,147	53,730,006	85	53,724,956	5,135
	sample	13,625,134	463,080	11,632,767	2,455,447	13,816,840	682,791	11,097,633	3,401,998
s100	**established**	26,766,113	6,160	26,743,969	28,304	27,405,620	4	27,405,566	58
	sample	39,182,475	543,806	36,747,991	2,978,290	40,141,226	682,872	37,417,023	3,407,075
c1	**established**	65,948,588	495,013	63,491,960	2,951,641	65,833,027	610,574	63,181,978	3,261,623
	sample	0	54,953	0	54,953	0	54,953	0	54,953
c5	**established**	65,164,767	338,336	63,491,960	2,011,143	65,073,357	429,746	63,181,978	2,321,125
	sample	783,821	211,630	0	995,451	759,670	235,781	0	995,451

		s10		s100		c1		c5	
		established	**sample**	**established**	**sample**	**established**	**sample**	**established**	**sample**
q1	**established**	52,323,454	13,625,134	26,766,113	39,182,475	65,948,588	0	65,164,767	783,821
	sample	86,886	463,080	6,160	543,806	495,013	54,953	338,336	211,630
q5	**established**	51,859,193	11,632,767	26,743,969	36,747,991	63,491,960	0	63,491,960	0
	sample	551,147	2,455,447	28,304	2,978,290	2,951,641	54,953	2,011,143	995,451
v1	**established**	53,730,006	13,816,840	27,405,620	40,141,226	65,833,027	0	65,073,357	759,670
	sample	85	682,791	4	682,872	610,574	54,953	429,746	235,781
v5	**established**	53,724,956	11,097,633	27,405,566	37,417,023	63,181,978	0	63,181,978	0
	sample	5,135	3,401,998	58	3,407,075	3,261,623	54,953	2,321,125	995,451
s10	**established**	53,730,091	0	27,405,624	26,324,467	52,410,299	41	52,409,956	384
	sample	0	14,499,631	0	14,499,631	14,033,302	54,912	13,093,147	995,067
s100	**established**	27,405,624	0	27,405,624	0	26,772,270	3	26,772,270	3
	sample	26,324,467	14,499,631	0	40,824,098	39,671,331	54,950	38,730,833	995,448
c1	**established**	52,410,299	14,033,302	26,772,270	39,671,331	66,443,601	0	65,503,103	940,498
	sample	41	54,912	3	54,950	0	54,953	0	54,953
c5	**established**	52,409,956	13,093,147	26,772,270	38,730,833	65,503,103	0	65,503,103	0
	sample	384	995,067	3	995,448	940,498	54,953	0	995,451

Source: Author's calculations based on UN Comtrade.
Notes: There are a total of 68,229,722 exports flows. Blocks containing quantity-based criteria sum to 66,498,554 as there are 1,731,168 flows with missing quantity data. The **s0** definition discussed in the text is not included in the table as it assumes there are no samples, *i.e.* all flows are classified as established products

Appendix B. ***Transition Matrices by Definition of 'Sample': Exports and Imports***

Each table shows the transition matrix for a given definition of what constitutes a trading sample. Data is presented row-wise, *i.e.* for exports and definition **v1** in Table B.1 below, if a product is not exported in the current year (*i.e.* is in state **o**), for the subsequent period the probability to remain not exported is 0.8552, the probability to become a sample (to transition to state **s**) is 0.0035, and the probability to become an established product (**X**) is 0.1413. Please refer to section 2.2 in the main text for details on the calculation of the probabilities.

The probability for a non-existing flow to remain non-existing (*i.e.* to go from state **o** to state **o**) as expected is constant across the definitions. Non-existing flows will never be tagged as samples or mature products in the definitions used here, and indeed no 'sensible' definition of a sample would do so. The small differences in the estimated transition probabilities are due to missing quantity data in the dataset. All value-based definitions (**v1**, **v5**, **s0**, **s10** and **s100**) have the same estimated probability for the transition from state **o** to state **o**, and so do the definitions which use quantity information (**q1**, **q5**, and **c1** and **c5**).

Table B.1. **Transition Matrices for Different Definitions of 'Trading Sample'**

World

Exports

v1	o	s	X
o	0.8552	0.0035	0.1413
s	0.6199	0.0132	0.3668
X	0.2467	0.0034	0.7499
v5	**o**	**s**	**X**
o	0.8552	0.0168	0.1280
s	0.5980	0.0598	0.3421
X	0.2321	0.0166	0.7513
q1	**o**	**s**	**X**
o	0.8582	0.0028	0.1391
s	0.6283	0.0782	0.2935
X	0.2520	0.0024	0.7456
q5	**o**	**s**	**X**
o	0.8582	0.0136	0.1282
s	0.5623	0.1491	0.2886
X	0.2409	0.0131	0.7460
s0	**o**	**s**	**X**
o	0.8552	n.a.	0.1448
s	n.a.	n.a.	n.a.
X	0.2505	n.a.	0.7495
s10	**o**	**s**	**X**
o	0.8552	0.0623	0.0825
s	0.5166	0.2288	0.2545
X	0.1783	0.0624	0.7594
s100	**o**	**s**	**X**
o	0.8552	0.1255	0.0193
s	0.3654	0.5313	0.1033
X	0.0776	0.1334	0.7890
c1	**o**	**s**	**X**
o	0.8582	0.0003	0.1415
s	0.6656	0.0085	0.3259
X	0.2547	0.0003	0.7451
c5	**o**	**s**	**X**
o	0.8582	0.0049	0.1369
s	0.6153	0.0472	0.3374
X	0.2496	0.0048	0.7455

Imports

v1	o	s	X
o	0.8572	0.0034	0.1394
s	0.6306	0.0126	0.3567
X	0.2520	0.0034	0.7446
v5	**o**	**s**	**X**
o	0.8572	0.0164	0.1264
s	0.6059	0.0570	0.3371
X	0.2374	0.0163	0.7463
q1	**o**	**s**	**X**
o	0.8598	0.0028	0.1374
s	0.6460	0.0736	0.2803
X	0.2577	0.0023	0.7400
q5	**o**	**s**	**X**
o	0.8598	0.0137	0.1265
s	0.5726	0.1444	0.2829
X	0.2462	0.0132	0.7406
s0	**o**	**s**	**X**
o	0.8572	n.a.	0.1428
s	n.a.	n.a.	n.a.
X	0.2558	n.a.	0.7442
s10	**o**	**s**	**X**
o	0.8572	0.0640	0.0787
s	0.5144	0.2424	0.2432
X	0.1791	0.0652	0.7557
s100	**o**	**s**	**X**
o	0.8572	0.1237	0.0190
s	0.3645	0.5402	0.0954
X	0.0827	0.1307	0.7866
c1	**o**	**s**	**X**
o	0.8598	0.0003	0.1399
s	0.6842	0.0081	0.3078
X	0.2605	0.0002	0.7393
c5	**o**	**s**	**X**
o	0.8598	0.0048	0.1354
s	0.6261	0.0444	0.3295
X	0.2555	0.0046	0.7399

Source: Author's calculations based on UN Comtrade.

Table B.2. **Transition Matrices for Different Definitions of 'Trading Sample'**

Bulgaria

Exports

v1	o	s	X
o	0.8828	0.0034	0.1138
s	0.6747	0.0131	0.3122
X	0.3502	0.0046	0.6452
v5	**o**	**s**	**X**
o	0.8828	0.0162	0.1010
s	0.6496	0.0634	0.2870
X	0.3304	0.0222	0.6474
q1	**o**	**s**	**X**
o	0.8830	0.0016	0.1154
s	0.7115	0.0573	0.2313
X	0.3529	0.0015	0.6457
q5	**o**	**s**	**X**
o	0.8830	0.0098	0.1072
s	0.6454	0.1151	0.2395
X	0.3419	0.0103	0.6478
s0	**o**	**s**	**X**
o	0.8828	n.a.	0.1172
s	n.a.	n.a.	n.a.
X	0.3553	n.a.	0.6447
s10	**o**	**s**	**X**
o	0.8828	0.0556	0.0616
s	0.5777	0.2138	0.2085
X	0.2608	0.0781	0.6611
s100	**o**	**s**	**X**
o	0.8828	0.1044	0.0128
s	0.4449	0.4772	0.0779
X	0.1251	0.1652	0.7097
c1	**o**	**s**	**X**
o	0.8830	0.0002	0.1168
s	0.7391	0.0067	0.2542
X	0.3549	0.0003	0.6448
c5	**o**	**s**	**X**
o	0.8830	0.0042	0.1128
s	0.6723	0.0536	0.2741
X	0.3492	0.0050	0.6458

Imports

v1	o	s	X
o	0.8530	0.0039	0.1431
s	0.6034	0.0108	0.3858
X	0.2512	0.0041	0.7448
v5	**o**	**s**	**X**
o	0.8530	0.0190	0.1279
s	0.5674	0.0602	0.3724
X	0.2354	0.0207	0.7440
q1	**o**	**s**	**X**
o	0.8536	0.0023	0.1441
s	0.6550	0.0527	0.2922
X	0.2526	0.0019	0.7455
q5	**o**	**s**	**X**
o	0.8536	0.0145	0.1319
s	0.5628	0.1230	0.3142
X	0.2402	0.0149	0.7449
s0	**o**	**s**	**X**
o	0.8530	n.a.	0.1470
s	n.a.	n.a.	n.a.
X	0.2553	n.a.	0.7447
s10	**o**	**s**	**X**
o	0.8530	0.0742	0.0728
s	0.4736	0.2640	0.2624
X	0.1719	0.0829	0.7452
s100	**o**	**s**	**X**
o	0.8530	0.1338	0.0132
s	0.3286	0.5789	0.0926
X	0.0758	0.1766	0.7476
c1	**o**	**s**	**X**
o	0.8536	0.0003	0.1461
s	0.6477	0.0052	0.3472
X	0.2549	0.0002	0.7449
c5	**o**	**s**	**X**
o	0.8536	0.0057	0.1406
s	0.5931	0.0415	0.3654
X	0.2492	0.0059	0.7449

Source: Author's calculations based on UN Comtrade.

Table B.3. **Transition Matrices for Different Definitions of 'Trading Sample'**

Romania

Exports

v1	o	s	X
o	0.8686	0.0021	0.1293
s	0.6371	0.0111	0.3519
X	0.3232	0.0029	0.6739
v5	**o**	**s**	**X**
o	0.8686	0.0139	0.1175
s	0.6333	0.0473	0.3194
X	0.3068	0.0171	0.6762
q1	**o**	**s**	**X**
o	0.8730	0.0016	0.1254
s	0.6912	0.0679	0.2409
X	0.3223	0.0015	0.6762
q5	**o**	**s**	**X**
o	0.8730	0.0090	0.1180
s	0.6214	0.1205	0.2581
X	0.3129	0.0098	0.6773
s0	**o**	**s**	**X**
o	0.8686	n.a.	0.1314
s	n.a.	n.a.	n.a.
X	0.3260	n.a.	0.6740
s10	**o**	**s**	**X**
o	0.8686	0.0501	0.0812
s	0.5686	0.1918	0.2397
X	0.2524	0.0614	0.6861
s100	**o**	**s**	**X**
o	0.8686	0.1079	0.0235
s	0.4405	0.4537	0.1058
X	0.1321	0.1418	0.7262
c1	**o**	**s**	**X**
o	0.8730	0.0002	0.1268
s	0.7376	0.0000	0.2624
X	0.3243	0.0002	0.6755
c5	**o**	**s**	**X**
o	0.8730	0.0035	0.1235
s	0.6463	0.0393	0.3144
X	0.3199	0.0042	0.6758

Imports

v1	o	s	X
o	0.8282	0.0026	0.1693
s	0.5703	0.0094	0.4203
X	0.2107	0.0026	0.7867
v5	**o**	**s**	**X**
o	0.8282	0.0199	0.1520
s	0.5421	0.0611	0.3968
X	0.1955	0.0175	0.7870
q1	**o**	**s**	**X**
o	0.8365	0.0023	0.1612
s	0.6264	0.0772	0.2964
X	0.2141	0.0017	0.7842
q5	**o**	**s**	**X**
o	0.8365	0.0123	0.1512
s	0.5313	0.1487	0.3200
X	0.2050	0.0118	0.7832
s0	**o**	**s**	**X**
o	0.8282	n.a.	0.1718
s	n.a.	n.a.	n.a.
X	0.2130	n.a.	0.7870
s10	**o**	**s**	**X**
o	0.8282	0.0767	0.0951
s	0.4556	0.2537	0.2908
X	0.1413	0.0700	0.7888
s100	**o**	**s**	**X**
o	0.8282	0.1491	0.0228
s	0.3035	0.5809	0.1156
X	0.0575	0.1494	0.7931
c1	**o**	**s**	**X**
o	0.8365	0.0003	0.1632
s	0.6489	0.0097	0.3414
X	0.2161	0.0002	0.7837
c5	**o**	**s**	**X**
o	0.8365	0.0045	0.1590
s	0.5752	0.0412	0.3836
X	0.2121	0.0043	0.7835

Source: Author's calculations based on UN Comtrade.

Table B.4. **Transition Matrices for Different Definitions of 'Trading Sample'**

Germany

Exports

v1	o	s	X
o	0.8311	0.0018	0.1670
s	0.5531	0.0128	0.4341
X	0.1388	0.0013	0.8599
v5	**o**	**s**	**X**
o	0.8311	0.0234	0.1455
s	0.5344	0.0757	0.3899
X	0.1250	0.0135	0.8615
q1	**o**	**s**	**X**
o	0.8323	0.0011	0.1665
s	0.5417	0.0805	0.3777
X	0.1403	0.0006	0.8591
q5	**o**	**s**	**X**
o	0.8323	0.0137	0.1540
s	0.4701	0.1490	0.3810
X	0.1334	0.0091	0.8574
s0	**o**	**s**	**X**
o	0.8311	n.a.	0.1689
s	n.a.	n.a.	n.a.
X	0.1402	n.a.	0.8598
s10	**o**	**s**	**X**
o	0.8311	0.0827	0.0862
s	0.4236	0.2957	0.2806
X	0.0849	0.0483	0.8668
s100	**o**	**s**	**X**
o	0.8311	0.1488	0.0201
s	0.2512	0.6405	0.1083
X	0.0316	0.0901	0.8783
c1	**o**	**s**	**X**
o	0.8323	0.0001	0.1675
s	0.5882	0.0156	0.3962
X	0.1407	0.0001	0.8592
c5	**o**	**s**	**X**
o	0.8323	0.0062	0.1614
s	0.5342	0.0609	0.4049
X	0.1371	0.0037	0.8592

Imports

v1	o	s	X
o	0.8440	0.0012	0.1549
s	0.6096	0.0096	0.3807
X	0.1688	0.0010	0.8303
v5	**o**	**s**	**X**
o	0.8440	0.0174	0.1386
s	0.5958	0.0532	0.3511
X	0.1555	0.0115	0.8330
q1	**o**	**s**	**X**
o	0.8450	0.0009	0.1541
s	0.5523	0.0806	0.3672
X	0.1696	0.0006	0.8298
q5	**o**	**s**	**X**
o	0.8450	0.0088	0.1463
s	0.5201	0.1180	0.3619
X	0.1642	0.0066	0.8292
s0	**o**	**s**	**X**
o	0.8440	n.a.	0.1560
s	n.a.	n.a.	n.a.
X	0.1699	n.a.	0.8301
s10	**o**	**s**	**X**
o	0.8440	0.0661	0.0899
s	0.5091	0.2129	0.2779
X	0.1128	0.0449	0.8423
s100	**o**	**s**	**X**
o	0.8440	0.1350	0.0210
s	0.3343	0.5373	0.1284
X	0.0387	0.0946	0.8667
c1	**o**	**s**	**X**
o	0.8450	0.0001	0.1549
s	0.5983	0.0042	0.3975
X	0.1700	0.0001	0.8299
c5	**o**	**s**	**X**
o	0.8450	0.0036	0.1514
s	0.5920	0.0425	0.3655
X	0.1673	0.0024	0.8303

Source: Author's calculations based on UN Comtrade.

Table B.5. **Transition Matrices for Different Definitions of 'Trading Sample'**

USA

Exports

v1	o	s	X
o	0.8143	0.0048	0.1809
s	0.5467	0.0178	0.4355
X	0.2101	0.0042	0.7857
v5	**o**	**s**	**X**
o	0.8143	0.0227	0.1630
s	0.5455	0.0599	0.3946
X	0.1969	0.0188	0.7843
q1	**o**	**s**	**X**
o	0.8271	0.0039	0.1690
s	0.5516	0.0552	0.3932
X	0.2339	0.0036	0.7626
q5	**o**	**s**	**X**
o	0.8271	0.0181	0.1548
s	0.5208	0.1187	0.3605
X	0.2233	0.0165	0.7602
s0	**o**	**s**	**X**
o	0.8143	n.a.	0.1857
s	n.a.	n.a.	n.a.
X	0.2135	n.a.	0.7865
s10	**o**	**s**	**X**
o	0.8143	0.0799	0.1058
s	0.5058	0.1912	0.3030
X	0.1490	0.0630	0.7879
s100	**o**	**s**	**X**
o	0.8143	0.1640	0.0217
s	0.3481	0.5194	0.1325
X	0.0551	0.1434	0.8015
c1	**o**	**s**	**X**
o	0.8271	0.0005	0.1724
s	0.5868	0.0115	0.4017
X	0.2364	0.0004	0.7632
c5	**o**	**s**	**X**
o	0.8271	0.0069	0.1661
s	0.5590	0.0459	0.3951
X	0.2316	0.0062	0.7622

Imports

v1	o	s	X
o	0.8389	0.0038	0.1573
s	0.6130	0.0132	0.3738
X	0.1861	0.0026	0.8113
v5	**o**	**s**	**X**
o	0.8389	0.0176	0.1435
s	0.6050	0.0458	0.3493
X	0.1745	0.0117	0.8138
q1	**o**	**s**	**X**
o	0.8501	0.0057	0.1442
s	0.5964	0.0895	0.3141
X	0.2035	0.0040	0.7925
q5	**o**	**s**	**X**
o	0.8501	0.0202	0.1298
s	0.5391	0.1606	0.3004
X	0.1912	0.0154	0.7934
s0	**o**	**s**	**X**
o	0.8389	n.a.	0.1611
s	n.a.	n.a.	n.a.
X	0.1892	n.a.	0.8108
s10	**o**	**s**	**X**
o	0.8389	0.0691	0.0920
s	0.5325	0.1939	0.2736
X	0.1272	0.0466	0.8262
s100	**o**	**s**	**X**
o	0.8389	0.1388	0.0223
s	0.3639	0.5082	0.1279
X	0.0473	0.0948	0.8579
c1	**o**	**s**	**X**
o	0.8501	0.0006	0.1494
s	0.6323	0.0108	0.3568
X	0.2080	0.0004	0.7916
c5	**o**	**s**	**X**
o	0.8501	0.0063	0.1436
s	0.6082	0.0385	0.3533
X	0.2029	0.0047	0.7923

Source: Author's calculations based on UN Comtrade.

Table B.6. **Transition Matrices for Different Definitions of 'Trading Sample'**

China

Exports

v1	o	s	X
o	0.8301	0.0032	0.1667
s	0.5340	0.0123	0.4537
X	0.1642	0.0022	0.8337
v5	**o**	**s**	**X**
o	0.8301	0.0151	0.1548
s	0.5146	0.0469	0.4385
X	0.1554	0.0104	0.8342
q1	**o**	**s**	**X**
o	0.8300	0.0004	0.1695
s	0.5734	0.0605	0.3661
X	0.1682	0.0002	0.8316
q5	**o**	**s**	**X**
o	0.8300	0.0034	0.1666
s	0.5454	0.0835	0.3711
X	0.1660	0.0019	0.8320
s0	**o**	**s**	**X**
o	0.8301	n.a.	0.1699
s	n.a.	n.a.	n.a.
X	0.1665	n.a.	0.8335
s10	**o**	**s**	**X**
o	0.8301	0.0637	0.1062
s	0.4434	0.2007	0.3559
X	0.1185	0.0441	0.8374
s100	**o**	**s**	**X**
o	0.8301	0.1473	0.0226
s	0.2889	0.5464	0.1647
X	0.0415	0.1009	0.8576
c1	**o**	**s**	**X**
o	0.8300	0.0001	0.1699
s	0.5855	0.0024	0.4120
X	0.1685	0.0000	0.8315
c5	**o**	**s**	**X**
o	0.8300	0.0014	0.1686
s	0.5719	0.0277	0.4004
X	0.1674	0.0008	0.8317

Imports

v1	o	s	X
o	0.8432	0.0034	0.1534
s	0.5773	0.0109	0.4119
X	0.1858	0.0027	0.8115
v5	**o**	**s**	**X**
o	0.8432	0.0158	0.1409
s	0.5574	0.0475	0.3951
X	0.1754	0.0124	0.8122
q1	**o**	**s**	**X**
o	0.8435	0.0037	0.1528
s	0.5835	0.0846	0.3319
X	0.1870	0.0024	0.8107
q5	**o**	**s**	**X**
o	0.8435	0.0179	0.1386
s	0.5005	0.1653	0.3342
X	0.1770	0.0129	0.8101
s0	**o**	**s**	**X**
o	0.8432	n.a.	0.1568
s	n.a.	n.a.	n.a.
X	0.1887	n.a.	0.8113
s10	**o**	**s**	**X**
o	0.8432	0.0637	0.0931
s	0.4761	0.2126	0.3113
X	0.1338	0.0501	0.8161
s100	**o**	**s**	**X**
o	0.8432	0.1311	0.0257
s	0.3275	0.5255	0.1470
X	0.0603	0.1073	0.8323
c1	**o**	**s**	**X**
o	0.8435	0.0004	0.1561
s	0.6742	0.0042	0.3216
X	0.1895	0.0002	0.8103
c5	**o**	**s**	**X**
o	0.8435	0.0053	0.1512
s	0.5702	0.0358	0.3940
X	0.1857	0.0038	0.8105

Source: Author's calculations based on UN Comtrade.

Appendix C. ***Comparing the Three Definitions of Discovery***

Table C.1. **Flows Tagged/Not Tagged as a Discovery, by Trading Sample Definition**

CS Definition of Discovery

Exports				Imports			
sample	**discovery**	**number**	**percent**	**sample**	**discovery**	**number**	**percent**
v1	**not discovery**	65,543,498	96.06	**v1**	**not discovery**	66,232,284	95.71
	discovery	2,686,224	3.94		**discovery**	2,971,501	4.29
v5	**not discovery**	65,543,498	96.06	**v5**	**not discovery**	66,232,284	95.71
	discovery	2,686,224	3.94		**discovery**	2,971,501	4.29
q1	**not discovery**	65,443,304	95.92	**q1**	**not discovery**	66,154,061	95.59
	discovery	2,786,418	4.08		**discovery**	3,049,724	4.41
q5	**not discovery**	65,443,304	95.92	**q5**	**not discovery**	66,154,061	95.59
	discovery	2,786,418	4.08		**discovery**	3,049,724	4.41
s0	**not discovery**	65,543,498	96.06	**s0**	**not discovery**	66,232,284	95.71
	discovery	2,686,224	3.94		**discovery**	2,971,501	4.29
s10	**not discovery**	65,543,498	96.06	**s10**	**not discovery**	66,232,284	95.71
	discovery	2,686,224	3.94		**discovery**	2,971,501	4.29
s100	**not discovery**	65,543,498	96.06	**s100**	**not discovery**	66,232,284	95.71
	discovery	2,686,224	3.94		**discovery**	2,971,501	4.29
c1	**not discovery**	65,443,304	95.92	**c1**	**not discovery**	66,154,061	95.59
	discovery	2,786,418	4.08		**discovery**	3,049,724	4.41
c5	**not discovery**	65,443,304	95.92	**c5**	**not discovery**	66,154,061	95.59
	discovery	2,786,418	4.08		**discovery**	3,049,724	4.41

Source: Author's calculations based on UN Comtrade.
Note: Since the CS definition does not use samples, the numbers of discoveries found is the same for all definitions of trading sample. The differences in the values in the table are due to flows with missing quantity data.

Table C.2. **Flows Tagged/Not Tagged as a Discovery, by Trading Sample Definition**

KL Definition of Discovery

Exports				Imports			
sample	**discovery**	**number**	**percent**	**sample**	**discovery**	**number**	**percent**
v1	**not discovery**	67,752,924	99.30	**v1**	**not discovery**	68,687,736	99.25
	discovery	476,798	0.70		**discovery**	516,049	0.75
v5	**not discovery**	67,770,503	99.33	**v5**	**not discovery**	68,704,926	99.28
	discovery	459,219	0.67		**discovery**	498,859	0.72
q1	**not discovery**	67,623,669	99.11	**q1**	**not discovery**	68,571,053	99.09
	discovery	606,053	0.89		**discovery**	632,732	0.91
q5	**not discovery**	67,645,188	99.14	**q5**	**not discovery**	68,593,102	99.12
	discovery	584,534	0.86		**discovery**	610,683	0.88
s0	**not discovery**	67,748,713	99.30	**s0**	**not discovery**	68,683,765	99.25
	discovery	481,009	0.70		**discovery**	520,020	0.75
s10	**not discovery**	67,852,775	99.45	**s10**	**not discovery**	68,795,734	99.41
	discovery	376,947	0.55		**discovery**	408,051	0.59
s100	**not discovery**	68,067,194	99.76	**s100**	**not discovery**	69,025,303	99.74
	discovery	162,528	0.24		**discovery**	178,482	0.26
c1	**not discovery**	67,619,780	99.11	**c1**	**not discovery**	68,566,901	99.08
	discovery	609,942	0.89		**discovery**	636,884	0.92
c5	**not discovery**	67,627,039	99.12	**c5**	**not discovery**	68,574,044	99.09
	discovery	602,683	0.88		**discovery**	629,741	0.91

Source: Author's calculations based on UN Comtrade.

Table C.3. **Flows Tagged/Not Tagged as a Discovery, by Trading Sample Definition**

PS Definition of Discovery

Exports				Imports			
sample	**discovery**	**number**	**percent**	**sample**	**discovery**	**number**	**percent**
v1	**not discovery**	64,792,383	94.96	**v1**	**not discovery**	65,377,441	94.47
	discovery	3,437,339	5.04		**discovery**	3,826,344	5.53
v5	**not discovery**	65,060,531	95.36	**v5**	**not discovery**	65,662,415	94.88
	discovery	3,169,191	4.64		**discovery**	3,541,370	5.12
q1	**not discovery**	64,651,087	94.76	**q1**	**not discovery**	65,276,277	94.32
	discovery	3,578,635	5.24		**discovery**	3,927,508	5.68
q5	**not discovery**	64,880,362	95.09	**q5**	**not discovery**	65,526,228	94.69
	discovery	3,349,360	4.91		**discovery**	3,677,557	5.31
s0	**not discovery**	64,722,760	94.86	**s0**	**not discovery**	65,301,769	94.36
	discovery	3,506,962	5.14		**discovery**	3,902,016	5.64
s10	**not discovery**	65,968,824	96.69	**s10**	**not discovery**	66,720,920	96.41
	discovery	2,260,898	3.31		**discovery**	2,482,865	3.59
s100	**not discovery**	67,455,599	98.87	**s100**	**not discovery**	68,350,002	98.77
	discovery	774,123	1.13		**discovery**	853,783	1.23
c1	**not discovery**	64,601,891	94.68	**c1**	**not discovery**	65,221,910	94.25
	discovery	3,627,831	5.32		**discovery**	3,981,875	5.75
c5	**not discovery**	64,697,183	94.82	**c5**	**not discovery**	65,322,100	94.39
	discovery	3,532,539	5.18		**discovery**	3,881,685	5.61

Source: Author's calculations based on UN Comtrade.

The following three tables provide a cross-tabulation of flows according to whether they have or have not been tagged as a discovery by the respective pair of discovery definitions. For example, Table C.4 compares how the CS and KL definitions stack against each other. For sample definition v1, 65,268,668 flows have been tagged as 'not discoveries' by both the CS and KL definitions. There are 274,830 flows that have been tagged as a discovery by KL, and tagged as 'not discovery' by CS; 2,484,256 flows have been tagged as discoveries by CS and not by KL, and only 201,968 flows have been tagged as a discovery by both definitions.

Table C.4. **Crosstab of Discoveries by Different Discovery Definitions**

			CS vs. KL			
			Exports		Imports	
sample		kl =	not discovery	discovery	not discovery	discovery
v1	cs =	not discovery	65,268,668	274,830	65,957,661	274,623
	cs =	discovery	2,484,256	201,968	2,730,075	241,426
v5	cs =	not discovery	65,283,770	259,728	65,972,373	259,911
	cs =	discovery	2,486,733	199,491	2,732,553	238,948
q1	cs =	not discovery	65,052,954	390,350	65,760,223	393,838
	cs =	discovery	2,570,715	215,703	2,810,830	238,894
q5	cs =	not discovery	65,070,790	372,514	65,778,260	375,801
	cs =	discovery	2,574,398	212,020	2,814,842	234,882
s0	cs =	not discovery	65,264,986	278,512	65,954,302	277,982
	cs =	discovery	2,483,727	202,497	2,729,463	242,038
s10	cs =	not discovery	65,350,402	193,096	66,044,214	188,070
	cs =	discovery	2,502,373	183,851	2,751,520	219,981
s100	cs =	not discovery	65,482,117	61,381	66,174,878	57,406
	cs =	discovery	2,585,077	101,147	2,850,425	121,076
c1	cs =	not discovery	65,049,648	393,656	65,756,735	397,326
	cs =	discovery	2,570,132	216,286	2,810,166	239,558
c5	cs =	not discovery	65,055,742	387,562	65,762,789	391,272
	cs =	discovery	2,571,297	215,121	2,811,255	238,469

Source: Author's calculations based on UN Comtrade.

Table C.5. **Crosstab of Discoveries by Different Discovery Definitions**

			CS vs. PS			
			Exports		Imports	
sample		**ps =**	**not discovery**	**discovery**	**not discovery**	**discovery**
v1	**cs =**	**not discovery**	62,145,774	3,397,724	62,448,796	3,783,488
	cs =	**discovery**	2,646,609	39,615	2,928,645	42,856
v5	**cs =**	**not discovery**	62,551,380	2,992,118	62,884,134	3,348,150
	cs =	**discovery**	2,509,151	177,073	2,778,281	193,220
q1	**cs =**	**not discovery**	61,884,754	3,558,550	62,249,118	3,904,943
	cs =	**discovery**	2,766,333	20,085	3,027,159	22,565
q5	**cs =**	**not discovery**	62,196,851	3,246,453	62,591,849	3,562,212
	cs =	**discovery**	2,683,511	102,907	2,934,379	115,345
s0	**cs =**	**not discovery**	62,036,536	3,506,962	62,330,268	3,902,016
	cs =	**discovery**	2,686,224	0	2,971,501	0
s10	**cs =**	**not discovery**	63,748,959	1,794,539	64,274,013	1,958,271
	cs =	**discovery**	2,219,865	466,359	2,446,907	524,594
s100	**cs =**	**not discovery**	65,126,200	417,298	65,760,728	471,556
	cs =	**discovery**	2,329,399	356,825	2,589,274	382,227
c1	**cs =**	**not discovery**	61,818,016	3,625,288	62,174,902	3,979,159
	cs =	**discovery**	2,783,875	2,543	3,047,008	2,716
c5	**cs =**	**not discovery**	61,958,714	3,484,590	62,323,705	3,830,356
	cs =	**discovery**	2,738,469	47,949	2,998,395	51,329

Source: Author's calculations based on UN Comtrade.
Note: The table provides a cross-tabulation of flows according to whether they have or have not been tagged as a discovery by the respective pair of discovery definitions.

Table C.6. **Crosstab of Discoveries by Different Discovery Definitions**

			KL vs. PS			
			Exports		Imports	
sample		ps =	not discovery	discovery	not discovery	discovery
v1	kl =	not discovery	64,319,185	3,433,739	64,865,276	3,822,460
	kl =	discovery	473,198	3,600	512,165	3,884
v5	kl =	not discovery	64,618,478	3,152,025	65,181,910	3,523,016
	kl =	discovery	442,053	17,166	480,505	18,354
q1	kl =	not discovery	64,046,625	3,577,044	64,645,365	3,925,688
	kl =	discovery	604,462	1,591	630,912	1,820
q5	kl =	not discovery	64,305,521	3,339,667	64,926,293	3,666,809
	kl =	discovery	574,841	9,693	599,935	10,748
s0	kl =	not discovery	64,241,751	3,506,962	64,781,749	3,902,016
	kl =	discovery	481,009	0	520,020	0
s10	kl =	not discovery	65,648,250	2,204,525	66,376,117	2,419,617
	kl =	discovery	320,574	56,373	344,803	63,248
s100	kl =	not discovery	67,354,507	712,687	68,241,214	784,089
	kl =	discovery	101,092	61,436	108,788	69,694
c1	kl =	not discovery	63,992,107	3,627,673	64,585,211	3,981,690
	kl =	discovery	609,784	158	636,699	185
c5	kl =	not discovery	64,098,395	3,528,644	64,696,595	3,877,449
	kl =	discovery	598,788	3,895	625,505	4,236

Source: Author's calculations based on UN Comtrade.
Note: The table provides a cross-tabulation of flows according to whether they have or have not been tagged as a discovery by the respective pair of discovery definitions.

Table C.7. **Exports: Crosstab of Discoveries by Different Discovery Definitions, Ignoring Time**

CS vs. KL

v1	kl = 0	kl = 1
cs = 0	8,975,857	155,615
cs = 1	2,179,655	301,328
cs = 2	82,309	19,710
cs = 3	256	145
v5	**kl = 0**	**kl = 1**
cs = 0	8,986,274	145,198
cs = 1	2,185,636	295,347
cs = 2	83,477	18,542
cs = 3	269	132
q1	**kl = 0**	**kl = 1**
cs = 0	8,864,635	169,872
cs = 1	2,161,421	413,266
cs = 2	82,548	22,764
cs = 3	218	151
q5	**kl = 0**	**kl = 1**
cs = 0	8,875,857	158,650
cs = 1	2,170,514	404,173
cs = 2	83,740	21,572
cs = 3	230	139
s0	**kl = 0**	**kl = 1**
cs = 0	8,973,155	158,317
cs = 1	2,178,464	302,519
cs = 2	82,001	20,018
cs = 3	246	155
s10	**kl = 0**	**kl = 1**
cs = 0	9,025,544	105,928
cs = 1	2,223,506	257,477
cs = 2	88,558	13,461
cs = 3	320	81
s100	**kl = 0**	**kl = 1**
cs = 0	9,097,492	33,980
cs = 1	2,356,436	124,547
cs = 2	98,036	3,983
cs = 3	383	18
c1	**kl = 0**	**kl = 1**
cs = 0	8,862,388	172,119
cs = 1	2,159,989	414,698
cs = 2	82,341	22,971
cs = 3	215	154
c5	**kl = 0**	**kl = 1**
cs = 0	8,866,631	167,876
cs = 1	2,162,568	412,119
cs = 2	82,771	22,541
cs = 3	222	147

CS vs. PS

v1	ps = 0	ps = 1	ps = 2	ps = 3
cs = 0	7,859,618	1,231,091	40,507	256
cs = 1	536,108	1,872,326	72,240	309
cs = 2	21,451	54,731	25,789	48
cs = 3	161	200	40	0
v5	**ps = 0**	**ps = 1**	**ps = 2**	**ps = 3**
cs = 0	8,045,730	1,056,138	29,448	156
cs = 1	589,266	1,832,285	59,220	212
cs = 2	23,723	54,561	23,700	35
cs = 3	173	194	34	0
q1	**ps = 0**	**ps = 1**	**ps = 2**	**ps = 3**
cs = 0	7,742,281	1,249,714	42,215	297
cs = 1	520,347	1,978,199	75,842	299
cs = 2	20,199	57,720	27,345	48
cs = 3	144	184	41	0
q5	**ps = 0**	**ps = 1**	**ps = 2**	**ps = 3**
cs = 0	7,888,249	1,111,973	34,068	217
cs = 1	581,079	1,927,360	66,017	231
cs = 2	22,468	57,466	25,339	39
cs = 3	154	178	37	0
s0	**ps = 0**	**ps = 1**	**ps = 2**	**ps = 3**
cs = 0	7,807,720	1,279,540	43,918	294
cs = 1	526,218	1,878,725	75,710	330
cs = 2	21,026	54,755	26,185	53
cs = 3	158	201	42	0
s10	**ps = 0**	**ps = 1**	**ps = 2**	**ps = 3**
cs = 0	8,544,180	576,820	10,433	39
cs = 1	922,131	1,531,740	27,049	63
cs = 2	39,306	48,555	14,146	12
cs = 3	235	147	19	0
s100	**ps = 0**	**ps = 1**	**ps = 2**	**ps = 3**
cs = 0	9,024,658	105,837	974	3
cs = 1	1,842,043	635,412	3,522	6
cs = 2	80,579	19,072	2,367	1
cs = 3	356	44	1	0
c1	**ps = 0**	**ps = 1**	**ps = 2**	**ps = 3**
cs = 0	7,707,843	1,281,934	44,411	319
cs = 1	510,685	1,985,678	78,001	323
cs = 2	19,879	57,719	27,663	51
cs = 3	141	185	43	0
c5	**ps = 0**	**ps = 1**	**ps = 2**	**ps = 3**
cs = 0	7,776,361	1,217,803	40,062	281
cs = 1	527,152	1,973,775	73,474	286
cs = 2	20,490	57,769	27,008	45
cs = 3	143	184	42	0

PS vs. KL

v1	kl = 0	kl = 1
ps = 0	8,301,417	115,921
ps = 1	2,827,489	330,859
ps = 2	109,101	29,475
ps = 3	70	543
v5	**kl = 0**	**kl = 1**
ps = 0	8,537,439	121,453
ps = 1	2,628,573	314,605
ps = 2	89,592	22,810
ps = 3	52	351
q1	**kl = 0**	**kl = 1**
ps = 0	8,160,483	122,488
ps = 1	2,835,081	450,736
ps = 2	113,195	32,248
ps = 3	63	581
q5	**kl = 0**	**kl = 1**
ps = 0	8,369,056	122,894
ps = 1	2,663,219	433,758
ps = 2	98,021	27,440
ps = 3	45	442
s0	**kl = 0**	**kl = 1**
ps = 0	8,240,635	114,487
ps = 1	2,878,653	334,568
ps = 2	114,509	31,346
ps = 3	69	608
s10	**kl = 0**	**kl = 1**
ps = 0	9,384,692	121,160
ps = 1	1,910,702	246,560
ps = 2	42,522	9,125
ps = 3	12	102
s100	**kl = 0**	**kl = 1**
ps = 0	10,897,449	50,187
ps = 1	649,087	111,278
ps = 2	5,810	1,054
ps = 3	1	9
c1	**kl = 0**	**kl = 1**
ps = 0	8,116,239	122,309
ps = 1	2,871,819	453,697
ps = 2	116,807	33,311
ps = 3	68	625
c5	**kl = 0**	**kl = 1**
ps = 0	8,200,717	123,429
ps = 1	2,801,767	447,764
ps = 2	109,650	30,936
ps = 3	58	554

Source: Author's calculations based on UN Comtrade.

Note: The row and column names show how many times a relationship (reporter-partner-commodity) has been tagged as a discovery by the respective definition in any of the 17 years in the sample. For sample definition **v1**, 8,975,857 relationships have never been tagged as discoveries by either CS or KL; 155,615 relationships have been tagged once by KL and never by CS, *etc.*

Table C.8. **Imports: Crosstab of Discoveries by Different Discovery Definitions, Ignoring Time**

CS vs. KL

v1	kl = 0	kl = 1
cs = 0	9,350,148	158,733
cs = 1	2,414,930	335,951
cs = 2	88,492	21,218
cs = 3	253	147
v5	**kl = 0**	**kl = 1**
cs = 0	9,360,489	148,392
cs = 1	2,420,508	330,373
cs = 2	89,744	19,966
cs = 3	272	128
q1	**kl = 0**	**kl = 1**
cs = 0	9,260,769	170,179
cs = 1	2,388,947	439,556
cs = 2	87,186	22,856
cs = 3	238	141
q5	**kl = 0**	**kl = 1**
cs = 0	9,272,633	158,315
cs = 1	2,397,787	430,716
cs = 2	88,520	21,522
cs = 3	249	130
s0	**kl = 0**	**kl = 1**
cs = 0	9,347,650	161,231
cs = 1	2,413,784	337,097
cs = 2	88,170	21,540
cs = 3	248	152
s10	**kl = 0**	**kl = 1**
cs = 0	9,403,746	105,135
cs = 1	2,462,042	288,839
cs = 2	95,706	14,004
cs = 3	327	73
s100	**kl = 0**	**kl = 1**
cs = 0	9,476,056	32,825
cs = 1	2,609,334	141,547
cs = 2	105,616	4,094
cs = 3	384	16
c1	**kl = 0**	**kl = 1**
cs = 0	9,258,256	172,692
cs = 1	2,387,516	440,987
cs = 2	86,984	23,058
cs = 3	232	147
c5	**kl = 0**	**kl = 1**
cs = 0	9,262,580	168,368
cs = 1	2,389,897	438,606
cs = 2	87,417	22,625
cs = 3	237	142

CS vs. PS

v1	ps = 0	ps = 1	ps = 2	ps = 3
cs = 0	8,136,532	1,328,275	43,732	342
cs = 1	541,390	2,126,233	82,957	301
cs = 2	20,615	61,977	27,073	45
cs = 3	166	197	37	0
v5	**ps = 0**	**ps = 1**	**ps = 2**	**ps = 3**
cs = 0	8,336,229	1,140,330	32,095	227
cs = 1	596,916	2,084,004	69,745	216
cs = 2	22,969	61,741	24,966	34
cs = 3	180	188	32	0
q1	**ps = 0**	**ps = 1**	**ps = 2**	**ps = 3**
cs = 0	8,051,081	1,335,230	44,292	345
cs = 1	528,650	2,216,006	83,542	305
cs = 2	19,538	62,659	27,798	47
cs = 3	157	186	36	0
q5	**ps = 0**	**ps = 1**	**ps = 2**	**ps = 3**
cs = 0	8,212,391	1,182,904	35,399	254
cs = 1	593,294	2,161,897	73,081	231
cs = 2	21,815	62,392	25,791	44
cs = 3	176	171	32	0
s0	**ps = 0**	**ps = 1**	**ps = 2**	**ps = 3**
cs = 0	8,079,955	1,381,092	47,452	382
cs = 1	530,764	2,133,067	86,726	324
cs = 2	20,201	61,961	27,496	52
cs = 3	163	200	37	0
s10	**ps = 0**	**ps = 1**	**ps = 2**	**ps = 3**
cs = 0	8,907,804	590,193	10,827	57
cs = 1	996,117	1,721,714	32,994	56
cs = 2	41,377	53,866	14,456	11
cs = 3	248	138	14	0
s100	**ps = 0**	**ps = 1**	**ps = 2**	**ps = 3**
cs = 0	9,399,231	108,589	1,056	5
cs = 1	2,039,910	705,313	5,655	3
cs = 2	85,702	21,624	2,384	0
cs = 3	359	39	2	0
c1	**ps = 0**	**ps = 1**	**ps = 2**	**ps = 3**
cs = 0	8,012,299	1,371,540	46,735	374
cs = 1	518,485	2,223,883	85,815	320
cs = 2	19,272	62,578	28,143	49
cs = 3	156	187	36	0
c5	**ps = 0**	**ps = 1**	**ps = 2**	**ps = 3**
cs = 0	8,085,327	1,303,071	42,221	329
cs = 1	535,250	2,211,669	81,298	286
cs = 2	19,861	62,599	27,535	47
cs = 3	163	180	36	0

PS vs. KL

v1	kl = 0	kl = 1
ps = 0	8,582,108	116,595
ps = 1	3,148,367	368,315
ps = 2	123,265	30,534
ps = 3	83	605
v5	**kl = 0**	**kl = 1**
ps = 0	8,834,010	122,284
ps = 1	2,934,019	352,244
ps = 2	102,917	23,921
ps = 3	67	410
q1	**kl = 0**	**kl = 1**
ps = 0	8,477,475	121,951
ps = 1	3,137,434	476,647
ps = 2	122,150	33,518
ps = 3	81	616
q5	**kl = 0**	**kl = 1**
ps = 0	8,705,309	122,367
ps = 1	2,948,133	459,231
ps = 2	105,681	28,622
ps = 3	66	463
s0	**kl = 0**	**kl = 1**
ps = 0	8,516,051	115,032
ps = 1	3,204,620	371,700
ps = 2	129,088	32,623
ps = 3	93	665
s10	**kl = 0**	**kl = 1**
ps = 0	9,823,619	121,927
ps = 1	2,089,131	276,780
ps = 2	49,054	9,237
ps = 3	17	107
s100	**kl = 0**	**kl = 1**
ps = 0	11,477,102	48,100
ps = 1	706,351	129,214
ps = 2	7,937	1,160
ps = 3	0	8
c1	**kl = 0**	**kl = 1**
ps = 0	8,428,326	121,886
ps = 1	3,178,489	479,699
ps = 2	126,087	34,642
ps = 3	86	657
c5	**kl = 0**	**kl = 1**
ps = 0	8,517,758	122,843
ps = 1	3,103,588	473,931
ps = 2	118,705	32,385
ps = 3	80	582

Source: Author's calculations based on UN Comtrade.
Note: The row and column names show how many times a relationship (reporter-partner-commodity) has been tagged as a discovery by the respective definition in any of the 17 years in the sample.

DP/99/2015

Determinants of Interest Rate Spreads in Bulgaria

Petar Peshev

Abstract. This scientific paper analyses the factors affecting the interest rate spreads of commercial banks in Bulgaria. The Engle–Granger two-step method (1987) and the bounds testing approach of Pesaran et al. (1999, 2001) are applied. The main findings of the paper are that the interest rate spreads dynamics in Bulgaria in the period of Q1.2004 – Q3.2014 was under the long-term influence of factors such as: economic activity, market concentration, foreign ownership, external liabilities, profit margins, loan to asset ratio, loan to deposit ratio, house prices, inflation, interbank lending rates and stock prices. The short-term determinants of the interest rate spreads include: loan to deposit ratio, foreign ownership, unemployment and market concentration. The global financial crisis and its projection on the Bulgarian economy and the Corporate Commercial Bank's insolvency are among the developments increasing interest rate spreads.

Резюме. Тази научна разработка анализира факторите, влияещи върху лихвените спредове на търговските банки в България. Приложени са двустъпковият метод на *Engle – Granger* (1987) и методът на *Pesaran et al.* (1999, 2001). Основните изводи от изследването са, че динамиката на лихвените спредове в България в периода от първото тримесечие на 2004 г. до третото тримесечие на 2014 г. са под дългосрочното влияние на фактори, като: икономическата активност, пазарната концентрация, чуждестранната собственост, външните задължения, нормата на печалбата, съотношенията „кредити/активи" и „кредити/депозити", цените на жилищата, инфлацията, междубанковите лихвени проценти по кредити и цените на акциите. Краткосрочните детерминанти на лихвените спредове включват: съотношението „кредити/депозити", чуждестранната собственост, безработицата и пазарната концентрация. Световната финансова криза и нейната проекция върху българската икономика, както и неплатежоспособността на Корпоративна търговска банка са сред тенденциите, увеличаващи лихвените спредове.

Petar Peshev, Ph.D. is an assistant professor of Economics at the Economics Department of the University of National and World Economy, p.peshev@unwe.bg.

Contents

Introduction .....

Review of Related Literature .....

Data .....
- Methodology .....
- Correlation Matrix .....
- Unit Root Tests .....

Error Correction Models .....

Results .....
- Co-integration Vectors .....
- Engle and Granger Two-step Approach .....
- ARDL Bounds Testing Procedure .....

Summary and Conclusions .....

References .....

Appendices .....

Introduction

Interest rate spreads have traditionally been a major economic concept for bank centric financial systems, hence attracting analysts' attention. This subject has quite practical dimensions for the Bulgarian economy and the banking system respectively, with net interest income comprising an average of above 70% of net banking income over the last decade, while the related net commissions and fees income add another 21%, and investments' net income has only a minor prominence. In the last decade, the Bulgarian banking system has grown dramatically in size and economic significance, being the main source of economic agents' external debt financing (with domestic credit to GDP ratio hovering between 63 and 71% of GDP in the last five years) and supplying the widely recognized investment/saving opportunities.

Banks in Bulgaria operate under a traditional banking model, in which deposit-lending services to non-financial firms and households represent the core of their daily routines. Net interest income and net fee and commission income are the main sources of banks' income, while income from financial markets operations has a negligible share in the net revenue mix.

It is well known that the banking sector is prone to oligopolistic behavior due to economies of scale and due to the high entry barriers, regulations and supervision. In the Bulgarian banking system, the five biggest banks (by asset value) control over 50% of the market, having the potential to dictate banking service trends.

Over the last eight years, the interest rate spreads of non-financial companies in Bulgaria have been on average 101% higher than the mean values of other EU countries, while households' housing purchases interest rate spreads have been over 90% higher compared to the other EU countries. The divergence has widened in the last couple of years, and the interest rate spreads for Bulgarian non-financial companies and households' housing purchases equal 5.4 percentage points and 4.5 percentage points as of November 2014 respectively.[1] Countries with similar risk and development characteristics (CEE countries), even countries that have experienced a more severe slump in GDP, employment, or countries with severely deteriorated public finances (Greece, Portugal, Hungary, Latvia, *etc*.), which were supported by international financial institutions (IMF, World Bank,

[1] For more detailed information on interest rate spreads in the EU, see Tables 4 and 5 in the Appendices section of the paper.

etc.) in latest financial crisis, have much lower spreads between loans and deposits interest rates.

This study has been motivated by the fact that interest rate spreads in Bulgaria have been relatively high compared to other EU countries for a considerable time. The paper aims to econometrically justify aggregate level factors that affect interest rate spreads in Bulgaria in the short and long term.

Review of Related Literature

Net Interest Margin (NIM) is found to be the prevailing dependent variable in most of the research on the subject; however, some researchers perform an analysis on interest rate spreads.[2] Interest rate spread, derived as the difference between loan rates and deposit rates, is a good approximation for the net interest margin of emerging economies' banking systems, where the main banking activities are deposit-lending operations and interest rates on other types of bank assets and liabilities have a very limited effect on NIM.

The most existing papers on the topic investigate the interest margin determinants using bank level data, aiming to better parameterize the relationship by means of panel econometric techniques, segmenting the results for big and small banks and for the system as a whole.

The widely cited and acknowledged interest rate spreads model of Ho and Saunders (1981) is built upon the assumption that macroeconomic factors are incorporated into the microeconomic variables included into the model. By means of a cross-sectional analysis on US individual commercial banks, the authors infer that the pure spread (the constant value in their regression) and non-interest expenses minus non-interest revenue to assets ratio seem to have the biggest and most positive impact on banks' net interest margin (*ibid.*). The model, which uses bank-specific variables, is prone to many limitations. Its approach implies that banks manage the markup above interest rates on deposits solely on their balance sheet indicators, ignoring the general national and international economic environment which is not yet reflected in *ad hoc* and interim financial statements.

[2] The net interest margin as a ratio between net interest revenues and interest-bearing assets is the dependent variable in the works of Gelos (2006), Claeys and Vander Vennet (2004), Trinugroho *et al.* (2012), Poghosyan (2012), Raharjo *et al.* (2014) and many others. Brock and Suarez (2000), Männasoo (2012), Mihailov (2005, 2014) and other researchers explain the variation of the interest rate spreads determinants.

Macroeconomic variables such as GDP and inflation are among the popular explanatory variables in recent papers. The change of the gross domestic product variable or a proxy of it leads to smaller lending margins (see Mihailov (2005), Gelos (2006), Liebeg and Schwaiger (2007), Gul *et al*. (2011), Chortareas *et al*. (2012) and Poghosyan (2012). Improving economic activity enhances banks' confidence in the economy and tends to stimulate lending at the expense of lower lending rates and lending spreads. However, in the results of Mihailov (2014), a positive relationship between the business climate and the lending spreads is found, despite the prevailing hypothesis that banks tend to decrease lending spreads as the economic conditions improve. The positive relationship can be explained with the higher/lower required returns associated with improving/deteriorating economic conjuncture.

The general price level tends to increase the value of the dependent variable, as found in the works of Mihailov (2005), Horovath (2009), Gul *et al*. (2011), Poghosyan (2012), Raharjo P *et al*. (2014). However, inflation has a negative impact on NIM in the work of Brock and Suarez (2000). Higher general price level leads to higher lending interest rates amid a lower marginal deposit rate increase, resulting in higher lending margins. When deflation forces prevail, deposit interest rates decline faster than interest rates on new and existing loans.

It can be assumed that a combination of bank-specific and general economic activity data better explains the variation in NIM and interest rate spreads. This statement is supported by the results of the most recent studies on the subject.

Operating expenses and a proxy of the indicator (personnel expenses, *etc*.) are among the variables widening the spread between loan and deposit rates (see Brock and Suarez (2000), Mihailov (2005), Gelos (2006), Liebeg and Schwaiger (2007), Fungáčová and Poghosyan (2009), Horovath (2009), Poghosyan (2012), Männasoo (2012), Trinugroho *et al*. (2012), Mihailov (2014), Raharjo *et al*. (2014), Ugur and Erkus (2014). Rising administrative and overall operating expenses motivate banks to compensate through higher net interest margins. Claeys and Vander Vennet (2004), however, found that operating costs decrease the net interest margin, contradicting the prevailing results and supporting the hypothesis that higher operating costs, especially higher personnel expenses, lead to lower NIM because of higher staff efficiency and lower asymmetric information side effects. The managerial efficiency diminishes the dependent variable, according to the paper of Chortareas *et al*. (2012).

Market concentration is frequently found to be a factor for the dynamics of NIM and its proxy interest rate spreads. NIM is in a positive association with market concentration according to the research findings of Demirgüç-Kunt and Huizinga (1999), Gelos (2006), Liebeg and Schwaiger (2007), Deans and Stewart (2012) and Poghosyan (2012). This dependency can be justified by an assumption that banks with a strong market presence widen the lending spread to increase their profitability. Banks in Turkey, however, seem to negatively manage their market share net interest margins (see Ugur and Erkus, 2014). In the scientific paper of Horovath (2009), market share has a negative effect on NIM as well. Authors claim that the tradeoff of gaining additional market share are reduced loan rates, hence a decline in the net interest income (*ibid*.). Market power, measured by the Lerner index, has a strong positive and statistically significant impact on NIM, according to the results of Trinugroho *et al*. (2012).

Net interest margin is positively affected by the foreign ownership of banks in emerging economies (see Demirgüç-Kunt and Huizinga (1999) and Ugur and Erkus (2014). Foreign parent banks require their subsidiaries to increase their NIM, thus increasing the speed of return on their investments. According to the findings of Männasoo (2012) and Trinugroho *et al*. (2012), there is a negative association between the two variables, *i.e*. foreign ownership increases the overall banking system's efficiency and leads to tighter lending spreads. However, state-owned banks are a factor for higher NIM, according to the scientific paper of Trinugroho *et al*. (2012).

Brock and Suarez (2000), using a panel model for Latin American countries, have found that non-performing loans lead to lower NIM, but in the model for Colombia bad loans are in a positive association with the dependent variable. The negative association is also revealed in the research of Fungáčová and Poghosyan (2009), Trinugroho *et al*. (2012) and Mihailov (2014). Poghosyan (2012) has found that a higher non-performing loan to assets ratio leads to higher NIM. It can be expected that aggravating loan quality would lead to higher lending rates; however, banks seem to narrow rates, mostly due to the faster decline of loan rates, eventually supporting borrowers' overall loan servicing capacity.

Capital adequacy (capital to assets and other capital ratios) leads to a higher dependent's value according to Claeys and Vander Vennet (2004), Gelos (2006), Fungáčová and Poghosyan (2009), Chortareas *et al*. (2012), Trinugroho *et al*. (2012), Raharjo P *et al*. (2014), Ugur and Erkus (2014). In contrast, Brock and Suarez (2000), Horovath (2009), Gul *et al*. (2011) and Poghosyan (2012) have found that capital adequacy reduces net interest margins. Both hypotheses have their grounds and reasoning, but may be

valid in different stages of the business cycle. An upward trend in capital ratios leads to higher profits demand, for which net interest income is the main source. Moreover, banks tend to increase their capital position as a result of higher spreads between interest rates received on assets and paid on liabilities. Higher capital ratios may lead to lower lending rates and spreads, because banks have already increased their capital, preparing themselves for improved lending activity, thus being able to observe the ratio between capital and risk-weighted assets.

Liquid assets and a proxy of the indicator lead to smaller lending margins, as banks accumulate less liquid assets in exchange of higher lending activity, supported by lower lending rates and lending margins. In times of economic distress banks are active on the liabilities side in their demand for liquidity, *i.e.* banks narrow the lending spread and collect higher interest rates on deposits, as suggested in the works of Brock and Suarez (2000), Fungáčová and Poghosyan (2009) and Männasoo (2012), while the results of Poghosyan (2012) point to a positive association. Mihailov (2005) finds that reserves as a representation of liquid assets are in a negative correlation to interest rate spread values as well.

According to Horovath (2009), a risk proxy, such as the loan to assets ratio, is in a positive association with NIM in the Czech banking system. A similar dependency is also supported by the results in EU and CEE EU member countries in the work of Claeys and Vander Vennet (2004). Risk aversion (debt to equity ratio) leads to a higher interest margin in the study of Liebeg and Schwaiger (2007).

Many other parameters lead to changes of net interest margins. The change in loans leads to higher NIM (see Gul *et al*. 2011). The ratio of fees to interest earning assets also has a considerable positive effect on the interest rate spread, according to the study of Männasoo (2012). Using a bank sample for EU and CEE EU member countries, Claeys and Vander Vennet (2004) find that short-term interest rates are positively associated with NIM.

Profitability, measured through the return on assets ratio, is in a positive association with net interest margins (see Raharjo P *et al*., 2014).

Williams and Rajaguru (2013) examine banks' net interest income in Australia as a decreasing function of the fees and commission income, with banks compensating the diminishing net interest income by raising fees and commissions.

According to the study of Demirgüç-Kunt and Huizinga (1999), bank assets to GDP ratio is in negative associations with the dependent variable.

Authors also find that indices of credit rights, law and order and corruption are positively related to interest margins in developing countries.

Data

Most of the scientific papers on the subject use a panel econometric technique, combining cross-sectional data from individual banks and macroeconomic variables. In this paper, a different approach has been applied. An aggregate bank-specific and general economic data is used to explain the variation of interest rate spreads. Aggregate data is not only more easily obtainable and less time-consuming to analyze and regress, but would be a better choice for a further study, in which different national economies are covered. The task is accomplished through standard time series regressions, *i.e.* using dynamic error correction models.

The interest rate spread has been calculated as the difference between the weighted average of interest rates on the loans and that on the deposits of non-financial companies and households. The approach used for calculating the interest rate spread is similar to the one used by the World Bank and ECB.[3] The data sources are the publicly available statistical databases of the Bulgarian National Bank and the European Central Bank.

All variables except LOG(YR) are seasonally adjusted using Census X-13 seasonal adjustment software. All variables except some interest rate indicators are in a natural logarithm form.

The analysis is performed with quarterly data, from the beginning of 2004 to the end of the third quarter of 2014. Each series set consists of 43 observations.

Table 1. Variables and Description

Variables	Description
RMFHBS	Interest rate spreads, outstanding amounts at the end of the period – weighted average of commercial banks' loan interest rates minus deposit interest rates (outstanding amounts at the end of the period), Source: BNB, Own calculations;

[3] Interest rate spreads, presented as lending margins in the ECB and WB datasets, are measured as the difference between MFIs' interest rates for new business loans to households and non-financial corporations and a weighted average rate of new business on deposits of households and non-financial corporations. For non-euro area countries, rates for loans and deposits in both euro and national currency are taken into account (see ESRB risk dashboard, 2014, and WB WDI database, 2014).

RMFHBV	Interest rate spreads, business volumes, *i.e.* new business (flows) – weighted average of commercial banks' loan interest rates minus deposit interest rates (new business on loans). Source: BNB, Own calculations;
BDLNS	A ratio between the restructured plus non-performing loans of non-financial companies and households overdue by 90 days or more divided by total loans of non-financial companies plus households. Source: BNB, Own calculations;
CAP	A ratio between commercial banks' common equity and assets. Source: BNB, Own calculations;
EXPTONII	Administrative expenses (personnel included) to net interest income ratio. Source: BNB, Own calculations;
EXTLIAB	External liabilities to assets ratio. Source: ECB, Own calculations;
FOWN	A ratio of foreign-owned local bank assets to total bank asset. Source: BNB, Own calculations;
MS5	Market share of the five biggest banks. Source: BNB, Own calculations;
HHI5	Herfindahl-Hirschman index of the assets of the five biggest banks. Source: BNB, Own calculations;
LNSTOASTS	Loans to non-financial companies and households to assets ratio. Source: ECB, BNB, Own calculations;
LNSTODPTS	Loans to non-financial companies and households to counterpart deposits. Source: ECB, BNB, Own calculations;
ROA	Net income to assets ratio. Source: BNB, Own calculations;
ROE	Net income to equity ratio. Source: BNB, Own calculations;
EONIA	A composite index of 1-day interbank lending interest rates. Source: ECB;
EURBR3M	A composite index of 3-month interbank lending interest rates. Source: ECB;
HPR	House prices index, deflated. Source: NSI, Own calculations;
SFX	SOFIX Bulgarian stock exchange index, deflated. Source: BSE, Own calculations;
UNMPL	Unemployment rate. Source: NSI;
YR	Gross domestic product in 2010 prices. Source: NSI;
PRH	Harmonized consumer price index. Source: NSI;
D1	Dummy variable for the start of the recession in Bulgaria, provoked by the last world financial crisis, starting from the beginning of Q1 2009;
D2	Dummy variable for Corporate Commercial Bank's failure, beginning in June 2014 (Q2 2014).

Methodology

From a methodological perspective, several steps are taken before structuring regression models deciphering the factors affecting interest rate spreads in Bulgaria. First, a correlation analysis is applied for identifying potential regressors and their impact on the dependent variable. Second, unit root tests are performed in levels and, if needed, in first differences. Third, co-integration tests, revealing long-term functional dependencies, are applied. Fourth, error correction models with added co-integrating vector (the error correction term) are built, possessing the ability to discover the factors that explain the short and long-term associations of regressors with the dependent variable. Fifth, results are tested for significance and then commented.

Depending of the order of integration, *i.e.* I(0) or I(1), a Engle and Granger (1987) two-step procedure or Pesaran *et al*. (1999, 2001) procedure can be used. The method of Pesaran *et al*. is more tolerate whether the variables are I(0) or I(1), and is also preferred due to the small size of the sample.

Correlation Matrix

For determining the probable strength of relationships between different pairs of dependent and independent variables, a standard correlation matrix has been used. It has been assumed that for an acceptable correlation coefficient, the p-value for accepting the null hypothesis should be less than 5%.

Correlation analysis is applied for identifying probable relationships between dependent variables and their explanatory variables with corresponding p-values below 5%. These initial and elemental relationships help the process of selecting regressors and researching their potential impact on the explained variable and their significance. Although these individual relationships are not indicative for the causation, the sign and the value of the individual coefficient and for its significance, being part of a regression based on multiple regressors, they are supposed to give cursory relation expectations. Correlation coefficients and their significance can be seen in the Table 2.

Correlation coefficients and their corresponding p-values suggest that interest rate spreads, measured through Ln(RMFHBS), are in a positive (possessing a correlation coefficient above 0.5) and significant correlation with EONIA, EURBR3M, Ln(HPR), Ln(ROA), Ln(ROE) and Ln(SFX). Interest rate spreads Ln(RMFHBV) correlate positively (with correlation coefficient exceeding 0.5) with Ln(EXPTONII).

Ln(RMFHBS) is in a negative correlation, with a coefficient below -0.5, with Ln(BDLNS), Ln(CAP), Ln(PRH) and Ln(UNMPL). Respectively, Ln(RMFHBV) is in a negative correlation of such magnitude with Ln(EXTLIAB), Ln(HHI5), Ln(LNSTOASTS), Ln(LNSTODPTS), Ln(MS5), Ln(PRH) and Ln(YR).

The value of correlation coefficients between Ln(HPR), Ln(UNMPL), Ln(FOWN) and both dependent variables are rather contradictory, *i.e.* each variable diverges in terms of the sign of the coefficient in a pair with Ln(RMFHBS) and Ln(RMFHBV), which is hard to be interpreted solely in terms of correlation.

While many correlation coefficients are statistically significant, their meaning is between -0.5 and 0.5. They are not ignored in the regression, but expected coefficient signs are challenged more.

Table 2. Correlation Matrix and p-values

VARIABLES	LN(RMFHBS)	LN(RMFHBV)	VARIABLES	LN(RMFHBS)	LN(RMFHBV)
LN(BDLNS)	-0.68	-0.48	Ln(ROA)	0.79	0.46
p-value	***0.00***	***0.00***	*p-value*	***0.00***	***0.00***
LN(CAP)	-0.69	-0.42	Ln(ROE)	0.80	0.47
p-value	***0.00***	***0.01***	*p-value*	***0.00***	***0.00***
LN(EXPTONII)	0.00	0.82	EONIA	0.87	0.27
p-value	*0.99*	***0.00***	*p-value*	***0.00***	*0.08*
LN(EXTLIAB)	0.01	-0.69	EURBR3M	0.85	0.15
p-value	*0.93*	***0.00***	*p-value*	***0.00***	*0.34*
LN(FOWN)	0.42	-0.28	LN(HPR)	0.59	-0.29
p-value	***0.00***	***0.07***	*p-value*	***0.00***	***0.06***
LN(MS5)	0.23	-0.57	LN(SFX)	0.79	0.49
p-value	*0.14*	***0.00***	*p-value*	***0.00***	***0.00***
LN(HHI5)	0.23	-0.53	LN(UNMPL)	-0.52	0.38
p-value	*0.13*	***0.00***	*p-value*	***0.00***	***0.01***
LN(LNSTOASTS)	-0.32	-0.86	LN(PRH)	-0.58	-0.74
p-value	***0.04***	***0.00***	*p-value*	***0.00***	***0.00***
LN(LNSTODPTS)	-0.21	-0.83	LN(YR)	-0.13	-0.74
p-value	*0.18*	***0.00***	*p-value*	***0.40***	***0.00***

Note: Numbers in bold mark the variables with p-values lower than 0.05.

Unit Root Tests

Results from Augmented Dickey-Fuller (ADF) and Philip-Perron (PP) unit root tests reveal that all of the variables are stationary in first differences I(1) except Ln(LNSTOASTSECB), Ln(PRH), Ln (ROA) and Ln(BDLNS), which are level stationary under one of the three options (intercept, intercept

and trend, no trend and intercept) and first difference stationary in the other two options. In some instances, ADF and PP point to slightly different results; nevertheless, as a whole, results from both type of unit root tests support each other.

For example, the ADF unit root test rejects the null hypothesis in first differences for the intercept option for Ln(CAP), while neither the level result, nor the two of the three options of the first differences ADF test reject the null. The PP test, however, strongly rejects the hypothesis of unit roots presence in the first differences for the same variable.

It can be summarized that variables don't have unit roots in their first differences; however, there are a few exceptions for level stationary variables under one of the three assumptions: intercept, intercept and trend, no trend and no intercept.

Error Correction Models

In this paper, a standard time series econometric modeling has been applied, since most of the level data failed to reject the null hypothesis of having unit root in levels. Error correction models using the two-step approach of Engle and Granger (1987) and the ARDL bounds testing approach of Pesaran and Shin (1999, 2001) have been applied.[4] The purpose of using both error correction methods is not only for the results to support each other, but also that some of the variables may have level stationary characteristics despite the fact that unit root tests point to an overall first difference data rejection of the null hypothesis. For the purpose of parameterizing level stationary, *i.e.* I(0) and first difference stationary data, *i.e.* I(1), the Pesaran *et al.* (1999, 2001) bounds testing approach is more suitable. However, a combination of I(0) and I(1) data may be integrated to lead to the integration of order I(1).[5]

Non-stationary time series require special methods for analysis. Dynamic models with error correction term may be the most appropriate in this case. Error correction models traditionally have the following form:

$$\Delta Y_t = \alpha + \sum_{m=1}^{n} \beta.\Delta Y_{t-m} + \sum_{m=0}^{n} \gamma.\Delta X_{it-m} + \delta.ECT_{t-1} + \zeta.Trend + \eta.Exogs + \varepsilon_{t,} \quad (1)$$

$$ECT_{t-1} = \nu_{t-1} \Rightarrow Y_{t-1} = \alpha + \beta.X_{it-1} + \nu_{t-1} \quad (2)$$

[4] Sheriff and Amoako (2014) use an aggregate data ARDL bounds testing technique on analyzing interest rate spreads in Ghana. English (2002) uses Engle and Granger two-step method on aggregate data for analyzing net interest margins and interest rates on assets and liabilities separately.

[5] See Engle and Granger (1987) and The Royal Swedish Academy of Sciences (2003).

Table 3. ADF and PP Unit Root Test Results (p-values)

	Augmented Dickey-Fuller test						Philip-Perron test					
	Levels			First differences			Levels			First differences		
Series	intercept	intercept and trend	no trend and no intercept	intercept	intercept and trend	no trend and no intercept	intercept	intercept and trend	no trend and no intercept	intercept	intercept and trend	no trend and no intercept
Ln(RMFHBS)	0.32	0.20	0.78	**0.00**	**0.00**	**0.00**	0.30	0.48	0.72	**0.00**	**0.00**	**0.00**
Ln(RMFHBV)	0.14	0.91	0.86	**0.00**	**0.00**	**0.00**	0.11	0.77	0.85	**0.00**	**0.00**	**0.00**
Ln(BDLNS)	0.94	0.45	**0.04**	0.09	0.28	**0.02**	0.89	0.46	0.12	**0.00**	**0.00**	**0.00**
Ln(CAP)	0.30	0.15	0.64	**0.00**	0.58	0.17	0.65	0.19	0.64	**0.00**	**0.00**	**0.00**
Ln(EXPTONII)	0.16	0.76	0.77	**0.00**	**0.00**	**0.00**	0.10	0.50	0.72	**0.00**	**0.00**	**0.00**
Ln(EXTLIAB)	0.41	0.35	0.71	**0.04**	**0.00**	**0.00**	0.12	0.35	0.43	**0.00**	**0.00**	**0.00**
Ln(FOWN)	0.53	0.66	0.56	**0.01**	**0.05**	**0.00**	0.53	0.65	0.56	**0.03**	0.22	**0.00**
Ln(MS5)	0.47	0.79	0.60	**0.00**	**0.00**	**0.00**	0.39	0.71	0.60	**0.00**	**0.00**	**0.00**
Ln(HHI5)	0.36	0.70	0.66	**0.00**	**0.00**	**0.00**	0.27	0.58	0.66	**0.00**	**0.00**	**0.00**
Ln(LNSTOASTS)	**0.00**	0.19	0.42	**0.00**	**0.00**	**0.00**	0.40	0.93	0.38	**0.00**	**0.00**	**0.00**
Ln(LNSTODPTS)	0.43	0.88	0.11	**0.03**	**0.05**	**0.00**	0.45	0.93	0.12	**0.04**	**0.05**	**0.00**
LOG(ROA)	0.62	0.64	0.88	**0.00**	**0.00**	**0.00**	0.29	**0.02**	0.77	**0.00**	**0.00**	**0.00**
LOG(ROE)	0.73	0.78	0.86	**0.00**	**0.00**	**0.00**	0.52	0.18	0.75	**0.00**	**0.00**	**0.00**
EONIA	0.49	0.24	0.16	**0.04**	0.14	**0.00**	0.73	0.61	0.24	**0.03**	0.11	**0.00**
EURBR3M	0.38	0.19	0.16	0.07	0.21	**0.01**	0.71	0.62	0.27	0.06	0.19	**0.01**
Ln(HPR)	0.20	0.23	0.55	0.19	0.62	**0.02**	0.34	0.17	0.76	0.16	0.54	**0.02**
Ln(SFX)	0.46	0.41	0.53	**0.00**	**0.02**	**0.00**	0.61	0.57	0.57	**0.00**	**0.02**	**0.00**
Ln(UNMPL)	0.52	0.56	0.60	**0.05**	0.13	**0.00**	0.54	0.66	0.66	**0.06**	0.13	**0.00**
Ln(PRH)	**0.01**	1.00	0.99	**0.01**	**0.00**	**0.00**	0.08	1.00	1.00	**0.00**	**0.00**	**0.00**
Ln(YR)	0.09	0.62	0.98	**0.00**	**0.00**	**0.00**	0.11	0.59	0.97	**0.00**	**0.00**	**0.00**

Note: The table presents the p-values for accepting the null hypothesis of presence of a unit root under each assumption. A 5% threshold is applied for rejecting the null. Bold values reject the null at 0.05 or lower.

Where: α – constant, $\Delta \mathbf{Y}_t$ – first differences of the dependent variable, $\Delta \mathbf{Y}_{t-m}$ – lagged first differences of the dependent variable, with the first lag being the smallest, $\Delta \mathbf{X}_{it-n}$ – lagged first differences of the ith explanatory variable (with the zero lag being the smallest), $\mathbf{ECT}_{t-1}$ – the error-correction term (the coefficient of ECT should have a negative sign, representing the error correction mechanism) is the residual of the OLS regression of level variables, and represent the long-run associations, **Trend** – deterministic trend (if existing), **Exogs** – exogenous variables (dummy variable and other, if existing), ε_t – residual of the dynamic model, ν_{t-1} – white noise/the residual of the OLS regression in levels.

Under the Engle and Granger (1987) approach, the residual of the co-integrating equation (regressing levels of variables) revealed in equation (2) is tested for the presence of unit roots, using the Augmented Dickey-Fuller test on levels, but applying the more restrictive critical values provided by Davidson and MacKinnon (1993). Failing to accept the null hypothesis of ADF test validates the long-term relationship, *i.e.* a co-integrating vector exists, and a dynamic model like the one presented in equation (1) can be structured.

The ARDL dynamic models can be presented either in the form of an equation (1), or in the following functional schematic form:

$$\Delta Y_t = \alpha + \sum_{m=1}^{n} \beta.\Delta Y_{t-m} + \sum_{m=0}^{n} \gamma.\Delta X_{it-m} + Y_{t-1} - \alpha_y - \beta_y.X_{it-1} + \varepsilon_t \quad (3)$$

Using the bounds testing procedure developed by Pesaran *et al*. (1999, 2001), a dynamic model with differenced and level data is constructed. Coefficients of level data are tested for whether they are equal to zero, using a standard Wald-test procedure. Testing for co-integration is performed using the asymptotic critical value bounds for the F-statistic presented in the papers of Pesaran *et al*. (1999, 2001).

For selecting the number of lags in each model under the ARDL bounds testing procedure, an EViews Add-in has been employed (see Yashar, 2014). Low values of Schwarz information criteria and Akaike information criteria and high F-statistic on the coefficients' Wald tests are the three main factors for selecting the lag length.

Results

Using both the co-integration and error correction approaches, nine different error correction models have been built and investigated in the analysis. Results of the two models with highest explanatory power have

first been analyzed, *i.e.* model A and model B, followed by an analysis of the results of models C to I.

Models A and B best suit the Engle and Granger approach and the ARDL bounds testing procedure. Model B has the highest explanatory power in terms of R^2 adjusted, residual test results, lowest information criteria (Akaike and Schwarz info criteria) and T-stats of regressors' coefficients, while Model A better fits the ARDL bounds testing approach under these assumptions. Results from both models under both approaches seem to validate each other, which of course does not negate the results of models C to I, which yield valuable information. Models A and B would be the first to comment on in each of the subsections below, while findings from the rest of the models will be summarized at the end of the two subsections, *i.e.* the Engle and Granger two-step approach section and the ARDL bounds testing procedure sections respectively.

Co-integration Vectors

The same co-integrating vector is implemented in each model (*e.g.* A, B, C, I) under the Engle and Granger two-step method and the ARDL bounds testing procedures. All co-integration equations can be found at Table 6 in the Appendices section.

Models A and B

First, the co-integrating vectors of models A and B are discussed, followed by the analysts for the rest. Results reveal that factors that have a serious impact on interest rate spreads in Bulgaria are: external bank liabilities, non-performing loans, common equity, market concentration, interbank borrowing rates, house prices, inflation, and to a lesser extent, the stock market.

External liabilities in the case of Bulgaria are mostly due to foreign parent banks. The co-integrating vectors reveal that external bank indebtedness widens the interest rate spreads, and *vice versa*. Repayment of banks' external liabilities urges banks to accumulate buffers *via* increasing the lending margin. Parent banks also insist on higher returns on their receivables, thus making their subsidiaries widen the spread between lending and deposit rates.

The non-performing loan to gross loan ratio leads to higher interest rate spreads. Higher bad loans ratios stimulate the banks to accumulate buffers for covering loan write-downs and for supporting capital adequacy and liquidity.

The equity to assets ratio contributes to higher interest rate spread, *i.e.* a positive relationship exists. Higher capital requirements and buffers are financed through higher interest rate spreads and through other sources (hybrid debt-capital instruments issuance, common and preferred stock issuance, *etc.*), but deposit-lending activity is the core business of the Bulgarian banks.

In the long run, market share indicators are in a positive association with interest rate spreads, confirming the hypothesis that in less competitive markets banks naturally require higher interest rate spreads, aiming at higher returns on equity and assets. Market competition promotes lower lending spreads in the long run.[6]

Higher global interbank lending rates are a factor for higher interest rate spreads. Interbank lending rates are one of the main elements of bank financing sources. In times of economic distress, banks tend to become more risk-averse in lending to each other, *i.e.* interbank lending rates are considered as proxy for risk aversion, resulting in a positive change of lending rates. In favorable economic conditions, interest rates in the economy tend to increase, including interbank lending rates and banks demand higher returns, for which wider lending rates are an option.

Rising house prices lead to more expensive collaterals and boost banks' confidence, which translates into a higher willingness to borrow and lend, resulting in lower interest rate spreads.

Price levels have a negative impact on interest rate spreads. It can be assumed that higher inflation makes customers demand higher deposit interest rates. However, they cannot be fully transferred into higher lending rates, *i.e.* interest rate spreads decrease.

The Bulgarian stock exchange benchmark SOFIX has been included in the regression and reveals that stock prices have a positive effect on the value of interest rate spreads, despite the low coefficient. Rising stock prices signal that booming economic conditions are underway or are expected to come, making banks raise interest rate spreads to boost profitability.

Stock prices have a small effect on interest rate spreads, which however contradicts the effect of property prices. One would expect property prices and stock prices, depicting the wealth effect, to have at least a parallel impact on the dependent variable. Property prices have a bigger effect on the economy, because of the high-ownership rates, supported also

[6] Mihailov (2014) does not find any significant interrelation between lending spreads and market concentration in Bulgaria.

by the low market cap to GDP ratio and the low penetration of financial instrument investments in the country.

Models C to I

In model G, the administrative and personnel expense to net interest income ratio is in a positive association with dependent variable, as would be expected. Administrative inefficiency is compensated through wider interest rate spreads, supporting the results of the panel study on Bulgarian banks performed by Mihailov (2005, 2014). However, in Model C, the administrative and personnel expenses compared to the net interest income lead to lower interest rate spreads, contradicting the results in model G. The results in model C can be interpreted with the assumption that banks are inclined to expand lending at the expense of lowering interest rate spreads. This new lending requires the involvement of further resources (personnel, office premises, equipment, *etc.*). The long-term results for the administrative expense ratio are ambiguous.

The external liabilities in models C, D, G and H are in a negative association with interest rate spreads, contradicting the findings in model B. It can be assumed that the funds supplied mostly by parent banks lower bank's fund costs and increase reserves and liquidity, allowing banks to lower rates on the loans lent, following a strategy for loans' market share increase and/or for attracting better-quality borrowers with better conditions on loans. These long-term results seem to negate the results derived in model B. Often short-term results oppose long-term results, as can be seen in the dynamic models under both approaches, *i.e.* a single factor can be positively associated with the dependent variable in the short term and negatively in the long term.

Models C and H reveal that higher net income ratios (ROE and ROA) prompt banks to expand the spread between loan and deposit rates, demanding bigger returns on assets and equity, as net interest income is the pillar of Bulgarian banks' profitability. These findings are consistent with the results of Raharjo P *et al.* (2014).

Non-performing loans in models C to I negate the results of Model A by having negative coefficients in the long term. It would be expected that bad loans lead to higher lending margins for accumulating bad loan buffers. Negative coefficients may be appropriate under the hypothesis that in the long term banks try to attract high-quality loans through offering better loan conditions (lowering interest rates, for example). A similar negative relation is found to be viable in the panel study of Mihailov (2014) as well.

In the long term, when banks increase the loans *versus* the deposits of non-financial companies and households, they amplify their risk tolerance, thus requiring a smaller interest rate spread (see models E, F and I). The loan to deposit ratio may be considered a risk proxy, being in negative relationship with risk aversion. The lending spread is in positive relationship with risk aversion. These results are in contrast to the positive correlation between the interest margins and the loan to deposit ratio in the research of Trinugroho *et al*. (2012).

The loan to assets ratio leads to higher lending margins, signaling that banks are becoming more risk-averse with a higher weight of loans in their balances. The positive association is comparable to the interdependencies derived from Horovath (2009) and Claeys and Vander Vennet (2004).

Long-term results for foreign ownerships are inconclusive, with a positive coefficient for the parameter in model G and negative in H and I.

Bank concentrations ratios (Market share and HHI) in models E and I confirm the results of model A, leading to wider interest rate spreads, like in the findings of Demirgüç-Kunt and Huizinga (1999), Gelos (2006), Liebeg and Schwaiger (2007), Deans and Stewart (2012) and Poghosyan (2012).

Interbank borrowing rates in models D and F confirm the results derived in models A and B, resulting in interest rate spreads, yielding findings similar to those of Claeys and Vander Vennet (2004).

Lower unemployment rates in the country boosts banks' confidence, leading to lower lending-deposit rate differential (see models D, E and F). The general economic activity, measured through GDP, is a strong factor with a negative meaning for the dependent variable. Higher GDP translates in lower interest rate spreads (see models F to I), confirming the findings of Mihailov (2005), Gelos (2006), Liebeg and Schwaiger (2007), Gul *et al*. (2011), Chortareas *et al*. (2012) and Poghosyan (2012).

A long-term trend is found to be significant in models B, D and F.

Engle and Granger Two-step Approach

Models A and B

The error correction terms are statistically significant and have a highly negative impact on the model, representing the error correction mechanism, *i.e.* reducing the disequilibrium in the dynamic model with each subsequent quarter (since quarterly data is analyzed). Dynamic models structured under the Engle and Granger two-step approach can be found

in the Appendices section in Table 7, while co-integrating vectors (ECT) are presented in Table 6 in the same section.

The interest rate spreads are in a positive and statistically significant relation to their lagged values.

It is common variables in the dynamic model that reveal short-term relations to have negative/positive coefficients and the opposite dependencies to develop in the long term.[7] Banks, as economic agents in general, have different behavior reasoning in the long and short term.

The first and second lags of the first differences of market share of biggest five banks by assets have a negative effect on regressand, while in the long term bigger market share results in larger spreads. It can be interpreted that in the short term banks are concerned with expanding their market share, which stimulates them to be more aggressive in terms of attracting deposits and new lending, thus maintaining or increasing their market share.

External bank liabilities, due mostly to parent banks, lead to larger interest rate spreads in the short term as well as in the long term, suggesting that banks through wider interest rate spreads secure returns in demand by foreign creditors by accumulating funds for their repayment.

Current changes of international interest rates on interbank loans have a positive effect on the dependent variable, and so do the third and fourth lags. The first lag of the regressor, however, has a diminishing effect on the dependent variable. In general, short-term results, presented in the dynamic models, confirm the co-integration findings. It can be summarized that interbank interest rate spreads can be considered as a risk proxy having a multiplying effect on the dependent variable.

House prices with a zero lag lead to wider interest rate spreads, while the first lag of the first differences has a negative effect on interest rate spreads. The coefficients of the house prices variable seem to negate each other with prevalence of the zero lag coefficient value, hence the long-term results should be considered concerning this regressor.

The recession dummy has a positive and statistically significant impact on interest rate spreads, and so does the dummy for Corporate Commercial Bank's turmoil.

The presented models pass the residual tests and have good explanatory power in terms of adjusted R^2 and F-stat. Testing for homoscedasticity is accomplished by using a Breusch-Pagan-Godfrey test and, if needed,

[7] See Pesaran *et al.* (1999, 2001).

White, ARCH and other tests are utilized. For accepting the null hypothesis of no serial correlation, a Breusch-Godfrey Serial Correlation LM Test is applied and, if needed, further investigation and confirmation is coupled with the correlogram of residuals, Q-stat probability, and a correlogram of squared residuals tests. A Jarque-Bera test is performed for normality of the distribution.

Models C to I

The error correction mechanism functions in models C to I as well, *i.e.* error correction terms have negative and statistically significant coefficients.

Interest rate spreads are in a positive association with their lagged first differences, confirming the results of models A and B.

Administrative and personnel expenses expressed as a ratio to net interest income lead to lower interest rate spreads in the short term, supported by the results of models C and I, with the first lag of Model C staying unconfirmed by other models and opposing the long-term results. It can be assumed that banks employ more resources in the short term, aiming to increase the size of their loan portfolios. Banks are in a position to expand lending at the expense of lowering interest rate spreads. New lendings require the involvement of new resources.

In model C, the return on assets ratio leads to smaller values for the dependent variable, in contrast to the long-term results. It can be assumed that lower profitability margins motivate banks to widen the lending spread, thus achieving better income results in the future and offsetting lower returns in the past. Higher return ratios, however, lead to lower interest rate spreads in the near term, with banks aiming to expand credit and deposit activity at the expense of lower lending margins, probably targeting bigger market share.

In the short term, foreign ownership leads to higher interest rate spreads. Banks are required by their foreign parent institution to achieve higher net and operating margins, thus being able to recover their investment according to their projections (see model H).

In contrast to the short-term results in model A, the dependent variable is in negative association with external liabilities (see models C, G and H). When banks acquire abundant funds by their foreign owners, they tighten the spread between interest rates on loans and deposits.

In the short term, the capital adequacy proxy leads to smaller interest rate spreads in its third lag and to a positive dependent value in its fourth lag, *i.e.* the results are inconclusive.

Unemployment in the short term leads to higher interest rate spreads. Deterioration in the labor market results in higher mark-ups on loans.

In dynamic models, the coefficients of first differences of GDP are inconclusive. The zero lag of the regressor has a negative effect on the regressand (see models D, F, G, H and I), while the third lag has a positive effect on interest rate spreads (see models G, H and I). It can be assumed that current GDP stimulates banks in requiring larger interest rate spreads, either because banks expect the situation to change for worse, *i.e.* being rational, or because banks require higher returns as economic conditions improve. However, the third lag of the regressor leads to a smaller value of the dependent, implying that GDP growth in the near past boosted banks' confidence and they increased lending activity at the expense of lowering interest rate spreads. A GDP decline leads to wider interest rate spreads, *i.e.* banks accumulate buffers, expecting balance sheet deterioration.

Growing consumer prices lead to smaller interest rate spreads, while their fourth lag values are in positive association with interest rate spreads (see model E). This contradiction can often be seen in dynamic error correction models when zero lag and other lags are employed, *i.e.* the same variables have positive/negative coefficients in the zero lag, while other lags' (1 to n) coefficients have negative/positive values. Banks behave differently not only in their long-term and short-term strategies, *e.g.* in the short term they may be differently motivated by the same factors.

The logic revealed in models A and B regarding the dummy variables also applies to models C, D, E and F.

Models C to I have a good explanatory power in terms of R^2 adjusted coefficients and residual stability; however, models A and B are preferred for their higher R^2 adjusted values and for their lower information criterion values.

ARDL Bounds Testing Procedure

Models A and B

The same co-integration equations have been employed under the ARDL bounds testing approach as in the Engle and Granger two-step models above, so the same reasoning applies.[8]

An EViews Add-in developed by Yashar (2014) has been employed for selecting the number of lags in each model under the ARDL bounds testing procedure. Low values of Schwarz information criteria and Akaike information criteria and high F-statistic of the Wald-test are the three main factors for selecting lag length. Dynamic models have the following lag structure:

- Model A: ARDL(3,3,3,3,2,1) [LOG(RMFHBS) = f(EONIA, LOG(BDLNS), LOG(HPR), LOG(MS5), LOG(SFX))],
- Model B: ARDL(5,0,1,3,4) [LOG(RMFHBS) = f(EONIA, LOG(EXTLIAB), LOG(CAP), LOG(PRH))].

The error correction mechanism is functioning in the ARDL dynamic models as well, having negative coefficients of lagged (one interval) level data of the dependent variable. All level data represent variables from the co-integrating equations and can be replaced with error correction terms. However, dynamic models will have slightly different short-term results if the first differences of co-integrated level variables are replaced by a single ECT formed by the same variables.

It was already stated that short-term findings frequently contrast long-term results, as cited in the co-integrating equations. A given variable can have a negative effect on the dependent variable in the long term and lead to an increase in the short term.

Dynamic models demonstrate that interest rate spreads are in a positive and very significant relationship with their lags.

By contrast, when compared with the long-term findings, market concentration is in a negative association with interest rate spreads in the short term, confirming the short-term findings of the Engle and Granger two-step dynamic model.

External liabilities lead to smaller interest rate spreads in the short term, contradicting the short-term results of the Engle and Granger two-step model B. Model B has a bigger explanatory power under the EG two-step

[8] Dynamic models under the ARDL bounds testing approach can be found in the Appendices section in Table 8, while co-integrating vectors (ECT) are presented in Table 6 in the same section.

approach and is supposed to have a bigger weight in weighing these contradictory findings.

As it can be expected, non-performing and restructured loans lead to a bigger dependent variable, *i.e.* banks acquire buffers in the short term for overcoming loan losses and to maintain sufficient liquid and capital buffers. The short-term and long-term findings coincide.

In the short term the capital adequacy proxy is in a strong positive association with interest rate spreads, supporting the long-term results. Capital buffer accumulation requires banks to widen the interest rate spreads, as net interest income is the main source of income and self-financing for local banks.

Interbank lending rates lead to smaller interest rate spreads in the short term, negating the positive coefficients for the zero, third and fourth lags in the two-step dynamic model under the Engle and Granger two-step approach. If lower interbank lending rates lead to higher interest rate spreads, it can be assumed that banks lower interest rates on deposits at a faster rate, being able to take advantage of interbank funds availability and of sources of income anchored to their rates.

The first lagged house price variable leads to lower interest rate spreads, while the third lag coefficient has a positive value, making implications inconclusive.

In the short term, the harmonized price index is in strong negative association with the regressand, with lags 3 and 4 being statistically significant and having negative coefficients.

Model A has bigger explanatory power under the ARDL bounds testing approach when compared to the Engle and Granger two-step approach in terms of the dynamic model's ability to explain the variation of the dependent variable (R^2 adjusted). The residual of the model has passed the serial correlation, heteroscedasticity and normality tests. However, model B suits examination under the Engle and Granger two-step approach better than model A.

Model A and model B findings validate each other to a bigger extent through confirmation under both types of co-integration techniques.

Models C to I

Models C to I have the following lag structure, as suggested by the EViews Add-in developed by Yashar (2014):[9]

- Model C: ARDL(3,5,2,4) [LOG(RMFHBS) = f(LOG(EXTLIAB), LOG(EXPTONII), LOG(ROA))]
- Model D: ARDL(3,2,3,3,3,4) [LOG(RMFHBV) = f(LOG(BDLNS), EURBR3M, LOG(EXTLIAB), LOG(UNMPL), LOG(YR))]
- Model E: ARDL(1,3,1,3,4,3) [LOG(RMFHBV) = f(LOG(BDLNS), LOG(UNMPL), LOG(YR), LOG(HHI5), LOG(LNSTODPTS))]
- Model F: ARDL(3,3,4,3,4,1,1) [LOG(RMFHBV) = f(LOG(BDLNS), EONIA, LOG(UNMPL), LOG(YR), LOG(LNSTOASTS), LOG(LNSTODPTS))] is transformed into ARDL(0,2,4,1(LAG 3),2(LAG1 AND LAG4),0,0) PASSING THE RESIDUAL TESTS,
- Model G: ARDL(3,1,4,4,4) [LOG(RMFHBV) = f(LOG(EXTLIAB), LOG(FOWN), LOG(YR), LOG(EXPTONII))]
- Model H: ARDL(1,1,3,2,4) [LOG(RMFHBV) = f(LOG(EXTLIAB), LOG(YR), LOG(ROE), LOG(FOWN))]
- Model I: ARDL(1,3,3,3,2,4) [LOG(RMFHBV) = f(LOG(YR), LOG(BDLNS), LOG(LNSTODPTS), LOG(FOWN), LOG(HHI5))]

The error correction mechanism functions well in the ARDL dynamic models with the coefficient of lagged (one interval) level variables being negative.

A positive association between the dependent variable and its lagged values in the dynamic model is revealed, confirming the findings of models A and B.

The expense to net interest income ratio leads to larger interest rate spreads in the long term and to smaller interest rate spreads in the short term in model C, while in model G, the opposite dependencies are revealed, *i.e.* the expense to net interest income ratio has a positive coefficient in the co-integrating equation and a negative coefficient in the dynamic model. Short-term results of model C are also corresponding to the main findings of Mihailov (2005, 2014). Model G results, however, are in line with Claeys and Vander Vennet (2004).

[9] All dynamic models under the ARDL bounds testing approach can be found in Table 8 in the Appendices section.

The return on assets ratio is in a positive relation with interest rate spreads in model C, while at 9% significance level the return on equity leads to larger dependent's values in model H. It can be assumed that in the near term banks decrease margins as profitability improves (in terms of return on assets ratio). Banks become more confident and more willing to lend, which results in lower lending rates and spreads.

Market concentration is in a negative association with interest rate spreads in the short term. Market share and HHI have been assigned negative coefficients in the dynamic models (models A, E and I). In a pursuit to increase market share, banks boost lending, thus decreasing the spread between loan and deposit rates.

In the short term, external liabilities' coefficients have opposing values in models B, C and D. While in models C and D, a short-term positive association exists, in model B, external liabilities lead to smaller interest rate spreads.

The non-performing loans parameter has a controversial impact on the dependent variable, having a negative coefficient in the short term in models E and I, and a positive coefficient in models A, D and F. Non-performing loans have a positive long-term effect on interest rate spreads in model A, while in models D, E, F and I, there is a negative association. While in models D and F the negative and positive signs are reversed between short-term and long-term dimension, models A, E and I don't share this feature.

The dynamic models reveal that in the short term, the loan to deposit ratio, employed as a risk proxy, has a significant positive effect on interest rate spreads, but a negative one in the long term.

In the short term, foreign ownership leads to higher interest rate spreads. Banks are required by their parent company to achieve higher net and operating margins, so that the parent organizations can recover investments as *per* projections. Co-integrating relations, however, reveal that foreign ownership leads to smaller interest rate spreads in models H and I and to higher interest rate spreads in model G.

Interbank lending rates have a contradicting effect on the regressand in the short term. The EONIA variable has a negative effect on the dependent in model A and leads to higher interest rate spreads in model F, while EURBR3M leads to higher interest rate spreads.

Unemployment leads to lower interest rate spreads in the short term. Banks reduce interest rate spreads, aiming to improve the quality of their loan

portfolio through attracting new borrowers and easing conditions to existing ones.

The change in GDP has an ambiguous impact on the dependent variable. In models D, F (lag 4), G and H, GDP is in a positive association with interest rate spreads, while the change of GDP leads to lower interest rate spreads in models E, F(lag 1) and I. However, the long-term results are persistent.

All presented models pass the serial correlation, heteroscedasticity and normality residual tests. Their ability to explain more than half of the variation in the dependent variable and the overall confirmation of results from the Engle and Granger two-step approach boost confidence in the results, despite some short-term results' divergence.

Model F follows a different lag structure from that suggested by the EViews Add-in developed by Yashar (2014). Instead of being an ARDL(3,3,4,3,4,1,1), it is an ARDL(0,2,4,3(lag 3 only),2(lag 1 and 4),0,0) for being in shape to overcome serial correlation issues.

Summary and Conclusions

Because of mixed economic, bank-specific and local factors, interest rate spreads in Bulgaria have been on average above the mean EU member states values for a considerable period. Bulgarian banks motivate their interest rate spreads by different factors in the short and in the long term, the biggest impact on interest rate spreads being from regressors like: economic activity, market concentration, foreign ownership, external liabilities, profitability, risk aversion, structural changes due global financial crisis and induced by local banking system shocks.

Sector competitiveness is important to interest rate spreads. Competition in the sector promotes lower interest rate spreads in the long term; however, in the short term, a less competitive bank sector promotes lower interest rate spreads, probably due to attempts at increasing market shares.

General economic activity has a strong inverted association with interest rate spreads. In periods of economic improvement, banks tend to diminish the interest rate spreads, and in time of economic deterioration, the opposite development is expected. Higher GDP, house prices, consumer prices and lower unemployment all lead to smaller interest rate spreads in the long term. In the short term, unemployment is distinctive to the dynamics in the spread between loan and deposit rates and has a positive impact on the dependent variable.

Banks act pro-cyclically in an economic sense, which is also confirmed by the negative dependency of interest rate spreads on the risk proxy variable loan to deposit ratio in the long term. However, in the short term, banks have a reverse response to the risk proxy variable of the loan to deposit ratio, revealing higher risk aversion.

External financing for banks, provided mostly by their foreign parent companies, usually leads to lower interest rate spreads in the long run. Bank privatization not only leads to transfer of know-how, but also to an increase in FDIs, especially the part provided by the bank's headquarters to Bulgarian subsidiaries, resulting in a lower interest rate spread in the long term. In the short term, however, foreign ownership stimulates larger interest rate spreads, as banks are required to achieve bigger returns on their assets and equity through a larger difference between the interest earned on loans and paid to deposits.

The accumulation of capital buffers, *i.e.* the increase in the capital adequacy ratio and net income margins, lead to an increase in the interest rate spreads in the long run. In the long term, higher loan to assets ratio also stimulates banks to demand higher interest rate spreads for achieving higher capital ratios due to the denominator growth (*i.e.* the risk-adjusted assets).

It can be assumed that banks widen interest rate spreads to absorb economic and financial shocks. Results under both econometric methods suggest that two structural changes events, namely global and local financial and economic shocks, play a key role in the interest rate spreads dynamics in Bulgaria in the analyzed time period. Dummy variables for the immediate effects of the last global financial crisis, with a severe impact in the first quarter of 2009 and Corporate Commercial Bank's crisis in the second quarter of 2014, have widened interest rate spreads.

References

Brock, P. L. and L. Rojas Suarez (2000). Understanding the Behavior of Bank Spreads in Latin America, *Journal of Development Economics* 63, pp. 113–134.

Chortareas, G., J. G. Garza-Garcia, C. Girardone (2012). Competition, Efficiency and Interest Rate Margins in Latin American Banking, Elsevier, *International Review of Financial Analysis*, Vol. 24, September 2012, pp. 93–103.

Claeys, S. and R. Vander Vennet (2004). Determinants of bank interest margins in Central and Eastern Europe: A comparison with the West, Elsevier, vol. 32(2), pp.197–216.

Davidson, R. and J. MacKinnon (1993). Estimation and Inference in Econometrics, Oxford University Press, London, p. 708, table 20.1.

Deans, C. and C. Stewart (2012). Banks' Funding Costs and Lending Rates, RBA Bulletin, October 2012, pp. 36–43.

Demirgüç-Kunt, A. and H. Huizinga (1999). Determinants of commercial bank interest margins and profitability: some international evidence, *The World Bank Economic Review*, 1999.

Engle, R. and C. Granger (1987). Co-integration and error correction: Representation, estimation and testing, *Econometrica* 55 (2), pp. 251–276.

English, W. (2002). Interest rate risk and bank net interest margins, *BIS Quarterly Review*, December 2002, pp. 67–82.

Fungáčová, Z. and T. Poghosyan (2009). Determinants of bank interest margins in Russia: Does bank ownership matter?, Bank of Finland, BOFIT Institute for Economies in Transition, Discussion Papers 22 • 2009.

Gelos, R. (2006). Banking Spreads In Latin America, IMF Working Paper, WP/06/44.

Gul, S., F. Irshad and K. Zaman (2011). Factors Affecting Bank Profitability in Pakistan, *The Romanian Economic Journal*, XIV, No. 39, March 2011.

Ho, T., A. Saunders (1981). The determinants of bank interest margins, theory and empirical evidence. *Journal of Financial and Quantitative Analysis* 16, pp. 581–600.

Horváth, R. (2009). The Determinants of the Interest Rate Margins of Czech Banks, *Czech Journal of Economics and Finance* (Finance a uver), Charles University Prague, Faculty of Social Sciences, Vol. 59(2), pp. 128–136.

Liebeg, D. and M. Schwaiger (2007). Determinants of Bank Interest Margins in Central and Eastern Europe, *Financial Stability Report*, issue 14, pp. 68–84.

Männasoo, K. (2012). Determinants of Bank Interest Spread in Estonia, Eesti Pank, Working Paper Series, 1/2012.

Mihailov, M. (2014). Factors Affecting Bank Interest Rate Spread Dynamics in Bulgaria (*in Bulgarian only*), Bulgarian National Bank, Discussion Papers 96/2014.

Mihailov, M. (2005). Interest Rate Spreads of Commercial Banks in Bulgaria (*in Bulgarian only*), Bulgarian National Bank, Discussion Papers 47/2005.

Pesaran, M. H., Y. Shin and R. Smith (1999). Bounds testing approaches to the analysis of long run relationships, Cambridge Working Papers in Economics, 9907, Faculty of Economics, University of Cambridge.

Pesaran, M. H., Y. Shin and R. Smith (2001). Bounds testing approaches to the analysis of level relationships, *Journal of Applied Econometrics*, John Wiley & Sons, Ltd., Vol. 16(3), pp. 89–326.

Poghosyan, T. (2012). Financial Intermediation Costs in Low-Income Countries: The Role of Regulatory, Institutional, and Macroeconomic Factors, IMF WP/12/140.

Raharjo P, D. Hakim, A. Manurung and T. Maulana (2014). The Determinant of Commercial Banks' Interest Margin in Indonesia: An Analysis of Fixed Effect Panel Regression, *International Journal of Economics and Financial Issues*, Vol. 4, No. 2, 2014, pp. 295–308.

Royal Swedish Academy of Sciences (2003). Time Series Econometrics: Cointegration and Autoregressive Conditional Heteroscedasticity, downloadable from: http://www-stat.wharton.upenn.edu/steele/HoldingPen/NobelPrizeInfo.pdf

Sheriff, I. and G. Amoako (2014). Macroeconomic Determinants of Interest rate Spread in Ghana: Evidence from ARDL Modelling Approach, *Journal of Finance and Bank Management*, Vol. 2, No. 2, pp. 115–132.

Trinugroho, I., A. Agusman, A. Tarazi (2012). Why Have Bank Interest Margins Been so High in Indonesia Since the 1997/1998 Financial Crisis?

Ugur, A. and E. Hakan (2010). Determinants of the Net Interest Margins of Banks in Turkey, *Journal of Economic and Social Research* 12 (2) 2010, pp. 101–118.

Yashar, T. (2014). ARDLBOUND: An EViews Add-in to perform ARDL bound test based on Pesaran *et al.* (1999, 2001).

Other sources:

Bulgarian National Bank (BNB) – *www.bnb.bg*.

European Central Bank (ECB) – *www.ecb.it*.

National Statistical Institute (NSI) – *www.nsi.bg*.

ESRB risk dashboard, December 2014, Section 3: *Credit risk*, page 3.

The Word Bank WDI database, 2014.

Appendices

Table 4. Non-financial Companies' Lending Margins

	2007	2008	2009	2010	2011	2012	2013	Nov. 2014
BG	4.3	2.7	2.2	4.1	4.4	3.3	4.6	5.4
PT	1.8	2.6	2.6	2.1	3.3	4.0	3.5	3.6
IE	2.3	2.0	1.3	1.6	1.6	2.0	2.8	3.5
GR	1.7	0.5	2.0	1.7	2.2	1.6	2.1	3.4
RO	3.4	1.6	2.8	2.0	2.7	3.1	3.5	3.3
SI	1.9	2.0	3.1	3.2	2.8	2.7	3.1	3.1
CY	2.0	0.9	1.9	2.6	2.7	2.4	3.4	2.9
LV	3.0	3.2	1.5	3.0	3.3	2.6	3.9	2.7
HR					3.8	2.8	2.7	2.6
LT	1.7	1.5	1.2	3.1	3.1	2.8	2.8	2.5
ES	1.2	0.8	0.8	0.6	1.3	0.8	1.9	2.5
MT	1.8	2.2	2.5	2.9	1.6	2.1	1.1	2.3
HU	1.6	0.2	2.0	2.0	2.3	1.5	1.9	2.3
EE	2.6	2.2	2.9	3.1	3.0	2.8	2.8	2.0
NL	0.9	0.9	1.8	2.0	1.9	1.9	1.9	1.7
PL	2.9	2.1	3.1	2.7	2.0	2.1	1.9	1.6
CZ	2.0	2.5	2.7	2.8	2.2	2.0	2.0	1.6
DE	1.4	1.6	2.1	2.2	2.0	1.7	1.8	1.5
BE	1.3	1.6	1.5	1.6	1.7	1.4	1.6	1.5
IT	1.3	1.5	1.3	1.5	1.3	1.3	1.7	1.4
SK	0.0	2.7	1.9	2.3	2.1	1.6	1.8	1.3
FI	0.8	1.2	1.5	1.6	1.6	1.5	1.8	1.3
SE	0.7	0.6	1.2	0.7	0.8	1.0	1.2	1.2
LU	1.2	1.6	1.6	1.8	2.0	1.7	1.6	1.1
AT	0.7	1.2	0.9	1.0	0.9	1.1	1.2	1.1
FR	0.9	1.5	1.4	1.2	1.2	1.1	1.3	1.1
DK	1.0	1.0	1.5	1.9	1.7	1.9	1.2	0.9
min	**0.0**	**0.0**	**0.0**	**0.0**	**0.0**	**0.0**	**0.0**	**0.0**
max	**4.3**	**3.2**	**3.1**	**4.1**	**4.4**	**4.0**	**4.6**	**5.4**
mean	**1.6**	**1.6**	**1.8**	**2.0**	**2.1**	**1.9**	**2.2**	**2.1**
median	**1.6**	**1.6**	**1.8**	**2.0**	**2.0**	**1.9**	**1.9**	**1.9**
BG above average	**169%**	**73%**	**17%**	**101%**	**110%**	**67%**	**111%**	**161%**

Note: BG – Bulgaria, PT – Portugal, IE – Ireland, GR – Greece, RO – Romania, SI – Slovenia, CY – Cyprus, LV – Latvia, HR – Croatia, LT – Lithuania, ES – Spain, MT – Malta, HU – Hungary, EE – Estonia, NL – Netherlands, PL – Poland, CZ – Czech Republic, DE – Germany, BE – Belgium, IT – Italy, SK – Slovakia, FI – Finland, SE – Sweden, LU – Luxemburg, AT – Austria, FR – France, DK – Denmark.

Source: ECB, Own calculations.

Table 5. Households Housing Purchases Lending Margins

	2007	2008	2009	2010	2011	2012	2013	Nov. 2014
HU	3.9	-0.1	3.8	4.5	4.8	5.1	5.5	4.8
BG	3.7	2.6	2.9	3.2	3.5	3.2	4.1	4.5
HR					2.5	2.3	2.8	3.4
RO	0.7	-5.4	-1.9	0.2	0.6	0.3	2.8	3.3
IE	1.1	1.3	1.4	1.6	0.7	2.4	2.7	3.3
LV	2.2	3.0	1.7	2.4	2.8	2.9	2.7	3.0
NL	0.9	2.2	4.2	3.4	3.4	3.6	3.1	2.8
PL	2.7	2.8	3.7	3.0	2.2	2.6	2.8	2.7
DK	1.5	1.7	2.3	2.7	1.7	2.2	1.9	2.7
BE	0.9	2.5	3.2	3.1	3.0	3.0	3.1	2.6
CZ	2.1	3.3	4.7	4.0	3.3	2.9	3.0	2.3
GB	0.1	2.4	3.2	2.9	2.2	2.9	2.4	2.3
ES	1.0	1.8	0.6	0.2	1.1	0.3	1.6	2.1
SI	2.4	2.3	1.5	1.4	1.5	0.5	1.7	2.1
SK	0.0	4.3	4.4	3.7	3.4	3.3	3.2	2.1
PT	0.8	1.2	0.8	0.0	1.1	1.7	1.7	2.0
DE	1.0	2.0	3.3	2.8	2.3	2.3	2.4	1.9
LU	0.8	1.8	1.6	1.6	1.6	2.0	1.8	1.8
IT	1.5	2.0	2.0	1.7	1.2	1.4	1.7	1.8
FR	0.6	2.2	2.8	2.0	1.9	2.2	2.1	1.7
EE	1.9	0.7	2.4	2.1	2.1	2.2	2.2	1.7
CY	1.2	0.8	0.3	1.0	1.4	0.7	2.4	1.6
AT	0.8	2.1	1.7	1.4	1.2	1.7	1.6	1.6
LT	0.9	-0.3	0.7	2.6	2.3	1.7	1.8	1.5
SE	0.9	1.2	1.2	1.4	1.3	1.2	1.1	1.3
MT	1.3	1.1	1.6	1.8	0.8	1.3	0.6	1.1
GR	0.0	0.2	1.7	0.2	-0.4	-1.6	0.0	0.9
FI	0.7	1.2	1.2	1.2	1.2	1.2	1.3	0.7
min	**0.0**	**-5.4**	**-1.9**	**0.0**	**-0.4**	**-1.6**	**0.0**	**0.7**
max	**3.9**	**4.3**	**4.7**	**4.5**	**4.8**	**5.1**	**5.5**	**4.8**
mean	**1.3**	**1.5**	**2.1**	**2.1**	**1.9**	**2.0**	**2.3**	**2.3**
median	**1.0**	**1.8**	**1.7**	**2.0**	**1.8**	**2.2**	**2.3**	**2.1**
BG above average	**202%**	**74%**	**38%**	**60%**	**88%**	**65%**	**83%**	**104%**

Note: BG – Bulgaria, PT – Portugal, IE – Ireland, GR – Greece, RO – Romania, SI – Slovenia, CY – Cyprus, LV – Latvia, HR – Croatia, LT – Lithuania, ES – Spain, MT – Malta, HU – Hungary, EE – Estonia, NL – Netherlands, PL – Poland, CZ – Czech Republic, GB – Great Britain, DE – Germany, BE – Belgium, IT – Italy, SK – Slovakia, FI – Finland, SE – Sweden, LU – Luxemburg, AT – Austria, FR – France, DK – Denmark.

Source: ECB, Own calculations.

Table 6. Co-integrating Equations

	MODEL A		MODEL B		MODEL C		MODEL D		MODEL E		MODEL F		MODEL G		MODEL H		MODEL I	
DEPENDENT VARIABLE	LOG(RMFHBS)		LOG(RMFHBV)		LOG(RMFHBS)		LOG(RMFHBS)		LOG(RMFHBV)		LOG(RMFHBV)		LOG(RMFHBV)		LOG(RMFHBV)		LOG(RMFHBV)	
EXPLANATORY VARIABLES	Coef.	p-value	Coef.	p-value	Coef.	p-value	Coef.	p-value	Coef.	p-value	Coef.	p-value	Coef.	p-value	Coef.	p-value	Coef.	p-value
CONSTANT	-0.99	**0.09**	7.97	**0.00**	-2.30	**0.00**	17.30	**0.02**			24.77	**0.00**			10.69	**0.00**		
LOG(EXPTONII)					-0.47	**0.00**							1.17	**0.00**				
LOG(EXTLIAB)			0.19	**0.00**	-0.17	**0.01**	-0.29	**0.01**					-0.54	**0.00**	-0.44	**0.00**		
LOG(ROE)															0.16	**0.00**		
LOG(ROA)					0.19	**0.00**												
LOG(BDLNS)	0.03	**0.02**					-0.22	**0.00**	-0.13	**0.00**	-0.16	**0.02**					-0.07	**0.01**
LOG(LNSTODPTS)									-1.05	**0.00**	-1.42	**0.04**					-1.44	**0.00**
LOG(LNSTOASTS)											1.52	**0.09**						
LOG(CAP)			1.05	**0.00**														
LOG(FOWN)													1.50	**0.01**	-1.05	**0.09**	-1.35	**0.02**
LOG(MS5)	0.66	**0.00**																
LOG(HHI5)									0.34	**0.11**							0.64	**0.01**
EURBR3M							10.01	**0.00**										
EONIA	9.35	**0.00**	13.18	**0.00**							11.85	**0.00**						
LOG(UNMPL)							0.59	**0.00**	0.41	**0.00**	0.45	**0.02**						
LOG(YR)							-2.16	**0.01**	-0.47	**0.00**	-2.82	**0.00**	-0.32	**0.00**	-1.49	**0.00**	-0.79	**0.00**
LOG(HPR_R1)	-0.29	**0.00**																
LOG(PRH)			-1.87	**0.00**														
LOG(SFX_R)	0.06	**0.01**																
TREND			0.02	**0.00**			0.02	**0.00**			0.02	**0.01**						

Note: P-values below or equal to 10% (p-values ≤ 0.10) are in bold.

Table 7. Engle and Granger Dynamic Models

	MODEL A		MODEL B		MODEL C		MODEL D		MODEL E		MODEL F		MODEL G		MODEL H		MODEL I	
DEPENDENT VARIABLE	ΔLOG(RMFHBS)		ΔLOG(RMFHBS)		ΔLOG(RMFHBS)		ΔLOG(RMFHBV)		ΔLOG(RMFHBV)		ΔLOG(RMFHBV)		ΔLOG(RMFHBV)		ΔLOG(RMFHBV)		ΔLOG(RMFHBV)	
EXPLANATORY VARIABLES	Coef.	p-value	Coef.	p-value	Coef.	p-value	Coef.	p-value	Coef.	p-value	Coef.	p-value	Coef.	p-value	Coef.	p-value	Coef.	p-value
CONSTANT			0.09	**0.00**					-0.05	**0.03**			-0.09	**0.02**				
ECT(-1)	-1.14	**0.00**	-2.56	**0.00**	-0.65	**0.00**	-1.05	**0.00**	-0.96	**0.00**	-1.19	**0.00**	-0.66	**0.00**	-0.67	**0.00**	-0.78	**0.00**
Δ(log(RMFHBS(-1)))	0.51	**0.00**			0.59	**0.00**			1.58	**0.00**								
Δ(log(RMFHBS(-2)))	0.39	**0.00**			0.44	**0.00**			0.32	**0.03**								
Δ(log(RMFHBV(-2)))															0.27	**0.04**	0.24	**0.06**
Δ(log(RMFHBV(-3)))															0.27	**0.04**	0.29	**0.03**
Δ(log(EXPTONII(-1)))					0.53	**0.00**							-1.02	**0.00**				
Δ(log(EXPTONII(-2)))					0.21	0.13							-0.92	**0.00**				
Δ(log(EXPTONII(-3)))					-0.36	**0.01**							-0.56	**0.03**				
Δ(log(ROA(-1)))					-0.08	**0.02**												
Δ(log(MS5(-1)))	-0.63	**0.02**																
Δ(log(MS5(-2)))	-0.44	**0.05**																
Δ(log(EXTLIAB))			0.48	**0.00**	-1.47	**0.00**							-0.90	**0.00**	-0.89	**0.00**		
Δ(log(EXTLIAB(-1)))			0.43	**0.00**														
Δ(log(EXTLIAB(-2)))			0.50	**0.00**														
Δ(log(BDLNS(-2)))									-0.20	**0.05**	-0.17	**0.08**						
Δ(log(CAP(-3)))									0.46	**0.01**								
Δ(log(CAP(-4)))									-0.92	**0.00**								
Δ(log(LNSTOASTS))											0.81	0.13						

Δ(log(FOWN(-1)))															2.48	**0.01**		
Δ(EONIA)			15.28	**0.00**														
Δ(EONIA(-1))	-6.44	**0.03**	-7.75	**0.02**														
Δ(EONIA(-3))			3.52	**0.05**														
Δ(EONIA(-4))			3.75	**0.03**														
Δ(log(UNMPL))							0.46	**0.06**										
Δ(log(UNMPL(-2)))											0.87	**0.00**						
Δ(log(YR))							-2.01	**0.01**			-1.83	**0.00**	-2.36	**0.00**	-2.72	**0.00**	-1.79	**0.01**
Δ(log(YR(-3)))													2.73	**0.00**	2.39	**0.00**	1.88	**0.01**
Δ(log(HPR_R1))	1.12	**0.00**																
Δ(log(HPR_R1(-1)))	-0.55	**0.02**																
Δ(log(PRH))									-3.23	**0.00**								
Δ(log(PRH(-4)))									1.33	**0.04**								
trend			-0.01	**0.00**									0.00	**0.10**				
D1	0.05	**0.06**	0.07	**0.01**	0.04	0.14	0.11	**0.07**			0.13	**0.03**						
D2			0.28	**0.00**					0.08	**0.01**								
R^2 adj.	0.69		0.79		0.53		0.56		0.51		0.62		0.59		0.57		0.55	
F-Stat	10.45		9.09		6.21		11.35		14.59		11.46		7.78		8.00		8.79	
Akaike info criterion	-3.81		-4.37		-3.45		-2.08		-1.96		-2.16		-2.00		-1.98		-1.98	
Schwarz criterion	-3.43		-3.60		-3.11		-1.87		-1.79		-1.91		-1.62		-1.69		-1.72	
DW stat	1.84		1.97		1.93		2.22		2.26		2.03		2.40		1.84		2.26	
Residual tests	H0 probability		H0 probability		H0 probability		H0 probability		H0 probability		H0 probability		H0 probability		H0 probability		H0 probability	

Ser. Corr. Test (type)	23% (Breusch-Godfrey Serial Correlation LM Test- 3 lags)	69% (Breusch-Godfrey Serial Correlation LM Test- 6 lags)	OK (Correlogram of residuals and Correlogram of squared residuals)	63% (4 lags)	18% (Breusch-Godfrey Serial Correlation LM Test- 4 lags)	32% (Breusch-Godfrey Serial Correlation LM Test- 4 lags)	OK (Correlogram of squared residuals)	66% (Breusch-Godfrey Serial Correlation LM Test- 6 lags)	61%(Breusch-Godfrey Serial Correlation LM Test- 6 lags)
Heteroscedasticity Test (type)	37% (ARCH-4 lags)	32% (Breusch-Pagan-Godfrey)	58% (Breusch-Pagan-Godfrey)	80% (ARCH-4 lags)	27% (Breusch-Pagan-Godfrey)	42% (Breusch-Pagan-Godfrey)	89% (Breusch-Pagan-Godfrey)	93% (Breusch-Pagan-Godfrey)	72% (Breusch-Pagan-Godfrey)
Jarque-Bera Test	47%	34%	57%	69%	23%	48%	88%	94%	97%

Note: P-values below or equal to 10% (p-values ≤ 0.10) are in bold.

Table 8. ARDL Dynamic Models

	MODEL A		MODEL B		MODEL C		MODEL D		MODEL E		MODEL F		MODEL G		MODEL H		MODEL I	
DEPENDENT VARIABLE	ΔLOG(RMFHBS)		ΔLOG(RMFHBS)		ΔLOG(RMFHBS)		ΔLOG(RMFHBV)		ΔLOG(RMFHBV)		ΔLOG(RMFHBV)		ΔLOG(RMFHBV)		ΔLOG(RMFHBV)		ΔLOG(RMFHBV)	
EXPLANATORY VARIABLES	Coef.	p-value	Coef.	p-value	Coef.	p-value	Coef.	p-value	Coef.	p-value	Coef.	p-value	Coef.	p-value	Coef.	p-value	Coef.	p-value
CONSTANT	-0.88	0.45	-9.69	**0.00**	-4.43	**0.00**			-13.91	0.27	-14.83	**0.04**	10.72	**0.07**	11.34	**0.02**	-12.45	0.31
LOG(RMFHBS(-1))	-1.50	**0.00**	-1.67	**0.00**	-1.46	**0.00**												
LOG(RMFHBV(-1))							-1.59	**0.00**	-2.17	**0.00**	-1.43	**0.00**	-1.21	**0.00**	-0.73	**0.00**	-2.12	**0.00**
LOG(BDLNS(-1))	0.12	**0.00**					-0.25	**0.00**	0.40	**0.03**	-0.53	**0.00**					0.21	**0.01**
LOG(CAP(-1))			-1.04	**0.02**														
LOG(EXTLIAB(-1))			0.18	**0.01**	-0.51	**0.00**	-0.06	0.83					0.14	0.60	0.22	0.40		
LOG(LNSTODPTS(-1))									-10.09	**0.00**	-2.52	**0.01**					-7.78	**0.00**
LOG(LNSTOASTS(-1))											3.05	**0.01**						
LOG(EXPTONII(-1))					-1.27	**0.00**							0.10	0.88				
LOG(ROE(-1))															-0.07	0.27		

LOG(ROA(-1))					0.22	**0.01**												
LOG(FOWN(-1))													-3.59	**0.01**	-1.97	**0.08**	1.28	0.26
LOG(MS5(-1))	1.03	**0.06**																
LOG(HHI5(-1))									7.18	**0.00**							5.34	**0.00**
EURBR3M(-1)							-25.80	**0.00**										
EONIA(-1)	15.16	**0.02**	11.31	**0.01**							-32.28	**0.00**						
LOG(HPR_R1(-1))	-0.66	**0.03**																
LOG(UNMPL(-1))							0.07	0.81	-0.84	0.15	0.18	0.44						
LOG(YR(-1))							-0.52	**0.00**	-4.06	**0.01**	1.21	**0.11**	-1.58	**0.02**	-1.46	**0.01**	-2.79	**0.00**
LOG(SFX_R(-1))	0.27	**0.00**																
LOG(PRH(-1))			0.58	**0.00**														
Δ(log(RMFHBS(-1)))	0.68	**0.11**	0.87	**0.02**	1.23	**0.00**			0.19	0.32								
Δ(log(RMFHBS(-2)))	0.56	**0.08**	0.44	**0.07**	0.93	**0.00**												
Δ(log(RMFHBS(-3)))	0.42	**0.01**	0.22	0.15	0.57	**0.00**												
Δ(log(RMFHBS(-4)))			0.25	0.12														
Δ(log(RMFHBS(-5)))			0.48	**0.01**														
Δ(log(RMFHBV(-1)))							0.51	**0.10**					-0.04	0.86	-0.14	0.51	0.27	0.17
Δ(log(RMFHBV(-2)))							0.43	**0.07**					0.36	**0.07**				
Δ(log(RMFHBV(-3)))							0.45	**0.02**					0.31	**0.06**				
Δ(log(EXPTONII(-1)))					1.24	**0.00**							-0.86	0.15				
Δ(log(EXPTONII(-2)))					0.82	**0.00**							-1.03	**0.07**				
Δ(log(EXPTONII(-3)))													-1.31	**0.01**				
Δ(log(EXPTONII(-4)))													-0.73	**0.03**				
Δ(log(ROA(-1)))					-0.22	**0.00**												
Δ(log(ROA(-2)))					-0.18	**0.00**												

Δ(log(ROA(-3)))					-0.12	**0.03**												
Δ(log(ROA(-4)))					-0.06	**0.10**												
Δ(log(ROE(-1)))															0.11	0.15		
Δ(log(ROE(-2)))															0.11	0.09		
Δ(log(MS5(-1)))	-1.07	**0.07**																
Δ(log(MS5(-2)))	-0.89	**0.01**																
Δ(log(MS5(-3)))																		
Δ(log(MS5(-4)))																		
Δ(log(HHI5(-1)))									-7.18	**0.00**							-5.26	**0.00**
Δ(log(HHI5(-2)))									-6.16	**0.00**							-4.40	**0.00**
Δ(log(HHI5(-3)))									-3.75	**0.00**							-2.60	**0.00**
Δ(log(HHI5(-4)))									-1.63	**0.01**							-1.07	**0.08**
Δ(log(EXTLIAB(-1)))			-0.07	0.74	0.33	**0.06**	0.40	0.29					-0.53	0.30	0.13	0.80		
Δ(log(EXTLIAB(-2)))			-0.73	**0.00**	0.15	0.34	0.42	**0.06**										
Δ(log(EXTLIAB(-3)))			-0.48	**0.00**	0.24	**0.04**	0.38	**0.03**										
Δ(log(EXTLIAB(-4)))					0.39	**0.01**												
Δ(log(EXTLIAB(-5)))					0.25	**0.02**												
Δ(log(BDLNS(-1)))	-0.03	0.59					0.30	**0.02**	-0.39	**0.10**	0.40	**0.00**					-0.09	0.51
Δ(log(BDLNS(-2)))	0.22	**0.00**					0.27	**0.04**	-0.53	**0.01**	0.14	0.21					-0.35	**0.03**
Δ(log(BDLNS(-3)))	0.13	**0.07**							-0.35	**0.03**							-0.36	**0.05**
Δ(log(CAP(-1)))			1.87	**0.00**														
Δ(log(LNSTODPTS(-1)))									5.03	**0.00**							4.51	**0.00**
Δ(log(LNSTODPTS(-2)))									3.74	**0.00**							2.78	**0.01**
Δ(log(LNSTODPTS(-3)))									2.59	**0.02**							2.44	**0.04**
Δ(log(FOWN(-1)))													3.46	**0.03**	3.40	**0.03**	0.37	0.80

Δ(log(FOWN(-2)))													2.84	**0.09**	0.68	0.59	-1.41	0.22
Δ(log(FOWN(-3)))													-0.07	0.96	1.34	0.28		
Δ(log(FOWN(-4)))													4.72	**0.01**	3.08	**0.02**		
Δ(EONIA(-1))	-12.24	**0.01**	-0.37	0.93							38.23	**0.00**						
Δ(EONIA(-2))	-8.69	**0.02**									21.56	**0.00**						
Δ(EONIA(-3))	-4.81	**0.09**									17.90	**0.00**						
Δ(EONIA(-4))											22.82	**0.00**						
Δ(EURBR3M(-1))							19.78	**0.00**										
Δ(EURBR3M(-2))							3.50	0.54										
Δ(EURBR3M(-3))							18.18	**0.01**										
Δ(log(UNMPL(-1)))							-0.80	0.14	1.29	0.11								
Δ(log(UNMPL(-2)))							-1.06	**0.05**										
Δ(log(UNMPL(-3)))							-1.74	**0.01**			-0.96	**0.01**						
Δ(log(YR(-1)))							1.74	0.15	0.79	0.51	-2.41	**0.00**	0.95	0.46	2.42	**0.05**	-0.51	0.58
Δ(log(YR(-2)))							1.75	0.11	-2.04	**0.03**			1.31	0.22	1.72	**0.10**	-2.34	**0.04**
Δ(log(YR(-3)))							2.36	**0.05**	-2.14	**0.03**			2.89	**0.01**	2.38	**0.02**	-2.10	**0.05**
Δ(log(YR(-4)))							2.14	**0.02**			1.90	**0.02**	2.31	**0.04**				
Δ(log(HPR_R1(-1)))	-0.83	**0.03**																
Δ(log(HPR_R1(-2)))	0.55	0.14																
Δ(log(HPR_R1(-3)))	0.95	**0.00**																
Δ(log(SFX_R(-1)))	-0.03	0.43																
Δ(log(PRH(-1)))			-0.19	0.82														
Δ(log(PRH(-2)))			-0.75	0.39														
Δ(log(PRH(-3)))			-2.50	**0.01**														
Δ(log(PRH(-4)))			-2.99	**0.00**														

R^2 adj.	0.86	0.62	0.65	0.81	0.69	0.76	0.57	0.48	0.67
Akaike info criterion	-4.56	-3.80	-3.88	-2.69	-2.18	-2.43	-1.84	-1.64	-2.11
Schwarz criterion	-3.62	-2.93	-2.96	-1.66	-1.23	-1.70	-0.89	-0.86	-1.12
DW stat	2.04	2.71	1.86	2.44	2.08	2.44	2.24	2.27	1.95
Residual tests	H0 Probability	H0 Probability	H0 Probability	H0 Probability	H0 Probability	H0 Probability	H0 Probability	H0 Probability	H0 Probability
Ser. Corr. Test (type)	34% (Breusch-Godfrey Serial Correlation LM Test- 8 lags)	25% (Breusch-Godfrey Serial Correlation LM Test- 9 lags) OK (Correlogram of squared residuals)	OK (Correlogram of residuals and Correlogram of squared residuals)	24% (Breusch-Godfrey Serial Correlation LM Test- 6 lags)	18% (Breusch-Godfrey Serial Correlation LM Test- 4 lags)	OK (Correlogram of squared residuals)	34% (Breusch-Godfrey Serial Correlation LM Test- 6 lags)	31% (Breusch-Godfrey Serial Correlation LM Test- 6 lags)	26%(Breusch-Godfrey Serial Correlation LM Test- 6 lags)
Heteroscedasticity Test (type)	27% (Breusch-Pagan-Godfrey)	73% (Breusch-Pagan-Godfrey)	95% (Breusch-Pagan-Godfrey)	36% (Breusch-Pagan-Godfrey)	82% (Breusch-Pagan-Godfrey)	57% (Breusch-Pagan-Godfrey)	82% (Breusch-Pagan-Godfrey)	23% (Breusch-Pagan-Godfrey)	40% (Breusch-Pagan-Godfrey)
Jarque-Bera Test	64%	95%	85%	88%	23%	27%	34%	48%	67%

Note: P-values below or equal to 10% (p-values ≤ 0.10) are in bold.

DP/100/2015

Empirical Evidence on the Relationship between Money Supply Dynamics and Prices in Bulgaria

Zornitsa Vladova, Mihail Yanchev

Abstract. This paper undertakes an empirical investigation of the relationship between the dynamics of money supply and prices in Bulgaria over the period 1998–2012. Starting from the hypothesis that the endogeneity of money supply determines a link from prices (and possibly inflation) through money demand to money supply, the paper attempts to address the issue of the possibility of an existence of a feedback effect of money aggregates (currency in circulation, M1 and M2) dynamics on inflation. The econometric analysis is based on three methods. The first one is dynamic cross-correlations which are performed after pre-whitening of the data and then applying univariate residual cross-correlation approach. The second one is Granger-causality tests in the framework of unrestricted VAR models which are conducted both on a rolling basis with a moving start date and fixed end date and with a moving window having a fixed length, with the aim of overcoming the problem of potential instability of coefficients. The third method is the Johansen cointegration technique which allows the analysis of both long-run relations between the variables, which are naturally related to the estimation of money demand models, as well as the short-run dynamics of money growth and inflation.

Резюме. В тази статия е направено емпирично изследване на взаимозависимостите между динамиката на паричното предлагане и цените в България в периода 1998–2012 г. Започвайки с хипотезата, че ендогенността на паричното предлагане предопределя връзката от цените (и възможно инфлацията) посредством паричното търсене към паричното предлагане, изследването се опитва да анализира възможната обратна връзка от динамиката на паричните агрегати (пари в обращение, M1 и M2) върху инфлацията. Иконометричният анализ се базира на три метода. Първият е динамична взаимна корелация, извършвана след предварително моделиране на автокорелацията в данните, след което се прилага едномерният подход на взаимни корелации на остатъчните величини. Вторият метод е тестът за причинност на Грейнджър в рамките на VAR модели без ограничения, който се прилага както с плъзгаща се начална и фиксирана крайна дата, така и с плъзгащ се прозорец, който има фиксирана продължителност, с цел да се преодолее проблемът с потенциалната нестабилност на коефициентите. Третият метод е техниката за коинтеграция на Йохансен, която дава възможност да се анализират както дългосрочните зависимости между променливите, което естествено е свързано с оценка на модели на парично търсене, така и краткосрочната динамика на растежа на паричните агрегати и инфлацията.

Zornitsa Vladova, Economic Research and Forecasting Directorate, Bulgarian National Bank, Corresponding author, vladova.z@bnbank.org.

Mihail Yanchev, Economic Research and Forecasting Directorate, Bulgarian National Bank (worked in the period 2012–2015), yanchev.mihail@gmail.com.

Contents

1. Introduction .....

2. Literature Review .....

3. Analysis of the Statistical Properties of the Data .....
   3.1. Dynamic Cross-correlations .....
   3.2. Granger-causality Tests .....
   3.3. Issues for Robustness of Granger-causality Tests .....

4. Estimates of Money Demand Functions with VECM .....
   4.1. Brief Overview of the Theoretical Framework .....
   4.2. Description of the Data and Methodology .....
   4.3. Results for Currency in Circulation .....
   4.4. Results for M1 .....
   4.5. Results for M2 .....

5. Conclusions .....

References .....

Appendix .....

1. Introduction

This paper undertakes an empirical analysis of the relationship between the dynamics of money supply and prices in Bulgaria. The motivation of the paper derives from the fact that this topic has received relatively little empirical research since the period of the introduction of the currency board arrangement in the country, despite its importance for understanding money demand behaviour and the potentially interdependent relationship between money supply and prices/inflation. The analysis of money demand, that has been rather limited over the last years, is by itself important as it would allow a better understanding of the dynamics of monetary aggregates within the framework of the currency board and thus a deeper understanding of the monetary transmission mechanism in the country. In turn, an analysis of the dynamics of monetary indicators and prices could potentially extend the range of factors used for forecasting inflation, thus improving inflation forecasts. Taking a more general view, the relevance of the topic relates to the future membership of Bulgaria in the euro area where the ECB gives a prominent role to monetary aggregates under the money pillar of its monetary policy framework that is aimed at achieving the primary objective of maintaining price stability.

Money growth impacts inflation developments in the long run as advocated by the quantity theory of money, according to which money growth precedes equal changes in the general price level rate of growth. Nevertheless, the empirical evidence of the relationship between money growth and inflation and the usefulness of money for predicting inflation has been controversial, while the practical use of monetary indicators in the conduct of monetary policy by central banks across the world has also been subject to wide discussions and presently stands far from close to consensus.

The introduction of the currency board arrangement in Bulgaria in 1997 played a pivotal role in the macroeconomic stabilization of the country following the severe financial crisis that started in 1996. The strong nominal anchor of the fixed exchange rate significantly lowered inflationary expectations and eliminated one of the main contributors to inflation in the pre-currency board period – the large exchange rate fluctuations and their quick pass-through to the price level. Another important distinguishing feature of the currency board was the elimination of the possibility for providing credit to the government and the very limited potential opportunity for extending loans to commercial banks under the strict requirements of the lender of last resort function of the Bulgarian National Bank (BNB).

Theoretically, orthodox currency board arrangements bring about an automatic money supply adjustment mechanism. The latter constrains the growth of base money primarily to balance of payment dynamics and corresponding changes in money demand, thus making money supply endogenous. The currency board introduced in Bulgaria belongs to the so-called new generation (or non-orthodox) currency board arrangements which are characterized by the presence of some discretionary powers of monetary authorities that provide opportunities for deliberately affecting or intervening into money supply dynamics (see *e.g.* Nenovsky and Hristov, 2002). In the specific case of Bulgaria these discretionary powers relate to the minimum required reserves and the lender of last resort role of the BNB. The presence of certain channels of discretionary monetary policy especially the minimum required reserves implies that the expansion of money supply may deviate from that determined by market forces, *i.e.* from the automatic money supply mechanism.[1]

One of the aspects of the relationship between money supply and inflation in Bulgaria, related to the effects of past growth of money supply on current period inflation, has been investigated in a number of research papers in the years immediately following the introduction of the currency board, with most papers concluding there were no monetary sources of inflation. However, the short time period available at that time as well as the concentration of econometric methods primarily on single equations can be considered as limitations of these studies. This raises the need for a more thorough analysis of the possible link between money and prices since 1997, including the consideration of issues that have not been investigated deeply so far – both the short- and long-run relation between the variables as well as their potential two-way relationship with the application of multi-equation econometric methods such as VARs.

These specific research issues which lead to new conclusions from empirical econometric analysis for Bulgaria as well as the results from the cointegrated VARs related to the estimation of money demand models for currency in circulation, M1 and M2, which are published for the first time for Bulgaria, represent the original contributions of the paper. Furthermore, the empirical analysis in the paper spans a relatively long period of time:

[1] For an empirical investigation of the validity of the presence of an automatic mechanism under the currency board arrangement (*i.e.* reserve money dynamics following balance of payments dynamics), see Nenovsky, N. *et al.* (2001) for the case of Bulgaria, Estonia and Lithuania and Nenovsky, N. and K. Hristov (2002) only for the case of Bulgaria. Krus (2012) examines the effect of the current account on the changes in the monetary base for a sample of countries with currency board arrangements, using a panel regression method.

starting from 1998 (and depending on data availability from 2000) and ending in the first quarter of 2012.

The present paper analyses the following two hypotheses. The endogeneity of money supply mechanism and the driving force of money demand for money supply make us start with the hypothesis about a link that goes from prices (and inflation) through money demand to money supply. While the case of endogeneity might imply that money supply is an inappropriate factor in models explaining inflation developments, it could be also hypothesized that a feedback effect from money aggregates dynamics on inflation is also possible. For example, this feedback effect could be driven by factors that are external to the operation of the currency board itself (*e.g.* strong capital inflows may be regarded as possible monetary sources of inflation).

The second hypothesis is related to the question whether any short-run disequilibrium between money supply and demand that results in excess money supply could potentially act as a pro-inflationary factor in Bulgaria. This proposition is made originally by Sepp (1995). In his note on inflation under the Estonian currency board, Sepp maintains that while in the long-run money supply and money demand coincide, which practically excludes the possibility for money supply from being a long-run factor for inflation, over the short-run the situation may differ. In particular, in the short-run disequilibria between money supply and money demand may exist, with money supply potentially either constraining or fueling inflation (in the first case when money supply is temporarily lower than money demand and in the second case when money supply temporarily exceeds demand).

The empirical methodology that will be used to address the two research hypotheses – the potential two-way relationship between monetary aggregates and price developments and the effects from possible short-run disequilibrium between money supply and demand on inflation – is based on the following approaches. The initial statistical properties of the data will be investigated with dynamic cross-correlations, while the potential two-way causality will be analysed with the standard Granger-causality tests in a VAR framework. Empirical testing of the second hypothesis requires constructing money demand models which in itself represents an important contribution of the paper to the generally scarce literature on money demand, including an analysis on stability of money demand and especially over the period since 2000. To assess the validity of short-run excess of money supply having an effect on inflation and to allow treatment of potential endogeneity in the relationship between money and prices, we employ vector error-correction models (VECM). The latter represent an

appropriate methodological approach as they make it possible to estimate a relationship between endogenous variables that is based both on their short-run dynamics and an adjustment to a possible long-run equilibrium.

The rest of the paper is structured as follows. In Section 2 we make a literature review of the research studies on money and prices/inflation, summarizing the results obtained so far for Bulgaria since 1998 as well as the studies for the countries with currency boards, specifically examining the Baltic States. Section 3 explores the time-series properties of the data on money and inflation, presents the methodological details of the applied econometric approach on dynamic cross-correlations and Granger-causality tests for the preliminary analysis of the data and reports the results. Section 4 presents the application of VECM in estimating money demand models for currency in circulation, M1 and M2. Section 5 summarizes the conclusions.

2. Literature Review

After the introduction of the currency board in Bulgaria a large number of studies were devoted to investigation of the impact of the new monetary policy framework on inflation. First, in 1998 Yotzov *et al.* report that over the first year of the operation of the currency board base money was not a determinant of inflation dynamics in contrast to the period before the currency board when the exchange rate, past inflation and base money represented the main inflation drivers. Yotzov (2000) finds that unlike the period from December 1991 to June 1997, in the period from July 1997 to December 1999 there was neither a short-term nor a long-term relationship between money supply (reserve money, M2 and M3) and inflation, based on the application of a single equation approach. Beck *et al.* (2003), who make an empirical assessment of the 'discipline' and 'confidence' effects of the established currency board on inflation, provide evidence that a confidence effect did occur, whereas the hypothesized lower effect of lagged money growth on the price level after the regime change found little support in inflation models. In a relatively recent study Arratibel *et al.* (2009) show that in an out-of-sample inflation forecasting exercise at a 12-quarter horizon over the period 2003Q3–2008Q2, single equation models for Bulgaria with monetary indicators included help outperform forecasts obtained with random walk models.

In the early period after the introduction of the new monetary policy regime there were a limited number of studies exploring money demand models for Bulgaria. Based on a single equation approach, Yotzov (2000)

reports a strong statistical effect of real income on real money demand but does not succeed in identifying the theoretically expected negative association between interest rates and money demand, arguing that interest rates had not yet started functioning as an opportunity cost for holding money. In a research work exploring specifically transaction demand for currency in circulation, Nenovsky and Hristov (2000), also employing single equation modelling, conclude that demand for currency in circulation is most significantly determined by the monetary expenditures of households (retail sales in the economy being a weaker determinant). The authors also discover a statistically significant negative elasticity of demand with respect to the BGN/USD exchange rate, as an opportunity cost of holding currency, while the impact of interest rates on time deposits has the expected negative relationship with currency demand but is relatively low. Slavova (2003) estimated narrow and broad money demand models for Bulgaria for three periods: April 1991 – April 1996, May 1996 – June 1997, July 1997 – December 2000. The empirical analysis, conducted again within a single equation approach, reveals different determinants of money demand during the three periods. In the last period the author finds stabilization of money demand, with standard determinants of money demand such as the wage rate and the Treasury bill rate gaining statistically significant explanatory power.

The paper by Komárek and Melecký (2001) is one representative of a number of studies on money demand modeling, taking due consideration of factors specific for the economic development of transition countries. The authors construct a money demand model for narrow money in the Czech Republic over the period 1993–2001, using foreign variables such as the real effective exchange rate and foreign direct investment in addition to the traditional set of explanatory variables. In the case of broad money, the authors conclude that all standard domestic money demand factors apart from inflation have the expected theoretical impact on M2 and furthermore there is also a significant influence from currency substitution and capital mobility represented by the return on foreign deposits in USD and on US assets.

Pikkani (2000) specifies a model for the monetary sector in Estonia, modeling real M2 demand as a function of real GDP and a three-period moving average of inflation over the period from March 1995 to August 1999. Pikkani finds that income elasticity of money demand is higher than 1 (2.02). The explanation provided by the author is related to the potential role of money as an accumulator of wealth due to the practical absence of substitutes for money in the period of estimation. Another potential

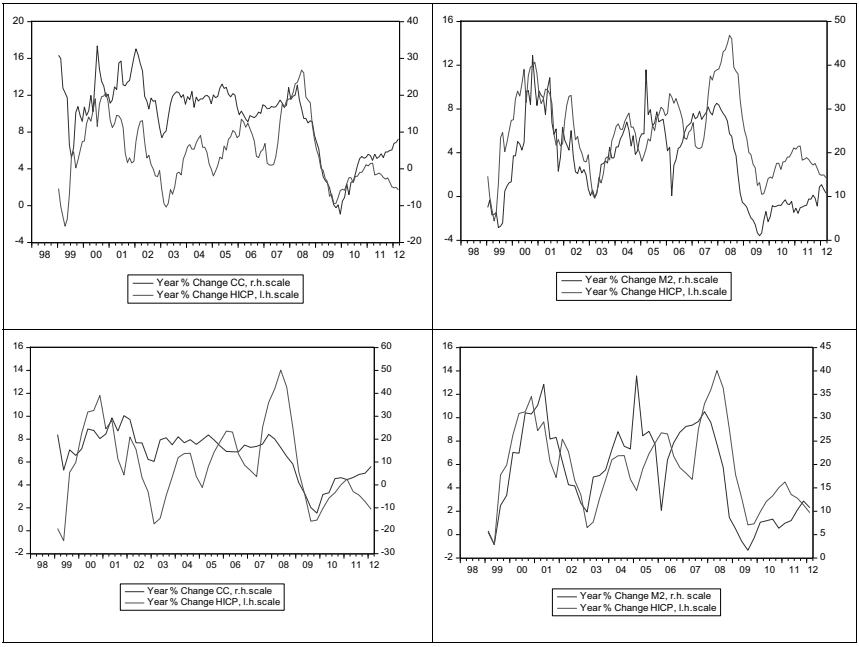
explanation is that income elasticity may also reflect initial conditions of low monetization, financial deepening, higher confidence and economic agents' optimistic expectations about the future. In a paper examining money demand in Latvia, Tillers (2004) establishes a stable money demand function using Johansen vector-error correction model for the period 1996–2003. The variables used for the analysis include real M2, real GDP and real long-term deposit interest rate. Siliverstovs (2007) constructs a money demand model for Estonia for the period 1995Q1–2006Q2 also with the Johansen cointegration procedure. He builds a model with real M2, real GDP, inflation rate as well as short- and long-term interest rates. The author discovers two cointegrating vectors: one characterizing a money demand function, and the other interpreted as a stationary spread between long- and short-term interest rates. As for the previous research for Estonia by Pikkani (2000) and the case of Latvia, the long-run income elasticity of money demand established by Siliverstovs is larger than unity (around 2).

The study by Dreger *et al.* (2006) investigates a long-run money demand relationship over the 1995–2004 period for 10 new EU Member States, using panel cointegration techniques. The authors do not manage to obtain such a relationship when employing the standard variables (broad money, real GDP and short-term market interest rate). Nevertheless, a stable long-run money demand cointegration link is established when the US dollar exchange rate is included. The authors find that income elasticity is larger than unity, while the elasticity with respect to the interest rate is negative with a relatively small magnitude.

Reimers and Roht (2007) construct a money demand function for Estonia using real M2 over the period from 1995 to 2006, applying various econometric techniques ranging from the system approach of Johansen to the Engle-Granger procedure and the dynamic OLS approach. The authors find unstable results when employing the Johansen approach but are able to obtain satisfactory results for a stable money demand function with the Engle-Granger approach. The preferred specification includes real GDP (with income coefficient greater than unity), euro area government bond rates with a maturity of 10 years and euro area money market rate.

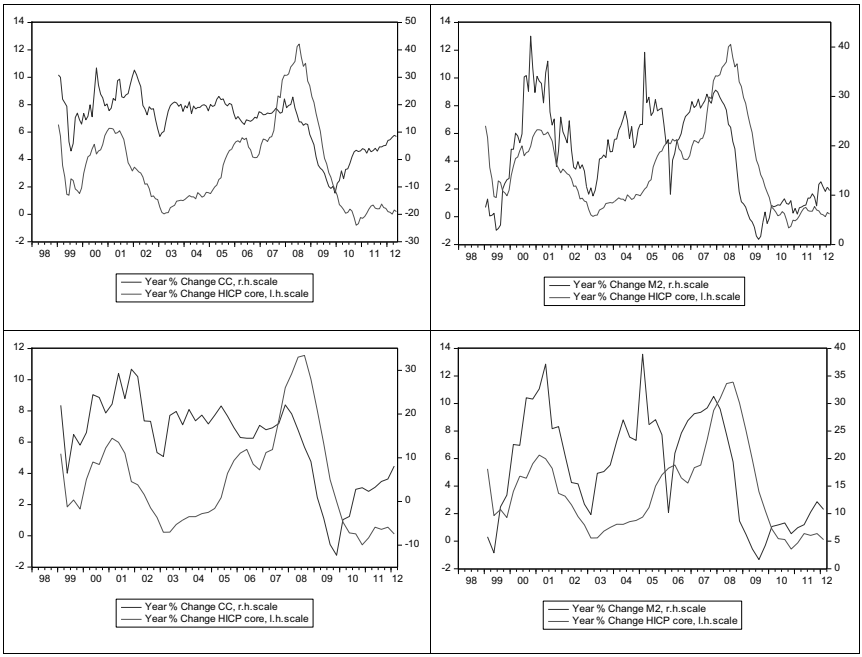
3. Analysis of the Statistical Properties of the Data

This section provides a preliminary analysis of the data focusing on their statistical properties over the period from January 1998 to March 2012, respectively for quarterly data from 1998Q1 to 2012Q1. We start with an investigation of the annual growth rates of monetary aggregates (currency in circulation, M1, M2) and the HICP index (overall and HICP excluding food, energy, administered prices and tobacco as a proxy for core inflation). Below we show the graphs for currency in circulation (CC) and M2 against overall HICP inflation. In the first line we employ monthly data, in the second line quarterly data respectively.



As seen from the graphs, despite the volatility of money growth and inflation, it appears that there is an association between the dynamics of the two series, perhaps more significantly for the case of M2 and HICP both when monthly and quarterly data are used. This finding seems to be confirmed for the developments of core inflation and the two selected monetary aggregates: currency in circulation and M2[2].

[2] The first line of the graphs presents monthly data and the second one quarterly data respectively.



Before proceeding with the empirical analysis, we seasonally adjust the following data series: currency in circulation (CC), monetary aggregate M1 (M1), monetary aggregate M2 (M2), the harmonized index of consumer prices (HICP) and HICP core (HICP_C). For the consumer price data series that are originally released on a monthly frequency, we transform them to quarterly by averaging, and for the monetary aggregates series we take an end-of-period transformation to obtain the respective quarterly frequency.

The preference for using quarterly rather than monthly data is determined by the following two reasons: 1) the substantial volatility present in the data with monthly frequency may be partly overcome with the quarterly aggregation. This argument makes it more appropriate to analyze dynamic cross-correlations of the data and Granger-causality tests based on data with quarterly frequency; 2) the empirical literature on the issue of cointegration reveals that cointegration depends on the total length of the sample rather than on the number of observations which does not give more power to statistical tests with monthly over quarterly data (see *e.g.* Otero and Smith (2000)).

The univariate time series properties of the seasonally adjusted data are examined by the Augmented Dickey-Fuller (ADF) test, the Phillips-

Perron (PP) test and the Kwiatkowski–Phillips–Schmidt–Shin (KPSS) test (see Appendix A1). The results point to the conclusion that the series are integrated of order one, even though for CC and M2 the tests are not conclusive. For robustness of the results, which are known to be characterized by low power especially in cases of structural breaks, we additionally employ the Clemente, Montañés and Reyes unit root test (1998) which can deal with two potential breaks in time series (see Appendix A2). The latter test reveals the presence of two structural breaks for most of the examined data (in some cases both in levels and in first difference form) and furthermore provides strong evidence for the treatment of all series as $I(1)$.

3.1. Dynamic Cross-correlations

The analysis of the relationship between monetary aggregates dynamics and inflation begins with dynamic cross-correlations that can help us to investigate the possible lead-lag links between the series.

The series are transformed by taking first differences of the respective log levels (on a quarterly basis, seasonally adjusted). When the series are transformed by differencing to ensure they are jointly covariance-stationary, then their interrelationship can be determined by examining either their cross correlation function or their cross spectrum. In this paper, we resort to the more frequently applied cross-correlation estimator. The cross-correlogram allows us to determine whether there is a one-way causality between the series or feedback occurring.

We use the following money aggregates: CC, M1 and M2, and the two price indicators: HICP and HICP_C. The table presented below shows the cross correlogram, *i.e.* the degree of co-movement of the first differences of the respective money and price indicators, that is estimated from 1998Q1 to 2012Q1. The correlation dynamics is estimated at lags and leads of 1, 2, 3, 4, 5, 6, 7, and 8 quarters of the inflation series. The contemporaneous correlation between the growth rate of monetary variables and inflation is presented when q=0.

Dynamic cross-correlations between the growth rate of selected monetary variables (*t*) and HICP inflation (*t+k*)
(first differences of the seasonally adjusted series)

	(q-8)	(q-7)	(q-6)	(q-5)	(q-4)	(q-3)	(q-2)	(q-1)	q	(q+1)	(q+2)	(q+3)	(q+4)	(q+5)	(q+6)	(q+7)	(q+8)
CC	-0.23	-0.13	-0.20	-0.02	-0.09	0.07	0.14	0.12	0.45**	0.15	0.18	0.26**	0.20	0.01	-0.07	0.14	0.08
M1	-0.22	-0.14	-0.25	-0.10	-0.05	-0.16	0.14	-0.04	0.22	0.23	0.22	0.37**	0.27**	0.28**	-0.03	0.04	0.09
M2	-0.22	-0.11	-0.21	-0.04	0.06	0.03	0.25**	0.18	0.37**	0.42**	0.25**	0.30**	0.15	0.18	-0.05	-0.08	-0.01

*Note: ** indicates statistical significance at 5% level.*

The table shows that the pattern observed in the data is the following: there is statistically significant contemporaneous relationship between money supply growth and inflation when CC and M2 are used. Moreover, inflation appears to act both as leading indicator to money supply (for M2) and as a variable that is lagging money supply dynamics for all money aggregates examined. When performing the same dynamic cross-correlations with the core consumer price index, the pattern is somewhat similar, however there seems to be much more evidence of a two-way relationship between developments in money and inflation. On the whole, when core inflation is examined, money-price relationships are somewhat less clearly defined than those for headline inflation.

Dynamic cross-correlations between the growth rate of selected monetary variables (*t*) and HICP core inflation (*t+k*)
(first differences of the seasonally adjusted series)

	(q-8)	(q-7)	(q-6)	(q-5)	(q-4)	(q-3)	(q-2)	(q-1)	q	(q+1)	(q+2)	(q+3)	(q+4)	(q+5)	(q+6)	(q+7)	(q+8)
CC	-0.38**	-0.26**	-0.33**	-0.20	-0.38**	-0.14	-0.03	0.09	0.26**	0.21	0.30	0.26**	0.24	0.25**	0.14	0.11	0.12
M1	-0.45**	-0.26**	-0.37**	-0.31**	-0.39**	-0.38**	-0.17	-0.16	0.07	0.13	0.28**	0.37**	0.37**	0.43**	0.31**	0.20	0.28**
M2	-0.30**	-0.19	-0.26**	-0.21	-0.19	-0.17	0.00	0.04	0.21**	0.31**	0.40**	0.43**	0.31**	0.35**	0.23	0.06	0.14

*Note: ** indicates statistical significance at 5% level.*

A serious drawback of dynamic cross-correlations that is generally not addressed in the empirical literature is the possibility for a spurious effect on the estimated correlation coefficients in the case of autocorrelation in the input series. As Haugh (1976) and Haugh and Box (1977) note in their research on checking the independence of two covariance stationary time series, the autocorrelation present in each of the series can inflate the variance of cross-correlation estimates above that expected when cross-correlating two white-noise series. For overcoming the problem of obtaining a likely distorted cross-correlation function with misleading inference on the pattern of cross-correlations, Haugh (1976) suggests first obtaining the appropriate univariate (*e.g.* ARMA) models for each of the series and afterwards estimating the cross-correlation function of the residual white noise series obtained by fitting each of the separate univariate models.

We apply Haugh's methodology by first pre-whitening money supply growth and inflation data, implementing the Box-Jenkins methodology[3] to the first differences of the seasonally adjusted series, and then proceeding with the automatic calculation of the sample cross-correlation function between the residual white-noise series. We identify Box-Jenkins models by following the standard guidelines for the behaviour of the sample

[3] We apply the Box-Jenkins methodology as detailed in 'Forecasting, Time Series and Regression: An Applied Approach' by Bowerman, B., R. O'Connel and A. Koehler, Fourth Edition, 2005.

autocorrelation function and the sample partial autocorrelation function. The appropriate univariate models are constructed and the necessary diagnostic checks to ensure that the residuals behave as white-noise processes are performed (see Appendix B and Appendix C).

We examine the Ljung-Box (1978) Q-statistic for high-order serial correlation of the residuals from the fitted models. Since the degrees of freedom are reduced by the number of estimated coefficients, we do an adjustment of the degrees of freedom on the modified Q-statistic for small samples.[4] The diagnostic tests on the residuals of the fitted models show that for all models the estimated statistic, estimated for 8 lags, is not statistically significant at the 5% significance level, *i.e.* the null hypothesis of zero autocorrelation in the residuals cannot be rejected.

We proceed with performing a cross-correlation of the residual series obtained from the estimated models. The residual cross-correlograms, given respectively for 8 lags and 8 leads, are presented below. The significance of each of the reported residual cross-correlation estimates is determined in the standard way by 2 standard deviations bounds of ±2 /(√(T)), where T is the number of observations.[5]

Estimated univariate residual cross-correlation function for the growth rate of selected monetary variables (*t*) and HICP inflation (*t+k*)

	(q-8)	(q-7)	(q-6)	(q-5)	(q-4)	(q-3)	(q-2)	(q-1)	q	(q+1)	(q+2)	(q+3)	(q+4)	(q+5)	(q+6)	(q+7)	(q+8)
CC	-0.02	-0.22	-0.22	0.23	-0.16	-0.16	0.08	-0.09	0.37**	0.05	0.18	0.20	0.06	0.14	-0.22	-0.13	0.17
M1	-0.08	-0.11	-0.19	0.01	-0.14	-0.17	0.07	-0.18	0.22	0.01	0.12	0.28**	0.28**	0.18	-0.05	0.06	0.12
M2	-0.21	-0.05	-0.04	-0.10	0.10	-0.26	0.10	-0.11	0.22	0.14	-0.05	0.50**	0.01	0.18	0.08	0.11	0.03

Estimated univariate residual cross-correlation function for the growth rate of selected monetary variables (*t*) and HICP core inflation (*t+k*)

	(q-8)	(q-7)	(q-6)	(q-5)	(q-4)	(q-3)	(q-2)	(q-1)	(q)	(q+1)	(q+2)	(q+3)	(q+4)	(q+5)	(q+6)	(q+7)	(q+8)
CC	-0.04	-0.23	-0.15	-0.04	-0.05	0.06	0.02	0.18	0.00	0.32**	0.15	0.04	0.03	0.22	-0.16	-0.15	0.15
M1	-0.23	-0.14	0.00	-0.19	-0.21	0.11	-0.06	0.08	0.03	0.12	0.23	0.30**	0.22	0.31**	-0.11	-0.13	0.19
M2	-0.22	-0.04	0.03	-0.04	-0.05	0.04	-0.11	0.11	0.28**	0.21	0.09	0.14	0.20	0.25	0.03	-0.20	0.28**

*Note: ** indicates statistical significance at 5% level.*

[4] This modification in the statistic, which is used to test whether a chosen set of autocorrelations (*e.g.* for 8 quarters) in the residuals is significantly different from zero, is dependent on whether a constant has been included and as well as on the number of estimated autoregressive and moving average terms in the respective models. In each case, the value of the estimated modified Q-statistic is compared with the corresponding critical value from the chi-square distribution.

[5] As Haugh (1976) shows, the known asymptotic distribution of lagged cross-correlation estimates between two independent white noise series also applies to lagged cross-correlation estimates between the two residual series obtained by fitting univariate ARMA models to each of the two uncorrelated series.

We obtain the following results. For overall inflation and its relationship with the growth rate of currency in circulation, we observe a contemporaneous correlation. As regards the relationship between M1 and HICP inflation and M2 and HICP inflation, monetary aggregates lead inflation by about 3–4 quarters. Moving to the link between the selected monetary aggregates and core inflation, it can be seen that M2 has a significant contemporaneous relationship with core inflation and all monetary variables appear to have leading properties as regards inflation up to 8 quarters ahead. Furthermore, it does not appear that there is any feedback from inflation to monetary growth.

In addition to just comparing each cross-correlation coefficient with its respective standard error, we also apply Haugh's *s* statistic (1976) which presents a formal test of the null hypothesis that money supply dynamics and inflation series are unrelated, considering all examined cross-correlations. This statistic is computed as follows:

$$s = n \sum_{k=-M}^{M} r_{12}(k)^2 \text{ (for large samples);}$$

$$s^* = n^2 \sum_{k=-M}^{M} (n - |k|)^{-1} r_{12}(k)^2 \text{(for small samples)}$$

where: n is the number of time observations; $-M$ is the number of lags and M is the number of leads (M is chosen arbitrarily); $k = \pm\ 0,1,2,3...M$; r_{12} is the cross-correlation estimate at the specific lag/lead between the two residual series of length n, or when $k = 0$ (in case of the contemporaneous period).

In essence, the statistic is estimated for the first M residual cross-correlations and therefore M has to be determined by the researcher.[6] We estimate Haugh's *s* statistic for a different numbers of lags and leads of the cross-correlation function between the residuals from the fitted ARMA models to the series. The results, summarized in Appendix D, reveal a statistically significant feedback relationship between CC growth and HICP inflation, M1 growth rate and HICP core inflation, M2 growth and HICP inflation, and M2 growth and HICP core inflation. Furthermore, it can be seen that there is strong evidence for M1 growth leading core inflation and M2 growth leading overall inflation as well as core inflation.

[6] The critical value for this statistic is derived from the chi-square distribution (with $2M+1$ degrees of freedom for the case when k is from –M to +M and with M degrees of freedom when k is from 1 to M) under the null hypothesis that the series are uncorrelated at all lags considered.

The preliminary results presented in this section point to the conclusion that estimation methods based on the unidirectional assumption between money growth and inflation will probably be associated with inappropriate or even erroneous inferences. This is because when single equations are used for examining the association between the two variables, the OLS estimator will produce equations that are not identified correctly and precisely. We continue the empirical investigation by application of the standard Granger-causality tests to the original series of monetary aggregates growth and inflation.

3.2. Granger-causality Tests

This section is dedicated to the results obtained by an extensive Granger-causality test procedure using the OLS estimator in an unrestricted VAR framework. As indicated by the unit root tests, most of the series contain structural breaks, including in their first difference form. For this reason, we employ these tests with a rolling sample to tackle the problem of possibly instable coefficients, but also because we assume that structural changes that have likely taken place over the period from the late 1990s to 2012 can affect the relationship between the dynamics of a specific monetary aggregate and inflation throughout time. Therefore, that relationship and the possible existence of Granger causality can be different for different time periods. Two separate approaches are taken.[7] One is to use a sample with a rolling start, meaning that the end point of the sample is fixed for all tests (2012Q1), but the starting point is gradually moving forward in time. The other approach is to use a moving window sample with fixed length of eight years, meaning that both the start and endpoint of the sample are gradually moved forward.[8] Inflation and monetary variables are expressed in first differences of the levels (seasonally adjusted) in logs so as to achieve stationarity.

A critical step is determining the lag structure of each tested VAR system. On the one hand, choosing too few lags can result in omitted variable bias, which will produce a specification error and the estimation of incorrectly specified coefficients. On the other hand, using too many lags is

[7] For a description of possible rolling methods applied in the empirical literature, see for example Levasseur, S. (2008). The author provides definitions of the following techniques: 1) 'pure rolling method' – an arbitrarily chosen length of window (a fixed number of quarters) is rolled over the sample under investigation; 2) 'recursive method' – the starting date is kept fixed and the window size grows as the end date is approached; 3) 'reverse recursive method' – the end date is fixed and the window size shrinks as the distance from the starting date increases. Here we apply the 'reverse recursive method' and the 'pure rolling method'.

[8] Eight years is picked with the thought of having at least around 30 observations for each test.

wasting valuable observations and such specification error can result in the estimation of unreliable coefficients with high standard errors. Since the lag structure is crucial for the correct estimation of the tested system, the Akaike Information Criterion (AIC) and the Schwartz Criterion (SC) were consulted. However, the number of lags in some cases was determined subjectively, irrespective of the results from the two information criteria, in order to avoid autocorrelation in the residuals. Dummy variables were also used subjectively for each set of variables in order to fulfill the criterion of normality in the distribution of residuals.

The testing procedure involved estimating over 200 different unrestricted VAR models and a summary of the results is presented in Appendix E. For the first sampling method of a moving start of the sample period several tendencies are observed: for different sub-samples starting between 2000Q1 and 2002Q1 and ending in 2012Q1, there are statistically significant results pointing out to causality running from CC to headline inflation; however, the opposite causality is observed in some of the sub-samples as well; for sub-samples starting from 2000Q4 to 2004Q4, with a few exceptions, the results show causality running from M1 to headline inflation; testing for causality between M2 and headline inflation yields very mixed results; tests for CC and core inflation are mostly resulting in causality running from the former to the latter; unambiguous results also point out to causality running from M1 to core inflation almost throughout the whole period; testing M2 and core inflation, again, produces mixed results, but they mostly point out to either causality running from M2 to core inflation or to two-way causality.[9]

For the second method with a moving sample window the findings can be summarized as follows: testing headline inflation and CC yields mostly inconclusive results; significant results point out to causality running from headline inflation to M1 for the sub-samples 1998Q1–2006Q1, 1999Q1–2007Q1 and 2000Q1–2008Q1; for three other later sub-samples causality goes from M1 to overall inflation; for the sub-sample 1998Q1–2006Q1 results point out to causality running from headline inflation to M2, however for two other later sub-samples we find causality from M2 to headline inflation; testing core inflation and CC shows that for three sub-samples currency in circulation granger causes core inflation; significant results point out that causality runs from M1 to core inflation for

[9] When we use the word 'causality' for short, we actually mean Granger causality and we do not refer to the existence of an actual cause and effect relationship, that is we refer to the traditional Granger causality understanding in the sense of one variable preceding another variable in time.

four of the seven sub-samples; results from tests of M2 and core inflation yield inconclusive results and no causality in most of the sub-samples.

It is evident that the two methods have some differences in their findings. However, a few common results can be outlined: both sampling methods yield test results on currency in circulation and core inflation pointing out to causality running from the former to the latter; causality from the monetary aggregate to core inflation as well as to headline inflation is found for tests on M1, with this causality being stronger for the case of core inflation; tests involving M2 yield very mixed results whether headline or core inflation is used and often the tests show two-way causality. It should be also noted that sub-samples starting more recently generally show causality running from the monetary aggregates to inflation, while the opposite is true for sub-samples starting earlier. Therefore, our testing procedure proves that granger causality exists and in the predominant number of cases causality goes from monetary growth to inflation. However, the nature of that causality changes throughout time and this result may be probably attributed to the structural breaks that were revealed for most of the series in the unit root tests. Furthermore, significant evidence of two-way causality between M2 and inflation is also found.

3.3. Issues for Robustness of Granger-causality Tests

Formulating conclusions based on the Granger-causality test results described above needs to be necessarily accompanied with an appropriate assessment of possible cointegration. In the presence of the latter, the inclusion of 'error-correction terms' in Granger-causality tests is required so as to avoid misspecification due to an omitted variable bias problem related to cointegration. There are two possible econometric options for dealing with potential cointegration between monetary aggregates and prices/inflation. The first one is to re-estimate Granger-causality tests, including where necessary co-integration links. The second option is to apply the system approach framework of Johansen (1990, 1991) which enables testing for more than one cointegrating relationship in a vector autoregressive model, possibly including not only money and prices, but also all standard variables related for example to money demand investigation in a multivariate analysis. One of the main weaknesses of the first option is that even if one omitted variable bias problem can be addressed by adding cointegrating term for the two variables, there may be other potential serious problems causing misspecification such as the issue related to omitting relevant explanatory variables in the cointegration relation or omitting whole cointegrating vectors.

Our preference for applying the Johansen procedure is driven by the following main reasons. First, this procedure is the natural way of incorporating long-run relations, one of them in the case of money and prices/inflation is from a theoretical point of view expected to be naturally related to the estimation of a full money demand function. The adjustment of monetary aggregates to the long-run money demand function can be interpreted as money supply reaction to money demand, provided the estimated model is correctly specified with all relevant variables for a money demand function. Furthermore, possible adjustment of price variables can be thought of as the case when any short-run disequilibrium between money supply and demand results in excess money supply potentially acting as a pro-inflationary factor which represents the second hypothesis postulated to be tested in the paper. In addition, the short-run dynamics of the VECM allows for testing the impact of past values of monetary variables growth on inflation (as well as the impact of past values of inflation on monetary variables growth) once any feedback from disequilibrium in the cointegrating relationship has been taken into account. This can also help investigate further the first hypothesis of the paper.

4. Estimates of Money Demand Functions with VECM

4.1. Brief Overview of the Theoretical Framework

The general theoretical framework for modeling money demand in a specification relevant for the long-run horizon can be summarized as given below:

$$\frac{M}{P} = f(Y, R) \tag{1}$$

where M represents the monetary aggregate in nominal terms; P is the price level measured usually either by the consumer price index (HICP/CPI) or the GDP deflator; Y is the scale variable characterizing the transaction/precautionary motive of holding money and could be represented by variables such as real consumption expenditures of households, real wages or real retail sales (when demand for currency in circulation is modeled) and real GDP or industrial production (when broader monetary aggregates are modeled); R represents a vector of returns, including *own rate* of money which is to be chosen depending on the composition of the specific monetary aggregate as well as different indicators that proxy the opportunity cost of holding money for the respective monetary aggregate

such as the return on assets *outside of money* and the (expected) rate of inflation, with the latter representing the opportunity cost of holding money rather than goods or real assets in general. According to the standard theories of money demand, the function f is increasing in Y, increasing in the components of R that relate to assets included in M and respectively decreasing in the components of R that relate to assets outside of M.

In the cointegrated VAR modeling that we are going to employ we treat prices as an endogenous element to the system of analyzed variables and use (1) in the following form:

$$M = f(P, Y, R) \tag{2}$$

The endogenous treatment of the price level allows us to test whether the theoretical hypothesis for long-run unit price elasticity is supported by the data without imposing this restriction beforehand as in (1). We thus depart from the general specification structure of most empirical studies which assume long-run price homogeneity and additionally assume the validity of the hypothesis that money supply is not an important factor affecting price dynamics, thus justifying the omission of prices from the vector of endogenous variables in (1). By including the price level among the set of endogenous variables and attempting to model nominal rather than real money demand, our approach broadly resembles that of Komárek and Melecký (2001) who construct a money demand model for the Czech Republic. The characteristic feature of their paper, presented in the literature review section, is the importance that the authors attach to transition-specific factors and one example of such factors is the preference for modeling money demand in nominal terms.[10]

In a log-linear form the model we estimate for CC, M1 and M2 is as follows:

$$m^d = \gamma_0 + \gamma_1 y + \gamma_2 R^{\text{out}} + \gamma_3 R^{\text{own}} + (\gamma_4 \Delta p) + \gamma_5 p$$

Expected signs: $\gamma_1 = 1$, $\gamma_2 \leq 0$, $\gamma_3 \geq 0$, $\gamma_4 \leq 0$, $\gamma_5 = 1$

4.2. Description of the Data and Methodology

This part of the paper reports the results of money demand models for CC, M1 and M2 obtained by applying the Johansen cointegrated VAR approach. We investigate a large dataset for the analysis of the three

[10] The authors acknowledge that for transition economies there is uncertainty with respect to the specific price variable to be used as a deflator for estimating real money holdings due to factors such as the severe shocks related to deregulation of prices, transformation of tax policies, liberalization of foreign trade, *etc.* that have occurred, affecting the specific evolution of different measures of price developments over time.

measures of money, being aware of the fact that the specific choices of explanatory variables representing Y, P and R in (2) may significantly affect the outcome of the cointegration test for money demand relations.[11]

Based on the detailed literature review, we select the following possible variables as the dataset to be tested for modeling CC money demand: HICP (as the price variable); real expenditures of households, real wages, real retail sales (as proxies for real sector developments); and the BGN/USD exchange rate and the interest rate on new term deposits (as opportunity costs of holding currency in circulation). The dataset for modeling M1 covers a broader range of variables: HICP, GDP deflator and producer price index (PPI) (as possible price variables); real GDP, industrial production, real retail sales (for approximation of real income) and FDI as a variable reflecting the substantial amount of foreign direct investment in the Bulgarian economy with a potential influence on demand for M1; interest rates on overnight deposits (as a proxy for the own rate of money due to the relatively large share of overnight deposits in the composition of M1); the interest rate on new term deposits, the BGN/USD exchange rate and the 3-month money market rate SOFIBOR (as proxies for the opportunity cost of holding M1). The dataset for modeling the broad monetary aggregate includes: HICP, GDP deflator and PPI; real GDP, industrial production; the interest rate on new term deposits (as a proxy for the own rate of money); the interest rate on loans, the 3-month money market rate SOFIBOR, the 3-month EURIBOR, the BGN/USD exchange rate, the 10-year Bulgarian government bond yield and the 10-year euro area government bond yield (as possible proxies for the opportunity cost of demand for M2). Furthermore, for the specifications of all money aggregates, it is additionally investigated whether the inflation rate could enter as a long-run determinant of money demand in its role of an opportunity cost of holding money rather than goods or real assets.

A few notes need to be made regarding the selection of variables. The initial period of low monetization immediately following the introduction of the currency board arrangement, the gradual development of the banking sector, the stock market and the government bonds market, implying an initial period with practical absence of alternatives for storing wealth, the changes in the composition of monetary aggregates, the process of

[11] See Ericsson (1998) for a thorough discussion of the key issues in the empirical modeling of money demand, including data measurement and data choice problems, parameter constancy, the adequate construction of the opportunity cost of holding money as well as various specification issues that have to be considered. Ericsson argues that in testing for cointegration across different measures of money, scale variables, and rates of return, only a few combinations might be cointegrated, even if the underlying cointegration relationship exists.

financial innovation, the presence of grey economy are all reasons why it may be difficult to find appropriate indicators for opportunity costs of holding money. For this reason, we start off with an investigation of a relatively large dataset of possible explanatory variables in the money demand models. It is worth noting that among the potential measures for the opportunity cost of holding M2 along with traditional indicators in the economic literature we add the interest rate on loans as suggested by Komárek and Melecký (2001) who analyzed broad money demand in the Czech Republic and Dabušinskas (2005) who estimated a broad money demand model for Estonia.

The period of empirical analysis is from 1998Q1 to 2012Q1. However, for most of the models investigated the period starts from 2000Q1 and for some from 2003Q1 because of data availability. We work with seasonally adjusted quarterly data, with the adjustment performed in Eviews using the Tramo/Seats procedure with the automated setup for the diagnostics and optimization. The procedure does not identify seasonality for interest rates and the producer price index. The aggregation for the data which is published on a monthly frequency is obtained by averaging. The only exception are the series of the monetary aggregates, for which as explained in Section 3, we take an end-of-period transformation.

Estimations of money demand models within the Johansen framework are conducted without testing the order of integration of the variables as this is found not to be a necessary prerequisite for the analysis.[12] We have identified in Section 3 that monetary aggregates as well as the HICP are $I(1)$ processes in levels and in addition, unit root tests applied to the other two price indicators employed in the analysis (the GDP deflator and the PPI) show that these series are also integrated of order one (see Appendix A1 and A2). Consequently, no exhaustive examination of the stochastic time-series properties of the rest of the variables is performed.

The initial VAR models that are tested include levels of the variables expressed in logs with only interest rates expressed in percentages. A few variables such as the BGN/USD exchange rate, the 3-month EURIBOR, the 3-month SOFIBOR and the FDI stock are treated as exogenous both because such an assumption appears plausible for most of them, but also because treating them as endogenous usually produced very large systems to be estimated and made the application of the Johansen method unreliable due to the relatively small dataset. The inclusion of the inflation rate

[12] As Johansen (1995), cited in Ahking (2002), argues in case where there are at least two non-stationary variables in the cointegration model that are integrated of the same order, the other variables could be both stationary and/or trend-stationary.

in the VAR models is achieved either by considering the annual rate of inflation or by taking lags of the first difference of the log-level of the harmonized consumer price index.

The selection of the lag length is done in accordance with the information criteria (maximum lag is set to 4), with strongest reliance given to the Schwarz criterion due to the small sample. The final decision on the lag length of the respective model is made once a well-specified model with serially-uncorrelated and normally-distributed residuals is obtained and the VAR displays dynamic stability. When normality of residuals is difficult to achieve, the minimum requirement is that the distribution of residuals satisfies the assumption of non-skewness.[13] We also implement the adjustment of the critical values of the trace test as a correction for finite sample sizes, as proposed by Reinsel and Ahn (1988) and cited in Cheung and Lai (1993).

As it is known, the power of the Johansen cointegration test may weaken when exogenous series are added. We therefore choose the specific number of cointegrating vectors for each model investigated based on the trace test in combination with the economic interpretability of results. Once the cointegrating rank is determined, the specific number of cointegrating vectors is imposed on the VAR system based on the general assumption that there are linear trends in the data, while the cointegration relation(s) contain a non-zero intercept but no deterministic trends. Estimation procedure then continues with imposing restrictions on the cointegrating vector(s) so as to ensure their unique determination. Tests for significance and stationarity of the variables in the cointegrating relation(s) are performed. Then restrictions based on the postulations of money demand theory for one of the cointegrating vectors are imposed and in case a second vector is determined, other restrictions such as those related to the term structure between interest rates or to aggregate demand relations have been tested. Weak exogeneity of some of the variables is also investigated. As a final step, specification tests are performed.

4.3. Results for Currency in Circulation

The empirical analysis of the currency in circulation established a VECM for the following variables: currency in circulation (cc), HICP (hicp), real retail sales[14] (tradec) and the interest rate on new term deposits (m_tdir_n). The

[13] We follow the recommendations by Hendry and Juselius (2000) who show that statistical inferences are valid even if there is empirical evidence for excess kurtosis and residual heteroscedasticity. The authors also suggest using the trace test for determining the number of cointegrating vectors as a more robust indicator than the maximum eigenvalue test.

[14] Total turnover index of retail sales, except of motor vehicles and motorcycles at constant prices

variables are seasonally adjusted and expressed in logarithm, apart from the interest rate variable. The inclusion of the inflation rate as an additional opportunity cost produced a model with unstable coefficients.

The unrestricted VAR passed successfully specification tests for residual autocorrelation, normality of residuals and heteroscedasticity when the model included two lags as well as two dummy variables.[15] The test for the VAR stability condition check pointed out to dynamic stability of the investigated model. The Johansen cointegration trace test results, given in Appendix F1 (Section 1), reveal an empirical evidence for 2 cointegrating relationships. We thus proceed with imposing rank = 2 on the unrestricted model and make the assumption that there are linear trends in the data series with no deterministic trends in any cointegrating relations.[16] We succeed in obtaining unique specifications for these cointegrating relationships with the following over-identifying restrictions which are jointly accepted (LR test of restrictions: Chi^2(2) = 0.484 [0.7850])[17]:

cointegrating vectors	β_1	β_2
m_tdir_n	0.073 (0.009)	0.112 (0.018)
log(cc_sa)	1.0	0
log(hicp_sa)	-1.0	1.0
log(tradec_sa)	-0.885 (0.019)	-1.0

adjustment coefficients	α_1	α_2
m_tdir_n	-1.179 (0.883)	-1.127 (0.305)
log(cc_sa)	-0.507 (0.086)	0
log(hicp_sa)	-0.183 (0.045)	-0.089 (0.017)
log(tradec_sa)	-0.119 (0.062)	-0.065 (0.024)

Note: standard errors in parentheses.

The first cointegration relationship, which can be seen in Column 2 of the table with the beta coefficients, can be interpreted as a money demand function. In the long run currency in circulation is driven by the increase in the price level (with accepted long-run price homogeneity) and in the retail sales volumes (coefficient of 0.885, slightly less than the unit elasticity as postulated by the quantity theory of money) and is influenced negatively by the interest rate on new term deposits (the semi-elasticity stands at -0.073). These results are in line with theoretical considerations. The generally low interest rate elasticity also corresponds to empirical results for EU New Member States. The loading coefficient for the *cc equation,* which is presented in the second column of the table with the alpha coefficients,

[15] For the fourth quarter of 2000 and the fourth quarter of 2008 respectively.

[16] We tested the hypothesis whether a deterministic trend can be included in the cointegration relations, but that hypothesis was rejected.

[17] The asymptotic null distribution of the test is denoted Chi^2(·) with degrees of freedom in parentheses and the asymptotic p-value given in the brackets.

is negative at -0.51 which points to relatively fast adjustment of *cc* to short-run deviations from long-run equilibrium. This corresponds to expectations as under the currency board demand and supply of currency in circulation should be balanced in a fast and automatic way. The loading coefficient for the *hicp* is negative at -0.18 which does not support the hypothesis that in cases of an excessive supply of currency in circulation inflation increases.

The second cointegration relationship is set for the *hicp* (the normalization is done on this variable). In the long run *hicp* is related positively to the real retail sales (a permanent increase in real retail sales by 1% increases the consumer price index by 1%) and negatively to the interest rate on new term deposits (an increase of interest rates by 1 percentage point decreases the hicp index by 0.11%). The loading coefficient for the *hicp* equation in this cointegration relationship is as expected negative and shows relatively slow adjustment of prices to disequilibrium.

The results from the recursively estimated LR test of restrictions, presented in Appendix F1 (Section 2), show that the imposed restrictions are valid for the whole period of estimation which attests to the stability of the results. In the addition, the recursively estimated unrestricted coefficients (on the interest rate and on the real retail sales in the first cointegration equation and on the interest rate in the second cointegration equation) are also stable (Section 3).

We then transform the model in a stationary form (that is into a model with the respective error-correction terms combined with short-run dynamics of the variables), being particularly interested whether the short-run dynamics of the *hicp* variable is influenced by the dynamics of the *cc* variable. We obtain a statistically significant positive coefficient of 0.21, showing that while the dynamics of *cc* is affected by the development of prices in the long run, there is also a feedback influence from *cc* to the dynamics of prices (see Appendix F1, Section 4). Furthermore, the equations from the short-run error correction model successfully pass all residual diagnostic tests both when run for the case of single equations and for the vector version of the tests (Section 5). The recursively estimated 1-step residuals of the model for each of the four variables display no problems with instability (Section 6).

4.4. Results for M1

For the monetary aggregate M1 we obtain an empirical evidence for a cointegrating relationship with the following variables: producer price index (ppi), industrial production index (indp) and the interest rate on new term deposits (m_tdir_n). M1 and industrial production are seasonally

adjusted and expressed in logarithm, no seasonal adjustment was required for the producer price index and this index is only expressed in logarithm, whereas for the interest rate variable no transformation was applied.

The cointegration of *m1* with *ppi* as the price variable can be accounted for by the fact that a predominant part of the composition of M1 consists of overnight deposits that are hold mostly by non-financial corporations rather than households over the period under investigation. An additional money demand determinant that was also explored was the exchange rate BGN/USD (*er)* which could potentially influence M1 holdings for example in the case when a depreciation of the national currency against the US dollar could theoretically be associated with increasing external demand for Bulgarian goods, respectively higher domestic industrial production, higher domestic inflation rate and a need for more money in the economy because of the larger amount of transactions.

We obtain a well-specified model with a VAR (1) when including as exogenous variables the second lag of the log-level of *er (log(er(-2))*, the first difference of the log-level of the index of Brent oil prices measured in euro (*oil_p*): *dlog(oil_p)* (to account for residual autocorrelation in the *ppi* equation) and two dummy variables[18]. This model produced the required assumptions for normality of residuals, lack of serial autocorrelation as well as for dynamic stability. In addition, the adequacy of the chosen model was supported by the lack of residual heteroscedasticity both for the single equations and for the overall model (based on the vector heteroscedasticity test). The application of the Johansen cointegration trace test and the small-sample correction of the trace statistic, as discussed in Section 4.2, provided evidence for one cointegrating vector (see Appendix F2, Section 1).

Imposing one cointegrating vector in the unrestricted VAR model and working with the assumption about linear trends in the data, without any linear trends in the long-run relation, we proceed by testing hypotheses about the statistical significance of the variables in the long-run relation.[19] The results from the tests (not reported here for reasons of brevity) showed that the coefficient on the interest rate is not significantly different from zero. This fact is in contrast with theoretical postulates about a negative coefficient as the interest rate on new term deposits represents an opportunity cost of holding money in the form of the two subcomponents of M1. For the purpose of ensuring an identified model which can be estimated, we start

[18] The dummy variables are for the third quarter of 2008 and the first quarter of 2009 respectively.

[19] We also estimated the same model with one cointegrating vector based on the assumption that a restricted trend lies in the cointegration space, but in this model the trend was not statistically significant.

with the hypothesis of an existing money demand function and impose the standard theoretical restriction of unit price elasticity in addition to the restriction of zero interest rate elasticity. While these two restrictions are not rejected, we explore the statistical significance of the loading coefficients in the *ppi* and *indp* equations and subsequently apply two other restrictions – for weak exogeneity of *ppi* and *indp*. The joint LR test for all these over-identifying restrictions shows that the restrictions cannot be rejected (LR test of restrictions: Chi^2(4) = 6.119 [0.1904]). The obtained cointegrated VAR model with these restrictions is presented below:

cointegrating vector	β_1
m_tdir_n	0
log(m1_sa)	1.0
log(ppi)	-1.0
log(indp_sa)	-2.385 (0.189)

adjustment coefficients	α_1
m_tdir_n	-1.206 (0.261)
log(m1_sa)	-0.159 (0.040)
log(ppi)	0.0
log(indp_sa)	0.0

Note: standard errors in parentheses.

As can be seen, in the long-run income elasticity exceeds one (2.385), which is similar to the results generally obtained for EU New Member States. The established cointegrating vector can be interpreted as a money demand function, which is also supported by the negative adjustment coefficient in the *m1* equation (-0.159). The interest rate on new term deposits adjusts to deviations from the long-run relation with a negative coefficient, implying that any excess money supply brings about a decrease in the interest rate. Based on the weak exogeneity of *ppi* in this model and the lack of effects from lagged differences of *m1* on *ppi* dynamics (due to the lack of first-differenced terms in the vector-error correction model), we can conclude that there is no feedback effect from *m1* to *ppi*.

The results from the recursively estimated LR test for the four restrictions discussed above, given in Appendix F2, Section 2, contest to the overall stability of restrictions with the exception of the period around the end of 2009 and the beginning of 2010. For that period we can tentatively argue that the downturn of the Bulgarian economy might have resulted in changes in the behavior of economic agents and this could potentially destroy the validity of the restrictions on the model. At the same time, however, the recursively estimated unrestricted coefficient for the income elasticity displays overall stability as shown in Appendix F2, Section 3.

Transforming the VECM model into a stationary short-run form, we note additional results. The rate of increase in the price of oil has a statistically significant positive influence on the *ppi* index, which can be accounted for by the relatively large weight of energy in the producer price index. The loading coefficient for *ppi* is negative (at -0.024) which does not lend support to the hypothesis of a pro-inflationary impact of excess supply of M1 (Appendix F2, Section 4). In Section 5 we report all required residual diagnostic tests and it can be seen that all test results are in general satisfactory.

4.5. Results for M2

Econometric modeling for a cointegrating relationship including M2 led to empirical evidence for a model including M2, the producer price index (*ppi*), the industrial production index (*indp*), the interest rate on new term deposits (*m_tdir_n*) and the interest rate on loans for new business (*ir_loans*). The endogenous variables were taken in logs and seasonally adjusted, apart from the two interest rates variables. No seasonal adjustment was required for the *ppi*.

To obtain a well-specified model we considered a number of alternatives for the lag structure as well as a number of dummy variables and additional exogenous variables to correct mostly for normality of residuals (targeting at the least at their non-skewness), as lack of residual serial correlation was hard to attain. The dynamic stability properties of the initial unrestricted VAR were also a requirement that had to be satisfied. The final specification included 2 lags of the endogenous variables and as exogenous to the system the following variables: the first lag of the exchange rate BGN/USD *(log(er(-1))*, the first difference of the logarithm of Brent crude oil price (in euro): *dlog(oil_p)*, the first difference of the logarithm of real GDP in the euro area (seasonally adjusted): *dlog(rgdp_ea_sa)* as well as dummy variables.[20] This specification produced a statistically adequate unrestricted VAR model.

For this model the Johansen cointegration trace test revealed the presence of 1 cointegrating vector (see Appendix F3, Section 1). Obtaining unique specification for the cointegrating relationship is achieved with the following over-identifying restrictions which are jointly accepted (LR test of restrictions: Chi^2(3) = 0.796 [0.8504]):

[20] The following dummy variables were included: for Q1 2005, Q2 2005, Q1 2007, Q4 2008 and Q1 2009.

cointegrating vector	β_1
m_tdir_n	-0.072 (0.011)
ir_loans	0.068 (0.012)
log(indp_sa)	-1.0
log(m2_sa)	1.0
log(ppi)	-1.454 (0.093)

adjustment coefficients	α_1
m_tdir_n	-1.026 (0.711)
ir_loans	0
log(indp_sa)	0.177 (0.075)
log(m2_sa)	-0.165 (0.038)
log(ppi)	0

Note: standard errors in parentheses.

The imposed restriction in the cointegrating relationship follows the conventional money demand models and is related to unit income elasticity. Additionally, the adjustment coefficients in the equations for *ir_loans* and for *ppi* are set to zero. For the remaining unrestricted coefficients, we note that *m_tdir_n* enters the long-run relationship as expected with a positive sign and *ir_loans* has the expected negative coefficient. Long-run price elasticity comes to 1.454. As regards the adjustment coefficients, we can observe the expected negative coefficient for the M2 equation. The industrial production index responds positively to deviation of M2 supply from the relationship with the long-run components of money demand, meaning that 'excess money supply' affects economic performance. The results from the recursively estimated LR test for the three restrictions show stability properties (Appendix F3, Section 2). The recursively estimated unrestricted coefficients in the cointegrating relationship also display overall stability properties with a slight instability observed at the end of 2009 (see Section 3).

Transforming the cointegrated VAR model to a system with $I(0)$ variables leads us to the model equations presented in Section 4. In the short-run equation for the *ppi* the previous period growth of M2 has a statistically significant positive effect. The equations from the short-run error correction models pass successfully almost all residual diagnostic tests: both when run for each of the single equations and as in the vector version form (see Appendix F3, Section 5). The only tests that are not satisfied are the test for residual autocorrelation in the industrial production equation, the test for heteroscedasticity in the equation for the interest rate on new term deposits as well as the vector tests for normality and heteroscedasticity. The recursively estimated 1-step residuals from the five short-run equations are constant as shown in the graphs in Section 6.

5. Conclusions

This paper has undertaken an extensive empirical analysis of the relationship between money supply dynamics and inflation, continuing and further deepening the research conducted in the first years of the introduction of the currency board arrangement in Bulgaria.

The first hypothesis tested in the paper was whether a feedback relationship exists from money aggregates dynamics on inflation. The results based on univariate residual cross-correlation approach point out to the fact that any estimation methods based only on the unidirectional assumption (from prices to money supply) are most probably associated with inappropriate or even erroneous inferences. The application of Granger-causality tests shows that granger causality exists in both directions and in the predominant number of cases causality goes from monetary growth to inflation. There is evidence that a one-way direction of relationship exists from M1 growth to HICP inflation and this relationship is particularly strong from M1 growth to HICP core inflation. We also obtain results on currency in circulation and HICP core inflation that generally point out to causality running from the former to the latter, with this evidence being somewhat weaker when currency in circulation and headline inflation are analyzed.

Apart from the new conclusions from empirical econometric analysis on the relationship between money supply dynamics and inflation, an additional important contribution of the paper includes the results from cointegrated VARs related to the estimation of money demand models for currency in circulation, M1 and M2. Our main conclusions about money demand models can be summarized as follows.

We obtain a stable long-run money demand relation for currency in circulation, HICP, retail trade turnover at constant prices and the interest rate on new term deposits, with all coefficients having the theoretically expected signs and unit price elasticity being confirmed. The adjustment of currency in circulation to deviations from long-run equilibrium is relatively fast (adjustment coefficient at around -0.5) which is in line with expectations as under the currency board arrangement demand and supply of currency in circulation should be balanced in a fast and automatic way. The short-run dynamics reveals feedback influence from money to price dynamics. The overall parameter constancy tests of the model do not show any problems regarding changes in the sample period.

Concerning M1 and M2, cointegrating money demand relationships were more difficult to obtain and the models explored remain with some specification problems especially with respect to stability of coefficients and

the choice of variables to be included. We obtain evidence for a long-run link among M1, the producer price index, the industrial production index and the interest rate on new term deposits. In this relationship unit price elasticity is confirmed, whereas the income elasticity of money demand, in line with findings for other new EU Member States, is higher than 1 (2.4). The main issues with the model relate to the insignificance of the interest rate variable in the long-run link and the instability of the cointegrating relationship at the end of 2009 which needs to be further investigated. The model does not lend support for a possible feedback effect from the short-run dynamics of M1 back to that of producer prices.

For M2 we reach the conclusion of a long-run money demand relationship with the producer price index, the industrial production index, the interest rate on new term deposits and the interest rate on loans for new business. All of the coefficients have the theoretically expected signs, unit income elasticity of money demand is obtained, while the long-run price elasticity is slightly higher than 1 (1.4). In the short-run equation for the producer price index, the previous period growth of M2 has a positive effect on the dynamics of producer price inflation. Similar to the model for M1, this model displays certain instability at the end of 2009.

Finally, regarding the second hypothesis investigated in the paper whether excess money supply in the short-run might act as a pro-inflationary factor, the empirical analysis based on the cointegrated VAR models does not lend support for such an effect for any of the monetary variables analyzed.

To summarize, the results of the paper point to the general conclusion of a two-way relationship between money supply and price dynamics in Bulgaria. The link from prices through money demand to money supply may be theoretically justified because of the endogeneity of the money supply mechanism under the currency board arrangement. At the same time, the paper also finds evidence for feedback effects from monetary growth to inflation.

Future areas of research on the relationship between money supply dynamics and inflation include: 1) exploring in more detail the properties of M1 and currency in circulation for improving forecasts of core inflation; 2) the application of Johansen cointegration analysis with structural instability to deal with the identified structural breaks in most monetary aggregates and price variables series as well as with the evidence for stability problems of money demand functions for the cases of M1 and M2; 3) the extension of estimation techniques for money demand models with methods such as the single-equation techniques of DOLS and DGLS with the aim of providing evidence for the robustness of results.

References

Ahking, F. (2002). Model Mis-specification and Johansen's Co-integration Analysis: An Application to the US Money Demand, *Journal of Macroeconomics* 24 (2002), pp. 51–66.

Arratibel *et al.* (2009). Inflation Forecasting in the New EU Member States, *ECB Working Paper Series*, No. 1015/ February 2009.

Beck *et al.* (2003). Inflation and the Bulgarian Currency Board, *Bulgarian National Bank Discussion Papers*, May 2003.

Bowerman, B., R. O'Connel and A. Koehler (2005). Forecasting, Time Series and Regression: An Applied Approach, Fourth Edition.

Cheung, Yin-Wong and Kon S. Lai (1993). Finite Sample Sizes of Johansen's Likelihood Ratio Tests for Cointegration, *Oxford Bulletin of Economics and Statistics*, 55, 3 (1993).

Dabušinskas, A. (2005). Money and Prices in Estonia, *Working Papers of Eesti Pank*, No. 7/2005.

Dreger *et al.* (2006). Long-run Money Demand in the New EU Member States with Exchange Rates Effects, *ECB Working Paper Series*, No. 628/ May 2006.

Enders, W. Applied Econometric Time Series, Second Edition, pp. 67–69.

Ericsson, N. (1998). Empirical Modeling of Money Demand, *Empirical Economics* (1998), 23, pp. 295–315.

Haugh, L. (1976). Checking the Independence of Two Covariance-Stationary Time Series: A Univariate Residual Cross-Correlation Approach, *Journal of the American Statistical Association*, Vol. 71, No. 354, pp. 378–385.

Haugh and Box (1977). Identification of Dynamic Regression (Distributed Lag) Models Connecting Two Time Series, *Journal of the American Statistical Association*, Vol. 72, No. 357, pp. 121–130.

Hendry, D. and K. Juselius (2000). Explaining Cointegration Analysis: Part II, University of Copenhagen, Department of Economics, *Discussion Papers*, November 2000.

Johansen, S. (1991). Estimation and Hypothesis Testing of Cointegration Vectors in Gaussian Vector Autoregressive Models, *Econometrica*, Vol. 59, No. 6, pp. 1551–1580.

Johansen, S. and K. Juselius (1990). Maximum Likelihood Estimation and Inference on Cointegration – with Applications to the Demand for Money, *Oxford Bulletin of Economics and Statistics*, 52, pp. 169–210.

Komárek and Melecký (2001). Demand for Money in the Transition Economy: The case of the Czech Republic 1993–2001, *Warwick Economic Research Papers*, No. 614/ December 2001.

Krus, N. (2012). The Money Supply in Currency Boards, Johns Hopkins Institute for Applied Economics, Global Health, and Study of Business Enterprise, Studies in Applied Economics, No. 3, November 2012.

Levasseur, S. (2008). Updating Empirical Evidence on Business Cycles Synchronization Between CEECs and the Euro area: How important is The Recent Period, OFCE, No. 2008-11, April 2008.

Ljung, G. and G. Box. On a Measure of Lack of Fit in Time Series Models, *Biometrika* (1978), 65, 2, pp. 297–303.

Nenovsky, N. and K. Hristov (2000). Currency Circulation after Currency Board Introduction in Bulgaria, *Bulgarian National Bank Discussion Papers*, October 2000.

Nenovsky, N. and K. Hristov (2002). New Currency Boards and Discretion: Empirical Evidence from Bulgaria, *Economic Systems*, XXVI (1): pp. 55–72.

Nenovsky, N., K. Hristov and M. Mihaylov (2001). A Simple Test of Currency Board Automatic Mechanism in Bulgaria, Estonia and Lithuania, *Journal des Economistes et des Etudes Humaines*, 11 (4): pp. 575–616.

Otero and Smith (2000). Testing for Cointegration: Power versus Frequency of Observation – Further Monte Carlo Results, *Economics Letters* 67 (2000), pp. 5–9.

Pikkani, R. (2000). The Monetary Sector under a Currency Board Arrangement: Specification and Estimation of a Model with Estonian Data, Bank of Finland, Institute for Economies in Transition, No. 10/2000.

Reimers, H. and K. Roht (2007). Estimates of Money Demand Functions for Estonia, *Transition Studies Review* (2007) 14 (3): pp. 425–439.

Sepp, U. (1995). Inflation in Estonia: The Effect of Transition, Review of Economies in Transition, Bank of Finland, Institute for Economies in Transition, No.2, 1995.

Sepp, U. (1995). A Note on Inflation under the Estonian Currency Board, Review of Economies in Transition, Bank of Finland, Institute for Economies in Transition No. 4, 1995.

Siliverstovs, B. (2007). Money Demand in Estonia, *DIW Berlin Discussion Papers*, No. 675/ March 2007.

Slavova, S. (2003). Money Demand During Hyperinflation and Stabilization: Bulgaria, 1991–2000, *Applied Economics*, Volume 35, Issue 11, 2003.

Schwert, G. William (1979). Tests of Causality: The Message in the Innovations, Carnegie-Rochester Conference Series on Public Policy, 1979, Volume 10, pp. 55–96.

Tillers, I. (2004). Money Demand in Latvia, *Bank of Latvia Working Papers*, 3/2004.

Yotzov, V. (2000). A Possible Approach to Simulate Macroeconomic Development in Bulgaria, *Bulgarian National Bank Discussion Papers*, November 2000.

Yotzov *et al.* (1998). The First year of the Currency Board in Bulgaria, *Bulgarian National Bank Discussion Papers*, August 1998.

Appendix

Appendix A1: Results from unit root tests without structural breaks

Level	Test specification	ADF (AIC)	ADF (BIC)	PP	KPSS	First difference	Test specification	ADF (AIC)	ADF (BIC)	PP	KPSS	Order of integration
log(cc_sa)	constant	-2.43	-2.70 **	-2.94 **	0.88 ***	dlog(cc_sa)	constant	-3.26 **	-3.26 **	-6.02 ***	0.58 **	I(1)
	constant and trend	-0.23	0.30	0.01	0.21 **							
log(m1_sa)	constant	-1.25	-1.25	1.38	0.89 ***	dlog(m1_sa)	constant	-3.03 **	-3.03 **	-6.07 ***	0.25	I(1)
	constant and trend	-1.44	-1.44	0.79	0.16 **							
log(m2_sa)	constant	-1.52	-1.52	-1.18	0.91 ***	dog(m2_sa)	constant	-2.32	-2.32	-6.43 ***	0.24	I(1)
	constant and trend	-1.56	0.15	-0.64	0.17 **							
log(hicp_sa)	constant	-1.58	-0.76	-0.65	0.91 ***	dlog(hicp_sa)	constant	-3.63 ***	-4.95 ***	-4.99 ***	0.10	I(1)
	constant and trend	-2.53	-2.06	-1.94	0.07							
log(hicp_c_sa)	constant	-0.91	-0.17	-0.89	0.88 ***	dlog(hicp_c_sa)	constant	-2.20	-3.76 ***	-4.15 ***	0.11	I(1)
	constant and trend	-3.87 **	-2.31	-1.62	0.12 *							
log(ppi)	constant	0.03	0.03	0.11	0.93 ***	dlog(ppi)	constant	-4.70 ***	-4.70 ***	-4.76 ***	0.08	I(1)
	constant and trend	-2.78	-2.78	-2.15	0.11							
log(gdp_p_sa)	constant	0.91	-0.19	-0.12	0.91 ***	dlog(gdp_p_sa)	constant	-8.96 ***	-8.83 ***	-9.05 ***	0.11	I(1)
	constant and trend	0.11	-3.14	-3.19 *	0.13 *							

Notes:
*** indicates statistical significance at 1% level, ** indicates statistical significance at 5% level, * indicates statistical significance at 10% level
cc - currency in circulation; m1 - monetary aggregate M1; m2- monetary aggregate M2; hicp - harmonized index of consumer prices; hicp_c - harmonized index of consumer prices (core); ppi - producer price index; gdp_p - GDP deflator; "_sa" indicates seasonal adjustment
No presence of seasonality was detected for the producer price index.
ADF, PP and KPSS indicate Augmented Dickey-Fuller, Phillips-Perron and Kwiatkowski–Phillips–Schmidt–Shin unit root tests respectively;
ADF (AIC) and ADF (BIC) denote ADF unit root tests based on Akaike and Schwarz information criterion respectively

Appendix A2: Results from Clemente, Montañés and Reyes unit root test (1998) with double mean shifts (null hypothesis of a unit root)

Level	Additive outliers		Innovational outliers		First difference	Additive outliers		Innovational outliers		Order of integration of the variable
	t-stat.	No. structural breaks	t-stat.	No. structural breaks		t-stat.	No. structural breaks	t-stat.	No. structural breaks	
log(cc_sa)	-3.30	2	-5.24	0	**dlog(cc_sa)**	-9.53 **	2	-9.30 **	2	I(1)
log(m1_sa)	-3.21	2	-2.85	1	**dlog(m1_sa)**	-8.07 **	1	-7.90 **	1	I(1)
log(m2_sa)	-2.41	2	-4.09	2	**dlog(m2_sa)**	-5.86 **	1	-9.29 **	1	I(1)
log(hicp_sa)	-2.75	2	-4.24	2	**dlog(hicp_sa)**	-6.05 **	2	-6.40 **	2	I(1)
log(hicp_c_sa)	-3.55	2	-7.46 **	2	**dlog(hicp_c_sa)**	-6.61 **	2	-7.05 **	2	I(1)
log(ppi)	-3.64	2	-2.44	1	**dlog(ppi)**	-6.12 **	0	-7.36 **	2	I(1)
log(gdp_p_sa)	-2.80	2	-1.63	0	**dlog(gdp_p_sa)**	-6.47 **	0	-9.52 **	0	I(1)

Notes:
cc- currency in circulation; m1 - monetary aggregate M1; m2- monetary aggregate M2; hicp - harmonized index of consumer prices; hicp_c - harmonized index of consumer prices (core); ppi - producer price index; gdp_p - GDP deflator; "_sa" indicates seasonal adjustment
No presence of seasonality was detected for the producer price index.
Number of structural breaks indicates significant structural breaks at 5% significance level.
** indicates statistical significance at 5% level

Appendix B: Correlograms

Correlogram of dlog(hicp_sa)
Sample: 1999Q1 - 2012Q1
Included observations: 53

Lag	AC	PAC	Q-Stat	Prob
1	0.352 **	0.352 **	6.95	0.01
2	0.113	-0.013	7.67	0.02
3	0.087	0.059	8.12	0.04
4	-0.015	-0.072	8.13	0.09
5	-0.091	-0.076	8.63	0.13
6	-0.149	-0.108	10.01	0.12
7	-0.024	0.085	10.05	0.19
8	0.027	0.032	10.10	0.26
9	0.098	0.105	10.73	0.29
10	0.021	-0.079	10.76	0.38
11	-0.242	-0.306 **	14.81	0.19
12	-0.232	-0.116	18.64	0.10
13	0.002	0.205	18.64	0.14
14	-0.118	-0.096	19.68	0.14
15	-0.142	-0.054	21.23	0.13
16	-0.032	-0.048	21.31	0.17

Correlogram of dlog(hicp_c_sa)
Sample: 1999Q1 - 2012Q1
Included observations: 53

Lag	AC	PAC	Q-Stat	Prob
1	0.827 **	0.827 **	38.34	0.00
2	0.680 **	-0.011	64.80	0.00
3	0.567 **	0.024	83.56	0.00
4	0.355 **	-0.375 **	91.05	0.00
5	0.222	0.100	94.05	0.00
6	0.072	-0.227	94.37	0.00
7	-0.070	0.040	94.68	0.00
8	-0.133	-0.025	95.83	0.00
9	-0.200	0.020	98.49	0.00
10	-0.243	-0.068	102.49	0.00
11	-0.302 **	-0.212	108.81	0.00
12	-0.389 **	-0.196	119.55	0.00
13	-0.400 **	0.057	131.22	0.00
14	-0.395 **	0.038	142.87	0.00
15	-0.424 **	-0.144	156.68	0.00
16	-0.410 **	-0.062	169.91	0.00

Correlogram of dlog(cc_sa)
Sample: 1999Q1 - 2012Q1
Included observations: 53

Lag	AC	PAC	Q-Stat	Prob
1	0.304 **	0.304 **	5.17	0.02
2	0.378 **	0.315 **	13.35	0.00
3	0.318 **	0.175	19.23	0.00
4	0.262	0.069	23.31	0.00
5	0.157	-0.058	24.81	0.00
6	0.221	0.075	27.85	0.00
7	0.073	-0.076	28.19	0.00
8	0.060	-0.061	28.42	0.00
9	0.167	0.144	30.28	0.00
10	-0.074	-0.182	30.65	0.00
11	0.037	0.015	30.74	0.00
12	0.045	0.062	30.89	0.00
13	0.066	0.085	31.20	0.00
14	0.032	0.027	31.28	0.01
15	0.049	-0.057	31.47	0.01
16	-0.101	-0.146	32.26	0.01

Correlogram of dlog(m1_sa)
Sample: 1999Q1 - 2012Q1
Included observations: 53

Lag	AC	PAC	Q-Stat	Prob
1	0.277 **	0.277 **	4.31	0.04
2	0.369 **	0.316	12.07	0.00
3	0.242	0.102	15.50	0.00
4	0.100	-0.093	16.09	0.00
5	0.121	0.010	16.98	0.01
6	0.082	0.041	17.39	0.01
7	0.109	0.071	18.14	0.01
8	-0.122	-0.241	19.11	0.01
9	0.036	0.038	19.19	0.02
10	-0.118	-0.052	20.14	0.03
11	-0.030	0.053	20.20	0.04
12	-0.077	-0.071	20.62	0.06
13	0.072	0.172	21.00	0.07
14	-0.092	-0.131	21.63	0.09
15	-0.039	-0.010	21.74	0.12
16	-0.119	-0.179	22.85	0.12

Correlogram of dlog(m2_sa)
Sample: 1999Q1 - 2012Q1
Included observations: 53

Lag	AC	PAC	Q-Stat	Prob
1	0.157	0.157	1.39	0.24
2	0.314 **	0.296 **	7.02	0.03
3	0.304 **	0.250 **	12.39	0.01
4	0.091	-0.053	12.88	0.01
5	0.091	-0.085	13.37	0.02
6	0.001	-0.103	13.37	0.04
7	-0.036	-0.054	13.46	0.06
8	-0.088	-0.068	13.96	0.08
9	-0.028	0.050	14.02	0.12
10	-0.015	0.085	14.03	0.17
11	0.002	0.069	14.03	0.23
12	0.064	0.064	14.32	0.28
13	0.076	0.042	14.75	0.32
14	-0.028	-0.135	14.81	0.39
15	0.033	-0.081	14.89	0.46
16	-0.078	-0.130	15.37	0.50

**indicates statistical significance at 5% level.

Appendix C: Box-Jenkins univariate models

Fitted univariate (ARMA) models to the first-differenced series of money and prices

1. HICP

dlog(hicp_sa) = 0.013 + 0.353*MA(1)
(0.002)[1] (0.127)
$x^2(6) = 6.73$ - Ljung and Box (1978) chi-square statistic[2], with $x^2[0.05](6) = 12.59$
SSE (sum of squared residuals) = 0.007; **S.E. of regression** = 0.012

2. HICP_C

dlog(hicp_c_sa) = 0.008 + 0.319*AR(2)+0.551*AR(3)-0.232*AR(4)+1.08*MA(1)+0.518*MA(2)
(0.004) (0.113) (0.106) (0.082) (0.133) (0.141)
$x^2(2) = 4.70$ - Ljung and Box (1978) chi-square statistic, with $x^2[0.05](2) = 5.99$
SSE (sum of squared residuals) = 0.001; **S.E. of regression** = 0.004

3. CC

dlog(cc_sa) = 0.029 + 0.558*AR(3)+0.349*MA(1)+0.580*MA(2)-0.458*MA(3)-0.038*dummy (Q1'2009)
(0.010) (0.074) (0.129) (0.105) (0.123) (0.017)
$x^2(2) = 1.49$ - Ljung and Box (1978) chi-square statistic, with $x^2[0.05](2) = 5.99$
SSE (sum of squared residuals) = 0.023; **S.E. of regression** = 0.022

4. M1

dlog(m1_sa) = 0.033 + 0.428*AR(2)+0.109*dummy(Q1'2005)
(0.008) (0.126) (0.032)
$x^2(5) = 4.14$ - Ljung and Box (1978) chi-square statistic, with $x^2[0.05](5) = 11.07$
SSE (sum of squared residuals) = 0.063; **S.E. of regression** = 0.035

5. M2

dlog(m2_sa) = 0.040 +
(0.007)
0.260*AR(1)+0.305*MA(1)+0.750*MA(2)+1.059*MA(3)+0.386*MA(5)+dummies
(0.143) (0.080) (0.097) (0.053) (0.128)
$x^2(1) = 3.02$ - Ljung and Box (1978) chi-square statistic, with $x^2[0.05](1) = 3.84$
SSE (sum of squared residuals) = 0.006; **S.E. of regression** = 0.011

[1] Standard errors in parentheses

[2] The Ljung and Box (1978) chi-square statistic can be used to test whether a group of autocorrelations is significantly different from zero. For more details, see Enders, W., "Applied Econometric Time Series", 2nd edition, pp.67-69 and Ljung, G. and G.Box, "On a measure of lack of fit in time series models", *Biometrika* (1978), 65, 2, pp.297-303
The application of the chi-square test is done for 8 lags of the residuals of all fitted univariate models but the M2 model for which the chi-square statistic is calculated for 9 lags.

Appendix D: Application of Haugh's *s* statistic[3]

Test of the Lack of Relationship between monetary growth (CC) and HICP inflation rate based on Haugh's (1976) *s* statistic for small samples (quarterly data)

$$s^* = n^2 \sum_{k=-M}^{M} (n - |k|)^{-1} r_{12}(k)^2$$

Type of Relationship	Lags (-M)	Leads (M)	Degrees of freedom	*s**	Significance
1. Feedback					
	-8	8	17	34.2	***
	-12	12	25	40.4	**
2. Monetary growth causes inflation					
		1 to 8	8	12.1	
		1 to 12	12	14.2	
3. Inflation causes monetary growth					
	-1 to -8		8	13.4	
	-1 to -12		12	16.6	

The test is performed for the cross-correlations between the residuals of the ARMA models fitted for CC growth rate and HICP inflation rate: resid_CC $_t$ and resid_HICP $_{t-k}$, k = ± 0,1,2,3....M; -M is the number of lags and M is the number of leads (M is chosen arbitrarily); r_{12} is the cross-correlation estimate at the specific lag/lead between the two residual series of length n. The asymptotic distribution of s* is chi-square with 2M+1 degrees of freedom when k is from –M to +M and with M degrees of freedom when k is from 1 to M. The null hypothesis is that the series CC growth and HICP inflation are unrelated. *** indicates statistical significance at 1% level, ** indicates statistical significance at 5% level, * indicates statistical significance at 10% level

Test of the Lack of Relationship between monetary growth (CC) and HICP core inflation rate based on Haugh's (1976) *s* statistic for small samples (quarterly data)

$$s^* = n^2 \sum_{k=-M}^{M} (n - |k|)^{-1} r_{12}(k)^2$$

Type of Relationship	Lags (-M)	Leads (M)	Degrees of freedom	*s**	Significance
1. Feedback					
	-8	8	17	22.4	
	-12	12	25	35.9	*
2. Monetary growth causes inflation					
		1 to 8	8	15.2	*
		1 to 12	12	18.8	*
3. Inflation causes monetary growth					
	-1 to -8		8	7.2	
	-1 to -12		12	17.1	

The test is performed for the cross-correlations between the residuals of the ARMA models fitted for CC growth rate and HICP core inflation rate: resid_CC $_t$ and resid_HICP_C $_{t-k}$. See also notes above for testing the relationship between CC growth rate and HICP inflation rate.

[3] For similar analysis see Schwert, G.William, "Tests of Causality: The message in the innovations", Carnegie-Rochester Conference Series on Public Policy, 1979, Volume 10, pp.55-96

Test of the Lack of Relationship between monetary growth (M1) and HICP inflation rate
based on Haugh's (1976) *s* statistic for small samples (quarterly data)

$$s^* = n^2 \sum_{k=-M}^{M} (n - |k|)^{-1} r_{12}(k)^2$$

Type of Relationship	Lags (-M)	Leads (M)	Degrees of freedom	*s**	Significance
1. Feedback					
	-8	8	17	26.1	*
	-12	12	25	29.6	
2. Monetary growth causes inflation					
		1 to 8	8	14.1	*
		1 to 12	12	16.0	
3. Inflation causes monetary growth					
	-1 to -8		8	9.0	
	-1 to -12		12	10.3	

The test is performed for the cross-correlations between the residuals of the ARMA models fitted for M1 growth rate and HICP inflation rate: resid_M1 $_t$ and resid_HICP $_{t-k}$. See also notes above for testing the relationship between CC growth rate and HICP inflation rate.

Test of the Lack of Relationship between monetary growth (M1) and HICP core inflation rate
based on Haugh's (1976) *s* statistic for small samples (quarterly data)

$$s^* = n^2 \sum_{k=-M}^{M} (n - |k|)^{-1} r_{12}(k)^2$$

Type of Relationship	Lags (-M)	Leads (M)	Degrees of freedom	*s**	Significance
1. Feedback					
	-8	8	17	33.6	***
	-12	12	25	39.8	**
2. Monetary growth causes inflation					
		1 to 8	8	22.7	***
		1 to 12	12	25.7	**
3. Inflation causes monetary growth					
	-1 to -8		8	10.8	
	-1 to -12		12	14.1	

The test is performed for the cross-correlations between the residuals of the ARMA models fitted for M1 growth rate and HICP core inflation rate: resid_M1 $_t$ and resid_HICP_C $_{t-k}$. See also notes above for testing the relationship between CC growth rate and HICP inflation rate.

Test of the Lack of Relationship between monetary growth (M2) and HICP inflation rate based on Haugh's (1976) *s* statistic for small samples (quarterly data)

$$s^* = n^2 \sum_{k=-M}^{M} (n - |k|)^{-1} r_{12}(k)^2$$

Type of Relationship	Lags (-M)	Leads (M)	Degrees of freedom	*s**	Significance
1. Feedback					
	-8	8	17	34.0	***
	-12	12	25	39.3	**
2. Monetary growth causes inflation					
		1 to 8	8	20.8	***
		1 to 12	12	23.2	**
3. Inflation causes monetary growth					
	-1 to -8		8	10.0	
	-1 to -12		12	12.7	

The test is performed for the cross-correlations between the residuals of the ARMA models fitted for M2 growth rate and HICP inflation rate: resid_M2 $_t$ and resid_HICP $_{t-k}$. See also notes above for testing the relationship between CC growth rate and HICP inflation rate.

Test of the Lack of Relationship between monetary growth (M2) and HICP core inflation rate based on Haugh's (1976) *s* statistic for small samples (quarterly data)

$$s^* = n^2 \sum_{k=-M}^{M} (n - |k|)^{-1} r_{12}(k)^2$$

Type of Relationship	Lags (-M)	Leads (M)	Degrees of freedom	*s**	Significance
1. Feedback					
	-8	8	17	27.5	**
	-12	12	25	36.5	*
2. Monetary growth causes inflation					
		1 to 8	8	17.9	**
		1 to 12	12	24.1	**
3. Inflation causes monetary growth					
	-1 to -8		8	4.9	
	-1 to -12		12	7.3	

The test is performed for the cross-correlations between the residuals of the ARMA models fitted for M2 growth rate and HICP core inflation rate: resid_M2 $_t$ and resid_HICP_C $_{t-k}$. See also notes above for testing the relationship between CC growth rate and HICP inflation rate.

Appendix E: Granger Causality Tests

GRANGER CAUSALITY TEST RESULTS - HICP INFLATION				
INDICATORS: Δln HICP & Δln CURRENCY IN CIRCULATION				
PERIOD		DIRECTION	LAG LENGTH	SIGNI-FICANCE
1998Q1	2012Q1	INCONCLUSIVE	1	
1998Q2	2012Q1	TWO-WAY	1	**
1998Q3	2012Q1	HICP → CC	2	*
1998Q4	2012Q1	HICP → CC	2	*
1999Q1	2012Q1	INCONCLUSIVE	1	
1999Q2	2012Q1	HICP → CC	3	*
1999Q3	2012Q1	HICP → CC	3	*
1999Q4	2012Q1	HICP → CC	3	*
2000Q1	2012Q1	CC → HICP	1	**
2000Q2	2012Q1	CC → HICP	2	**
2000Q3	2012Q1	CC → HICP	2	*
2000Q4	2012Q1	CC → HICP	2	*
2001Q1	2012Q1	CC → HICP	1	**
2001Q2	2012Q1	CC → HICP	2	**
2001Q3	2012Q1	CC → HICP	2	*
2001Q4	2012Q1	CC → HICP	2	**
2002Q1	2012Q1	CC → HICP	1	**
2002Q2	2012Q1	INCONCLUSIVE	2	
2002Q3	2012Q1	INCONCLUSIVE	2	
2002Q4	2012Q1	INCONCLUSIVE	2	
2003Q1	2012Q1	INCONCLUSIVE	1	
2003Q2	2012Q1	HICP → CC	2	*
2003Q3	2012Q1	INCONCLUSIVE	2	
2003Q4	2012Q1	CC → HICP	2	*
2004Q1	2012Q1	HICP → CC	2	**
2004Q2	2012Q1	CC → HICP	2	*
2004Q3	2012Q1	HICP → CC	2	*
2004Q4	2012Q1	HICP → CC	2	**

GRANGER CAUSALITY TEST RESULTS - HICP INFLATION				
INDICATORS: Δln HICP & Δln M1				
PERIOD		DIRECTION	LAG LENGTH	SIGNI-FICANCE
1998Q1	2012Q1	INCONCLUSIVE	1	
1998Q2	2012Q1	INCONCLUSIVE	2	
1998Q3	2012Q1	M1 → HICP	2	*
1998Q4	2012Q1	M1 → HICP	2	*
1999Q1	2012Q1	INCONCLUSIVE	1	
1999Q2	2012Q1	INCONCLUSIVE	2	
1999Q3	2012Q1	INCONCLUSIVE	2	
1999Q4	2012Q1	INCONCLUSIVE	2	
2000Q1	2012Q1	HICP → M1	2	*
2000Q2	2012Q1	HICP → M1	2	*
2000Q3	2012Q1	INCONCLUSIVE	2	
2000Q4	2012Q1	M1 → HICP	2	*
2001Q1	2012Q1	M1 → HICP	1	***
2001Q2	2012Q1	M1 → HICP	2	*
2001Q3	2012Q1	M1 → HICP	2	*
2001Q4	2012Q1	M1 → HICP	2	*
2002Q1	2012Q1	M1 → HICP	2	**
2002Q2	2012Q1	M1 → HICP	2	*
2002Q3	2012Q1	M1 → HICP	2	*
2002Q4	2012Q1	M1 → HICP	2	*
2003Q1	2012Q1	INCONCLUSIVE	1	
2003Q2	2012Q1	M1 → HICP	2	**
2003Q3	2012Q1	M1 → HICP	2	**
2003Q4	2012Q1	M1 → HICP	2	*
2004Q1	2012Q1	INCONCLUSIVE	2	
2004Q2	2012Q1	INCONCLUSIVE	2	
2004Q3	2012Q1	M1 → HICP	5	**
2004Q4	2012Q1	M1 → HICP	6	***

GRANGER CAUSALITY TEST RESULTS - HICP INFLATION				
INDICATORS: Δln HICP & Δln M2				
PERIOD		DIRECTION	LAG LENGTH	SIGNI-FICANCE
1998Q1	2012Q1	TWO-WAY	3	**
1998Q2	2012Q1	TWO-WAY	3	**
1998Q3	2012Q1	TWO-WAY	3	**
1998Q4	2012Q1	TWO-WAY	3	**
1999Q1	2012Q1	TWO-WAY	3	**
1999Q2	2012Q1	M2 → HICP	2	**
1999Q3	2012Q1	TWO-WAY	4	**
1999Q4	2012Q1	TWO-WAY	3	**
2000Q1	2012Q1	TWO-WAY	3	**
2000Q2	2012Q1	M2 → HICP	2	**
2000Q3	2012Q1	TWO-WAY	3	**
2000Q4	2012Q1	TWO-WAY	3	**
2001Q1	2012Q1	TWO-WAY	3	**
2001Q2	2012Q1	M2 → HICP	5	**
2001Q3	2012Q1	M2 → HICP	5	**
2001Q4	2012Q1	M2 → HICP	2	**
2002Q1	2012Q1	TWO-WAY	3	**
2002Q2	2012Q1	M2 → HICP	2	**
2002Q3	2012Q1	HICP → M2	4	**
2002Q4	2012Q1	HICP → M2	4	***
2003Q1	2012Q1	M2 → HICP	2	**
2003Q2	2012Q1	M2 → HICP	2	**
2003Q3	2012Q1	M2 → HICP	4	**
2003Q4	2012Q1	M2 → HICP	4	**
2004Q1	2012Q1	INCONCLUSIVE	2	
2004Q2	2012Q1	HICP → M2	3	**
2004Q3	2012Q1	HICP → M2	4	**
2004Q4	2012Q1	M2 → HICP	2	**

GRANGER CAUSALITY TEST RESULTS - CORE INFLATION				
INDICATORS: Δln CORE HICP & Δln CURRENCY IN CIRCULATION				
PERIOD		DIRECTION	LAG LENGTH	SIGNI-FICANCE
1998Q1	2012Q1	INCONCLUSIVE	2	
1998Q2	2012Q1	INCONCLUSIVE	5	
1998Q3	2012Q1	INCONCLUSIVE	2	
1998Q4	2012Q1	INCONCLUSIVE	2	
1999Q1	2012Q1	CORE HICP → CC	2	***
1999Q2	2012Q1	CC → CORE HICP	2	**
1999Q3	2012Q1	CORE HICP → CC	3	*
1999Q4	2012Q1	CC → CORE HICP	1	**
2000Q1	2012Q1	CC → CORE HICP	2	**
2000Q2	2012Q1	CC → CORE HICP	2	**
2000Q3	2012Q1	CC → CORE HICP	2	**
2000Q4	2012Q1	CC → CORE HICP	2	**
2001Q1	2012Q1	CC → CORE HICP	1	**
2001Q2	2012Q1	INCONCLUSIVE	5	
2001Q3	2012Q1	CC → CORE HICP	2	*
2001Q4	2012Q1	CC → CORE HICP	3	**
2002Q1	2012Q1	CC → CORE HICP	1	***
2002Q2	2012Q1	CC → CORE HICP	2	**
2002Q3	2012Q1	CC → CORE HICP	2	**
2002Q4	2012Q1	INCONCLUSIVE	3	
2003Q1	2012Q1	CORE HICP → CC	6	**
2003Q2	2012Q1	CORE HICP → CC	2	**
2003Q3	2012Q1	INCONCLUSIVE	5	
2003Q4	2012Q1	CC → CORE HICP	5	*
2004Q1	2012Q1	INCONCLUSIVE	5	
2004Q2	2012Q1	CORE HICP → CC	5	*
2004Q3	2012Q1	INCONCLUSIVE	5	
2004Q4	2012Q1	CC → CORE HICP	6	***

GRANGER CAUSALITY TEST RESULTS - CORE INFLATION				
INDICATORS: Δln CORE HICP & Δln M1				
PERIOD		DIRECTION	LAG LENGTH	SIGNI-FICANCE
1998Q1	2012Q1	M1 → CORE HICP	2	**
1998Q2	2012Q1	M1 → CORE HICP	2	**
1998Q3	2012Q1	M1 → CORE HICP	2	**
1998Q4	2012Q1	M1 → CORE HICP	2	**
1999Q1	2012Q1	M1 → CORE HICP	2	**
1999Q2	2012Q1	M1 → CORE HICP	2	***
1999Q3	2012Q1	M1 → CORE HICP	2	***
1999Q4	2012Q1	M1 → CORE HICP	2	***
2000Q1	2012Q1	M1 → CORE HICP	2	***
2000Q2	2012Q1	M1 → CORE HICP	3	***
2000Q3	2012Q1	M1 → CORE HICP	2	**
2000Q4	2012Q1	M1 → CORE HICP	2	**
2001Q1	2012Q1	M1 → CORE HICP	1	**
2001Q2	2012Q1	INCONCLUSIVE	3	
2001Q3	2012Q1	INCONCLUSIVE	3	
2001Q4	2012Q1	M1 → CORE HICP	3	***
2002Q1	2012Q1	M1 → CORE HICP	3	***
2002Q2	2012Q1	M1 → CORE HICP	2	***
2002Q3	2012Q1	M1 → CORE HICP	2	***
2002Q4	2012Q1	M1 → CORE HICP	2	***
2003Q1	2012Q1	M1 → CORE HICP	2	***
2003Q2	2012Q1	M1 → CORE HICP	2	***
2003Q3	2012Q1	M1 → CORE HICP	4	**
2003Q4	2012Q1	M1 → CORE HICP	3	***
2004Q1	2012Q1	INCONCLUSIVE	8	
2004Q2	2012Q1	INCONCLUSIVE	3	
2004Q3	2012Q1	M1 → CORE HICP	5	**
2004Q4	2012Q1	M1 → CORE HICP	6	***

GRANGER CAUSALITY TEST RESULTS - CORE INFLATION				
INDICATORS: Δln CORE HICP & Δln M2				
PERIOD		DIRECTION	LAG LENGTH	SIGNI-FICANCE
1998Q1	2012Q1	M2 → CORE HICP	5	**
1998Q2	2012Q1	M2 → CORE HICP	2	**
1998Q3	2012Q1	M2 → CORE HICP	5	**
1998Q4	2012Q1	M2 → CORE HICP	5	**
1999Q1	2012Q1	M2 → CORE HICP	5	**
1999Q2	2012Q1	M2 → CORE HICP	2	**
1999Q3	2012Q1	INCONCLUSIVE	4	
1999Q4	2012Q1	INCONCLUSIVE	4	
2000Q1	2012Q1	TWO-WAY	3	**
2000Q2	2012Q1	M2 → CORE HICP	4	*
2000Q3	2012Q1	TWO-WAY	3	**
2000Q4	2012Q1	M2 → CORE HICP	4	*
2001Q1	2012Q1	INCONCLUSIVE	4	
2001Q2	2012Q1	M2 → CORE HICP	2	**
2001Q3	2012Q1	M2 → CORE HICP	3	***
2001Q4	2012Q1	M2 → CORE HICP	3	***
2002Q1	2012Q1	M2 → CORE HICP	7	***
2002Q2	2012Q1	M2 → CORE HICP	5	**
2002Q3	2012Q1	INCONCLUSIVE	4	
2002Q4	2012Q1	TWO-WAY	5	**
2003Q1	2012Q1	TWO-WAY	5	**
2003Q2	2012Q1	TWO-WAY	5	**
2003Q3	2012Q1	TWO-WAY	6	***
2003Q4	2012Q1	TWO-WAY	6	***
2004Q1	2012Q1	INCONCLUSIVE	5	
2004Q2	2012Q1	TWO-WAY	6	***
2004Q3	2012Q1	INCONCLUSIVE	3	
2004Q4	2012Q1	TWO-WAY	3	**

Notes: Granger causality results are based on a rolling basis with a moving start and a fixed end-date.
Significance is reported based on p-values obtained from $\chi 2$ (Wald) statistics of the joint significance of each of the lagged endogenous variables in the equation. *** indicates statistical significance at 1% level, ** indicates statistical significance at 5% level, * indicates statistical significance at 10% level

Tests are determined to be inconclusive when Granger causality is found in none of the directions or when a two-way causality is found, but in the latter case coefficients are significant at more than 5% level of significance. Tests are also considered inconclusive where residual autocorrelation could not have been avoided in the specifications.

GRANGER CAUSALITY TEST RESULTS - HICP INFLATION				
INDICATORS: ΔIn HICP & ΔIn CURRENCY IN CIRCULATION				
PERIOD		DIRECTION	LAG LENGTH	SIGNI-FICANCE
1998Q1	2006Q4	HICP → CC	3	*
1999Q1	2007Q4	INCONCLUSIVE	3	
2000Q1	2008Q4	INCONCLUSIVE	2	
2001Q1	2009Q1	INCONCLUSIVE	2	
2002Q1	2010Q1	INCONCLUSIVE	2	
2003Q1	2011Q1	INCONCLUSIVE	2	
2004Q1	2012Q1	HICP → CC	2	**

GRANGER CAUSALITY TEST RESULTS - HICP INFLATION				
INDICATORS: ΔIn HICP & ΔIn M1				
PERIOD		DIRECTION	LAG LENGTH	SIGNI-FICANCE
1998Q1	2006Q1	HICP → M1	2	***
1999Q1	2007Q1	HICP → M1	2	***
2000Q1	2008Q1	HICP → M1	2	**
2001Q1	2009Q1	M1 → HICP	3	**
2002Q1	2010Q1	INCONCLUSIVE	3	
2003Q1	2011Q1	M1 → HICP	2	*
2004Q1	2012Q1	M1 → HICP	2	*

GRANGER CAUSALITY TEST RESULTS - HICP INFLATION				
INDICATORS: ΔIn HICP & ΔIn M2				
PERIOD		DIRECTION	LAG LENGTH	SIGNI-FICANCE
1998Q1	2006Q1	HICP → M2	5	**
1999Q1	2007Q1	INCONCLUSIVE	5	
2000Q1	2008Q1	INCONCLUSIVE	7	
2001Q1	2009Q1	M2 → HICP	6	*
2002Q1	2010Q1	INCONCLUSIVE	5	
2003Q1	2011Q1	M2 → HICP	2	*
2004Q1	2012Q1	INCONCLUSIVE	2	

GRANGER CAUSALITY TEST RESULTS - CORE INFLATION				
INDICATORS: ΔIn CORE HICP & ΔIn CURRENCY IN CIRCULATION				
PERIOD		DIRECTION	LAG LENGTH	SIGNI-FICANCE
1998Q1	2006Q1	INCONCLUSIVE	2	
1999Q1	2007Q1	INCONCLUSIVE	2	
2000Q1	2008Q1	CC → CORE	2	*
2001Q1	2009Q1	INCONCLUSIVE	2	
2002Q1	2010Q1	INCONCLUSIVE	2	
2003Q1	2011Q1	CC → CORE	3	**
2004Q1	2012Q1	CC → CORE	5	*

GRANGER CAUSALITY TEST RESULTS - CORE INFLATION				
INDICATORS: ΔIn CORE HICP & ΔIn M1				
PERIOD		DIRECTION	LAG LENGTH	SIGNI-FICANCE
1998Q1	2006Q1	INCONCLUSIVE	2	
1999Q1	2007Q1	INCONCLUSIVE	2	
2000Q1	2008Q1	M1 → CORE	2	*
2001Q1	2009Q1	M1 → CORE	2	***
2002Q1	2010Q1	M1 → CORE	2	***
2003Q1	2011Q1	M1 → CORE	2	***
2004Q1	2012Q1	INCONCLUSIVE	7	

GRANGER CAUSALITY TEST RESULTS - CORE INFLATION				
INDICATORS: ΔIn CORE HICP & ΔIn M2				
PERIOD		DIRECTION	LAG LENGTH	SIGNI-FICANCE
1998Q1	2006Q1	INCONCLUSIVE	2	
1999Q1	2007Q1	INCONCLUSIVE	3	
2000Q1	2008Q1	CORE → M2	6	**
2001Q1	2009Q1	INCONCLUSIVE	2	
2002Q1	2010Q1	INCONCLUSIVE	2	
2003Q1	2011Q1	TWO-WAY	2	**
2004Q1	2012Q1	INCONCLUSIVE	2	

Notes: Granger causality results are based on a moving window with fixed length of eight years. See also the notes for the Granger test results based on a rolling basis with a moving start and a fixed end-date presented above.

Appendix F: VECM RESULTS

F1: VECM for cc, hicp, real retail sales and interest rate on new term deposits

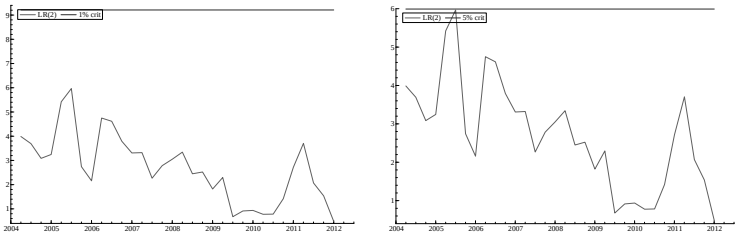
1. Unrestricted Cointegration Rank Test (Trace)

Sample: 2000Q3 - 2012Q1

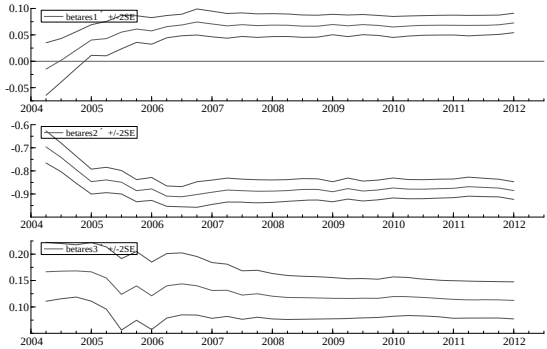
Hypothesized No. of CE(s)	Eigenvalue	Trace Statistic	Critical Value 0.05	Reinsel and Ahn (1988) correction of the critical value
None	0.67	105.16 **	47.86	57.67
At most 1	0.52	52.46 **	29.80	35.91
At most 2	0.25	17.78	15.49	18.67
At most 3	0.09	4.41	3.84	4.63

** indicates statistical significance at 5% level, considering sample size adjusted critical values

2. Recursively estimated likelihood ratio (LR) test statistic for over-identifying restrictions, scaled by the 1% and 5% critical values respectively



3. Recursively estimated coefficients in the two identified cointegrating vectors



4. Short-run dynamics estimated by OLS

   Cointegrating relations:

   CIa=1*log(cc_sa) - 1*log(hicp_sa) - 0.885*log(tradec_sa) + 0.073*m_tdir_n

   CIb=1*log(hicp_sa) - 1*log(tradec_sa) + 0.112*m_tdir_n

equation for dlog(cc_sa)

	Coefficient	Std.Error	t-value	Stat. significance
Constant	3.796	0.741	5.120	***
CIa(-1)	-0.536	0.103	-5.180	***
CIb(-1)	-0.020	0.040	-0.500	
dlog(cc_sa(-1))	0.203	0.136	1.500	
dlog(hicp_sa(-1))	-0.518	0.316	-1.640	
d(m_tdir_n(-1))	-0.013	0.011	-1.200	
dlog(tradec_sa(-1))	-0.050	0.337	-0.147	
d_04_2008	0.006	0.020	0.306	
d_04_2000	-0.034	0.022	-1.580	

sigma = 0.018 RSS = 0.012

equation for dlog(hicp_sa)

	Coefficient	Std.Error	t-value	Stat. significance
Constant	1.372	0.329	4.170	***

CIa(-1)	-0.187	0.046	-4.070	***
CIb(-1)	-0.092	0.018	-5.130	***
dlog(cc_sa(-1))	0.206	0.060	3.420	***
dlog(hicp_sa(-1))	-0.005	0.141	-0.036	
d(m_tdir_n(-1))	-0.002	0.005	-0.448	
dlog(tradec_sa(-1))	-0.728	0.150	-4.860	***
d_04_2008	-0.026	0.009	-2.990	***
d_04_2000	-0.001	0.010	-0.084	

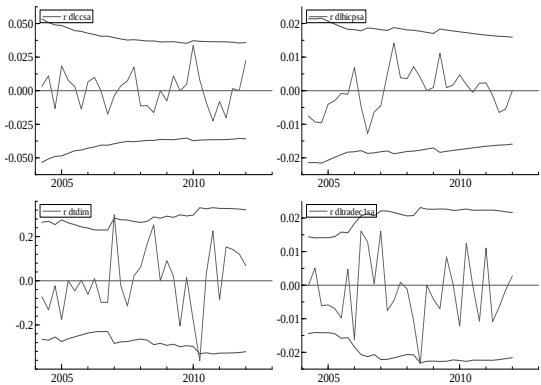
sigma = 0.008 RSS = 0.002

*Note: The short-run equations are presented only for the money and price variables. *** indicates statistical significance at 1% level, ** indicates statistical significance at 5% level, * indicates statistical significance at 10% level

5. Specification tests for the short-run error-correction model

```
Single-equation diagnostics using reduced-form residuals
dlog(cc_sa)      : AR 1-4 test:      F(4,34)   =   0.894 [0.4778]
dlog(cc_sa)      : ARCH 1-4 test:    F(4,39)   =   1.481 [0.2267]
dlog(cc_sa)      : Normality test:   Chi^2(2)  =   2.746 [0.2534]
dlog(cc_sa)      : Hetero test:      F(12,32)  =   0.628 [0.8028]
dlog(hicp_sa)    : AR 1-4 test:      F(4,34)   =   1.075 [0.3843]
dlog(hicp_sa)    : ARCH 1-4 test:    F(4,39)   =   0.270 [0.8957]
dlog(hicp_sa)    : Normality test:   Chi^2(2)  =   2.682 [0.2616]
dlog(hicp_sa)    : Hetero test:      F(12,32)  =   0.833 [0.6169]
d(m_tdir_n)      : AR 1-4 test:      F(4,34)   =   1.569 [0.2048]
d(m_tdir_n)      : ARCH 1-4 test:    F(4,39)   =   0.104 [0.9806]
d(m_tdir_n)      : Normality test:   Chi^2(2)  =   1.493 [0.4741]
d(m_tdir_n)      : Hetero test:      F(12,32)  =   1.309 [0.2616]
dlog(tradec_sa)  : AR 1-4 test:      F(4,34)   =   1.866 [0.1391]
dlog(tradec_sa)  : ARCH 1-4 test:    F(4,39)   =   0.070 [0.9906]
dlog(tradec_sa)  : Normality test:   Chi^2(2)  =   1.075 [0.5841]
dlog(tradec_sa)  : Hetero test:      F(12,32)  =   1.152 [0.3543]
Vector tests
Vector AR 1-4 test:      F(64,76)  =   1.104 [0.3384]
Vector Normality test:   Chi^2(8)  =   7.358 [0.4986]
Vector ZHetero test:     F(48,113) =   1.080 [0.3633]
Vector RESET23 test:     F(32,101) =   1.523 [0.0593]
```

6. Constancy of 1-step ahead recursively estimated residuals in the short-run model



F2: VECM for m1, ppi, industrial production index and interest rate on new term deposits

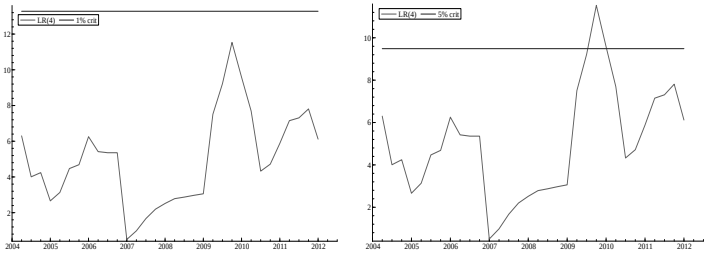
1. Unrestricted Cointegration Rank Test (Trace)

Sample: 2000Q2 - 2012Q1

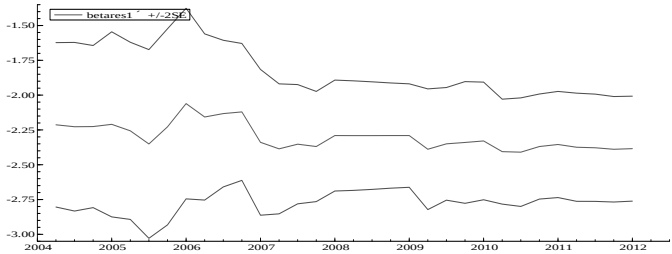
Hypothesized No. of CE(s)	Eigenvalue	Trace Statistic	Critical Value 0.05	Reinsel and Ahn (1988) correction of the critical value
None	0.67	70.52 **	47.86	52.21
At most 1	0.30	17.63	29.80	32.51
At most 2	0.01	0.70	15.49	16.90
At most 3	0.00	0.00	3.84	4.19

** indicates statistical significance at 5% level, considering sample size adjusted critical values

2. Recursively estimated likelihood ratio (LR) test statistic for over-identifying restrictions, scaled by the 1% and 5% critical values respectively



3. Recursively estimated unrestricted coefficient of the industrial production index variable in the identified cointegrating vector



4. Short-run dynamics estimated by OLS

Cointegrating relation:

CIa = 1*log(m1_sa) - 1*log(ppi) - 2.385*log(indp_sa)

equation for dlog(m1_sa)

	Coefficient	Std.Error	t-value	Stat. significance
Constant	-0.339	0.141	-2.410	**
CIa(-1)	-0.162	0.040	-4.010	***
log(er(-2))	0.108	0.032	3.350	***
dlog(oil_p)	0.041	0.037	1.130	
d_03_2008	-0.064	0.035	-1.830	*
D_01_2009	-0.086	0.036	-2.410	**

sigma = 0.033 RSS = 0.045

equation for dlog(ppi)

	Coefficient	Std.Error	t-value	Stat. significance
Constant	0.058	0.048	1.210	
CIa(-1)	-0.024	0.014	-1.760	*
log(er(-2))	-0.006	0.011	-0.534	
dlog(oil_p)	0.085	0.012	6.840	***
d_03_2008	0.034	0.012	2.890	***
D_01_2009	-0.049	0.012	-4.040	***

sigma = 0.011 RSS = 0.005

*Note: The short-run equations are presented only for the money and price variables. *** indicates statistical significance at 1% level, ** indicates statistical significance at 5% level, * indicates statistical significance at 10% level

5. Specification tests for the short-run error-correction model

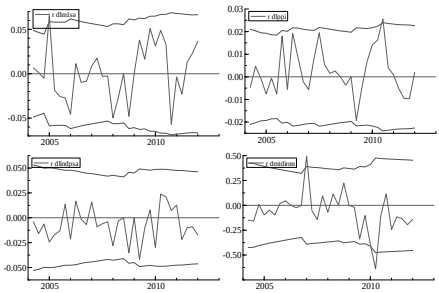
Single-equation diagnostics using reduced-form residuals

```
dlog(m1_sa)      : AR 1-4 test:      F(4,37)   =   1.738 [0.1623]
dlog(m1_sa)      : ARCH 1-4 test:    F(4,39)   =   0.405 [0.8035]
dlog(m1_sa)      : Normality test:   Chi^2(2)  =   2.378 [0.3046]
dlog(m1_sa)      : Hetero test:      F(6,38)   =   1.298 [0.2817]
dlog(ppi)        : AR 1-4 test:      F(4,37)   =   0.631 [0.6436]
dlog(ppi)        : ARCH 1-4 test:    F(4,39)   =   0.656 [0.6265]
dlog(ppi)        : Normality test:   Chi^2(2)  =   1.758 [0.4152]
dlog(ppi)        : Hetero test:      F(6,38)   =   1.402 [0.2393]
dlog(indp_sa)    : AR 1-4 test:      F(4,37)   =   3.494 [0.0162]*
dlog(indp_sa)    : ARCH 1-4 test:    F(4,39)   =   2.245 [0.0817]
dlog(indp_sa)    : Normality test:   Chi^2(2)  =   0.476 [0.7882]
dlog(indp_sa)    : Hetero test:      F(6,38)   =   1.414 [0.2347]
d(m_tdir_n)      : AR 1-4 test:      F(4,37)   =   0.240 [0.9137]
d(m_tdir_n)      : ARCH 1-4 test:    F(4,39)   =   0.919 [0.4626]
d(m_tdir_n)      : Normality test:   Chi^2(2)  =   3.196 [0.2023]
d(m_tdir_n)      : Hetero test:      F(6,38)   =   2.139 [0.0711]
```

Vector tests

```
Vector AR 1-4 test:      F(64,88)  =   1.239 [0.1749]
Vector Normality test:   Chi^2(8)  =   8.437 [0.3920]
Vector ZHetero test:     F(24,123) =   1.000 [0.4714]
Vector RESET23 test:     F(32,112) =   2.655 [0.0001]**
```

6. Constancy of 1-step ahead recursively estimated residuals in the short-run model



F3: VECM for m2, ppi, industrial production, interest rate on new term deposits, interest rates on loans

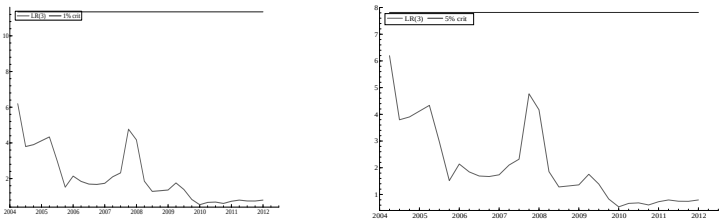
1. Unrestricted Cointegration Rank Test (Trace)

Sample: 2000Q3 2012Q1

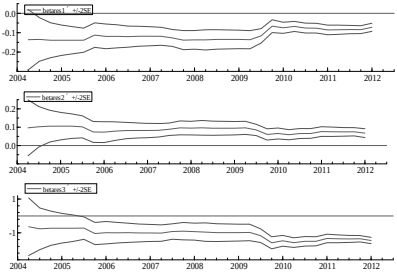
Hypothesized No. of CE(s)	Eigenvalue	Trace Statistic	Critical Value 0.05	Reinsel and Ahn (1988) correction of the critical value
None	0.67	96.23 **	69.82	88.69
At most 1	0.43	44.53	47.86	60.79
At most 2	0.21	17.96	29.80	37.85
At most 3	0.11	7.18	15.49	19.68
At most 4	0.03	1.64	3.84	4.88

** indicates statistical significance at 5% level, considering sample size adjusted critical values

2. Recursively estimated likelihood ratio (LR) test statistic for over-identifying restrictions



3. Recursively estimated unrestricted coefficients in the identified cointegrating vector



4. Short-run dynamics estimated by OLS

Cointegrating relation:

CIa=1*log(m2_sa) - 1*log(indp_sa) - 1.454*log(ppi) - 0.072*m_tdir_n + 0.068*ir_loans

equation for dlog(m2_sa)

	Coefficient	Std.Error	t-value	Stat. significance
Constant	0.930	0.319	2.910	***
CIa(-1)	-0.156	0.045	-3.470	***
dlog(indp_sa(-1))	0.200	0.076	2.630	**
d(m_tdir_n(-1))	-0.010	0.013	-0.744	
d(ir_loans(-1))	0.013	0.004	3.470	***
dlog(m2_sa(-1))	0.186	0.102	1.820	*
dlog(ppi(-1))	-0.330	0.186	-1.780	*
log(er(-1))	0.014	0.015	0.872	
dlog(oil_p)	0.018	0.021	0.851	
dlog(rgdp_ea_sa)	0.109	0.326	0.336	
d_01_2005	0.117	0.013	8.970	***
d_02_2005	-0.094	0.019	-4.930	***
d_01_2007	0.007	0.014	0.513	
d_04_2008	-0.037	0.023	-1.640	
d_01_2009	-0.035	0.021	-1.710	*

sigma = 0.012 RSS = 0.005

equation for dlog(ppi)

	Coefficient	Std.Error	t-value	Stat. significance
Constant	-0.042	0.274	-0.155	
CIa(-1)	0.016	0.038	0.415	
dlog(indp_sa(-1))	-0.027	0.065	-0.416	
d(m_tdir_n(-1))	0.002	0.011	0.140	
d(ir_loans(-1))	0.002	0.003	0.755	
dlog(m2_sa(-1))	0.183	0.087	2.090	**
dlog(ppi(-1))	0.490	0.159	3.080	***
log(er(-1))	-0.012	0.013	-0.874	
dlog(oil_p)	0.078	0.018	4.360	***
dlog(rgdp_ea_sa)	-0.542	0.279	-1.940	*
d_01_2005	-0.017	0.011	-1.520	
d_02_2005	-0.016	0.016	-0.960	
d_01_2007	0.004	0.012	0.307	
d_04_2008	-0.035	0.020	-1.790	*
d_01_2009	-0.034	0.018	-1.940	*

sigma = 0.011 RSS = 0.004

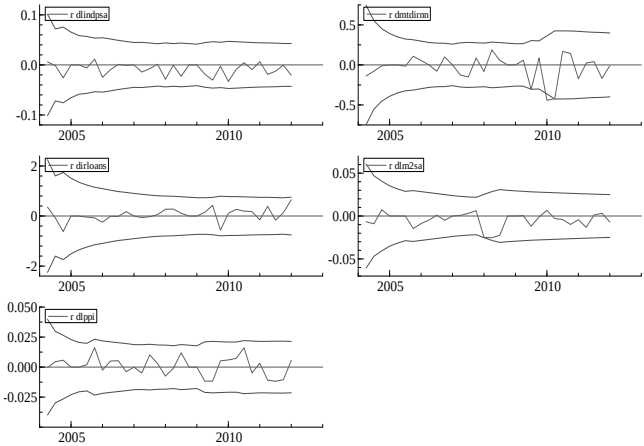
*Note: The short-run equations are presented only for the money and price variables. *** indicates statistical significance at 1% level, ** indicates statistical significance at 5% level, * indicates statistical significance at 10% level

5. Specification tests for the short-run error-correction model

```
Single-equation diagnostics using reduced-form residuals
dlog(indp_sa)    AR 1-4 test:      F(4,28)   =   3.598 [0.0173]*
dlog(indp_sa)    ARCH 1-4 test:    F(4,39)   =   0.707 [0.5922]
dlog(indp_sa)    Normality test:   Chi^2(2)  =   0.302 [0.8597]
dlog(indp_sa)    Hetero test:      F(18,23)  =   1.360 [0.2409]
d(m_tdir_n)      AR 1-4 test:      F(4,28)   =   0.511 [0.7283]
d(m_tdir_n)      ARCH 1-4 test:    F(4,39)   =   3.938 [0.0089]**
d(m_tdir_n)      Normality test:   Chi^2(2)  =   5.676 [0.0586]
d(m_tdir_n)      Hetero test:      F(18,23)  =   1.026 [0.4696]
d(ir_loans)      AR 1-4 test:      F(4,28)   =   2.094 [0.1083]
d(ir_loans)      ARCH 1-4 test:    F(4,39)   =   0.515 [0.7248]
d(ir_loans)      Normality test:   Chi^2(2)  =   3.267 [0.1953]
d(ir_loans)      Hetero test:      F(18,23)  =   0.905 [0.5801]
dlog(m2_sa)      AR 1-4 test:      F(4,28)   =   0.787 [0.5436]
dlog(m2_sa)      ARCH 1-4 test:    F(4,39)   =   0.845 [0.5042]
dlog(m2_sa)      Normality test:   Chi^2(2)  =   0.402 [0.8177]
dlog(m2_sa)      Hetero test:      F(18,23)  =   1.810 [0.0900]
dlog(ppi)        AR 1-4 test:      F(4,28)   =   0.792 [0.5402]
dlog(ppi)        ARCH 1-4 test:    F(4,39)   =   0.594 [0.6693]
dlog(ppi)        Normality test:   Chi^2(2)  =   3.477 [0.1758]
dlog(ppi)        Hetero test:      F(18,23)  =   1.408 [0.2176]

Vector tests
Vector AR 1-4 test:      F(100,43) =   1.020 [0.4835]
Vector Normality test:   Chi^2(10) =   19.46 [0.0348]*
Vector ZHetero test:     F(90,96)  =   1.469 [0.0322]*
Vector RESET23 test:     F(50,85)  =   1.176 [0.2523]
```

6. Constancy of 1-step ahead recursively estimated residuals in the short-run model



DP/102/2016

Analysis of the Bank Lending Survey Results for Bulgaria (for the 2003–2014 Period)

Tania Karamisheva

Abstract. This study examines descriptively and empirically the information content of the Bulgarian National Bank's quarterly bank lending survey for credit developments, focusing on lending to non-financial corporations. Carrying out an assessment on a macro and micro level we find that changes in demand estimated by survey data have statistically significant effect on corporate loan dynamics. The empirical research also reveals that other important factors are real GDP growth, the business climate, the share of bad and restructured loans, and bank-specific factors such as the capital to assets ratio and the interest spread between corporate loans and deposits.

Резюме. Статията изследва дескриптивно и емпирично информационното съдържание на Анкетата за кредитната активност на банките, провеждана от Българската народна банка на тримесечна база, за динамиката на кредита, като се фокусира върху кредитирането на нефинансовите предприятия. Извършвайки оценка на макро- и микрониво, установяваме, че промените в търсенето, оценени с данните от Анкетата, оказват статистически значимо влияние върху динамиката на корпоративните кредити. Емпиричното изследване също показва, че други съществени фактори са растежът на реалния БВП, бизнес климатът, делът на лошите и преструктурираните кредити в банковите портфейли, както и специфични за банките фактори, като съотношението „капитал/активи" и спредовете между лихвените проценти по кредитите и депозитите на предприятията.

Contents

1. Introduction .....

2. Literature overview .....

3. Banking system and credit developments in Bulgaria and the BNB policy after the introduction of the currency board ....

4. Survey results for Bulgaria .....
   4.1. Lending to enterprises .....
   4.2. Lending to households .....
   4.3. Comparison of bank lending survey data with other indicators .....

5. Empirical evidence .....
   5.1. Macro level .....
   5.2. Individual banks .....

6. Conclusions .....

References .....

Annex I .....

Structure and implementation of the Bank Lending Survey ....

Annex II .....

1. Introduction

Since the fourth quarter of 2003 the Bulgarian National Bank has conducted a regular quarterly bank lending survey among commercial banks in Bulgaria. The aim of the survey is to obtain additional qualitative information about changes in banks' lending policy and in the demand for loans, as well as to identify factors, affecting credit demand and banks' credit standards and terms. This additional information may be helpful to enhance understanding of the lending behavior of banks and the role of credit in the economy. Credit developments may have different implications for macroeconomic policy decisions, depending on whether their determinants are demand or supply side driven. The main contribution of the bank lending survey is in making distinction between loan demand and loan supply factors, as definitive conclusions about the exact determinants of changes in lending to enterprises and households cannot be drawn from the available monetary statistics. Thus, the findings of the survey can be useful for a complementary interpretation of existing monetary and interest rate statistics. They can also help to improve the forecasting of credit growth and economic developments.

In this paper we present the results from the bank lending survey and try to examine its information content for lending growth. We try to find a relation between the survey results and other macroeconomic variables such as real GDP growth, loan growth, gross fixed capital formation or industrial confidence. Furthermore we undertake an empirical analysis, first on a macro level using aggregate data on lending. In a next step, we construct a panel for which purpose we merge the individual banks' responses to the bank lending survey (BLS) questions with individual data on lending amounts for the surveyed banks.

The paper is organized as follows. Chapter 2 provides a summary of the main findings of different theoretical and empirical studies analyzing bank lending survey results and their role in explaining credit developments or changes in leading macroeconomic indicators. Chapter 3 provides a short overview of the main banking system and credit developments in Bulgaria before and during the global financial and economic crisis and the role of the BNB monetary policy. Chapter 4 follows with a discussion of the main bank lending survey results for Bulgaria and with a comparison of BLS results with other macroeconomic and financial data. Chapter 5 provides an empirical analysis both on a macro level and by individual banks. Chapter 6 concludes with some final remarks.

2. Literature overview

Credit developments are an important determinant of economic developments, and conditions in credit markets may affect the way monetary policy has an impact on the economy. In this respect, it is of importance to be able to distinguish between factors affecting the credit supply and those altering the demand for credit both of which influence the actual volume of credit. Available data from the monetary statistics on changes in bank lending provide information only on realized transaction volumes. However, they do not give an indication whether and to what extent these changes are influenced by the supply side or the demand side. The objective of the bank lending survey is to contribute to the closure of this gap and to enhance knowledge of developments in banks' lending policies. The qualitative results obtained from the survey should enable policy-makers to assess credit developments more accurately. The survey also provides the banks' assessment of the factors determining their potential changes in the supply of loans and those influencing changes in credit demand. Thus, the findings of the survey can be useful for a complementary interpretation of existing monetary and interest rate statistics. They can also help to improve the forecasting of credit growth and economic developments.

Several studies analyze the information content of bank lending surveys conducted in different countries, part of the Eurozone, the Eurozone as a whole, or the USA for the explanation of changes in credit activity or some real variables like GDP, consumption or investment. In part of the studies only a descriptive analysis is used, based on the graphical comparison of data collected via the bank lending survey and other macroeconomic data and focuses on finding some similar trends in their performance. Berg et al. (2005) for example, present the first results of the bank lending survey for the euro area, conducted since January 2003, and compare them with information derived from other sources. They compare BLS data on credit standards and real GDP growth or MFI loan growth and also carry out a comparison of BLS data and industrial confidence, consumer confidence or gross fixed capital formation. The graphical and descriptive analysis shows that even at this early stage of conducting the survey, it is possible to identify some systematic patterns in the results from the bank lending survey that prove to be in line with indicators retrieved from other sources. Mottiar and Monks (2007) undertake an analysis of the bank lending survey results for Ireland and compare them with aggregate euro area results. By means of graphical and descriptive analysis they also conclude that it is

possible to see some systematic patterns between the bank lending survey and other macro variables, in particular with regards to loan growth, gross fixed capital formation and consumer/industrial confidence.

Other part of the studies focus on an empirical analysis, using different econometric techniques and methods. Lown, Morgan and Rohatgi (2000) for example, using data obtained from the survey undertaken by the Federal Reserve, find that a strong correlation exists between tightening of credit standards and slowdowns in commercial lending and output. They find that the economy seems to grow more slowly during periods, in which banks tighten credit standards and that four of the five past recessions were preceded by sharply tighter standards. The chain of events following a standards' tightening resembles a credit crunch: commercial loans at banks plummet immediately and continue to fall until lenders ease up, output falls, and the federal funds rate, which is identified with the stance of monetary policy, is lowered. In a further study, using VAR analysis, Lown and Morgan (2002) find that fluctuations in credit standards are highly significant in predicting commercial bank loans, real GDP and inventory investment in the trade sector. They conclude that credit standards are more informative about future lending than loan rates, which is consistent with the idea that some sort of friction in lending markets leads lenders to ration loans via changes in standards more than through changes in rates. They also find a feedback from loans to standards, suggesting a sort of credit cycle. Higher loan levels cause tightening standards, perhaps because lenders conclude (or are told by supervisors) that standards are too loose. Tighter standards are followed by lower spending and loan levels, which eventually cause easing standards and higher spending and loan levels etc. Some of their negative findings are that shocks to the federal funds rate do not cause changes in standards, because lenders simply raise loan rates more or less in step with the funds rate. In the Monthly Report of Deutsche Bundesbank (January 2009) a simple regression analysis is undertaken in order to examine the explanatory content of BLS data on credit supply and demand for developments in lending to non-financial corporations in Germany. The regression analysis indicates importance of demand for developments in long-term lending, while the BLS supply variable lacks significance. In the case of long-term loans to enterprises the BLS demand is a robustly significant explanatory factor, which suggests that growth in long-term corporate lending in Germany has been determined in large part by demand-side factors. Bondt et al. (2010) examine empirically the information content of the euro area bank lending survey for aggregate credit and output growth. Using panel regression analysis they show that

the responses of the lending survey, especially those related to loans to enterprises, are a significant leading indicator for euro area bank credit and real GDP growth. Their results support the existence of a bank lending, balance sheet and risk-taking channel of monetary policy. These findings imply that not only changes in the official interest rate and in loan demand matter for credit and output, but also bank loan supply factors, the balance sheet position of borrowers and the risk perception in the economy. Finally, the authors discuss the implications for the 2008/2009 financial and economic crisis and come to the conclusion that the BLS responses provided an early and reliable signal about the deterioration of financing conditions and economic growth in the euro area. According to their panel estimates, the strong net tightening of credit standards and the increases in margins on average and riskier loans to enterprises during the crisis resulted in around one percentage point lower quarterly GDP growth in the euro area. Blaes (2011) undertakes an analysis of the role of bank-related factors in explaining the slowdown in bank lending to non-financial corporations in Germany during the recent financial and economic crisis. For the econometric panel analysis micro data on lending quantities and prices is used and matched to individual banks' survey responses. The main findings of the paper suggest that BLS indicators have a significant explanatory power for bank lending in the period 2003-2010. Both bank-related supply and demand-side factors prove to be important in explaining the sharp slowdown in lending after the collapse of Lehman Brothers. The results indicate that the dampening impact of the bank-related supply factor on loan developments occurred with a time lag of several quarters and was strongest from the third quarter of 2009 to the first quarter of 2010. During this period more than one third of the explained negative loan development was due to the restrictive adjustments of purely bank-side determinants, such as banks' capital costs, market financing conditions and liquidity position.

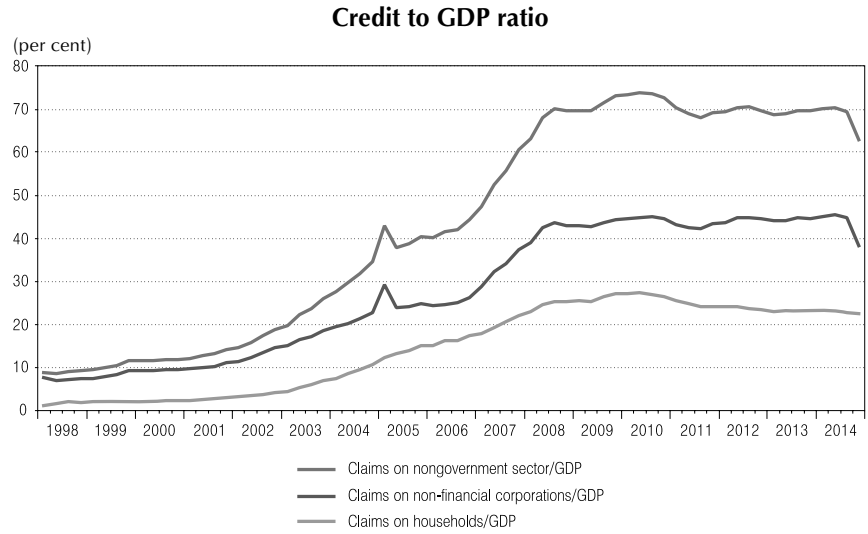
3. Banking system and credit developments in Bulgaria and the BNB policy after the introduction of the currency board

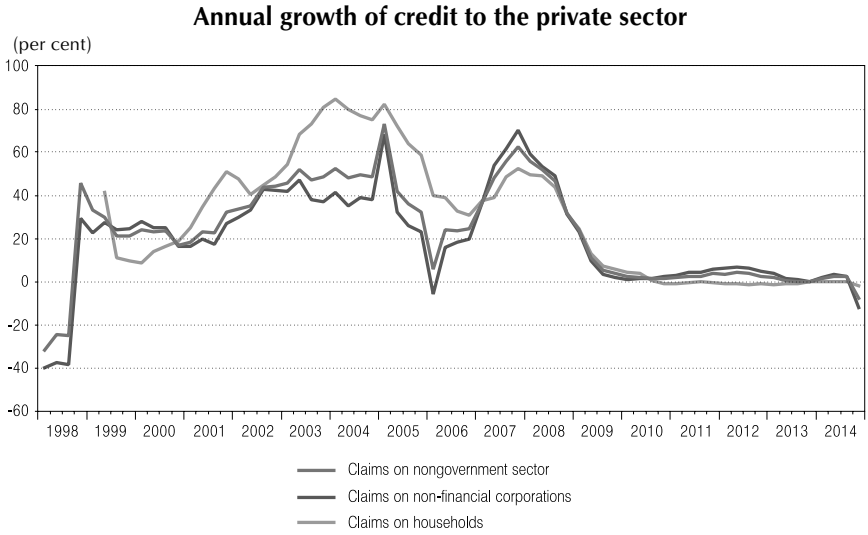
In this section we provide a short overview of the main banking system and credit developments in Bulgaria before and during the global financial and economic crisis along with a description of the Bulgarian National Bank policy over the period after the introduction of the currency board arrangement. The purpose for this is to set the context, in which we will later present the main results from the bank lending survey.

After several inconclusive attempts to stabilize the Bulgarian economy in 1991–1996 and a major financial crisis, which culminated in a short-lived hyperinflationary episode in December 1996–February 1997, a currency board in Bulgaria was introduced with the new Law on the Bulgarian National Bank of 10 June 1997. In the first several years after the adoption of the currency board, credit growth in Bulgaria was moderate and the credit to GDP ratio was low averaging 11% in the period 1998–2001. At that time the banking system in Bulgaria was characterized by comparatively high level of non-performing loans, low capitalization and liquidity constraints. There were also structural factors, which inhibited the expansion in bank lending associated with the fact, that the majority of banks were stated-owned and lacked the knowledge required for modern banking practices. Meanwhile, bank privatization was an important factor which started the gradual process of restructuring of the banking sector in Bulgaria.

From 2002 onwards a gradual credit expansion was observed and credit to GDP reached nearly 70% in late 2008. In the years before the collapse of Lehman Brothers there were two periods of high growth of credit to the private sector in Bulgaria: the first one from 2003 till 2005 and the second one in 2007. Rapid credit growth in these years was driven on the one hand by high loan demand, which was stimulated by the favorable domestic and external macroeconomic environment and the global upswing in the credit cycle, high expected return on investment and positive income convergence expectations. On the other hand, banks actively expanded their operations. An important factor which contributed to the deepening of the financial intermediation over the period was the privatization of many domestic banks by foreign financial institutions. Parent banks provided capital, liquidity and know-how to their subsidiary banks and

Chart 1: Credit developments in Bulgaria after the introduction of the currency board in 1997





Source: BNB, NSI.

their branches in Bulgaria, intending to boost their market share in the region where return on capital was very high. These processes prompted a strong competition among banks and certain easing of lending standards was observed. Another factor pushing credit growth was the signing of the Treaty of accession to the EU in 2005, which affected positively investor confidence about the development prospects of the country.

In this context, operating in a currency board and being unable to set interest rates, the Bulgarian National Bank pursued a consistent counter-cyclical policy mostly with macro-prudential and supervisory measures aimed at ensuring the stability of the banking system and at containing rapid credit growth. In the years of high economic growth before 2008, the BNB imposed very strict and conservative regulations for capital, liquidity, risk classifications and provisioning. Some of the macro-prudential measures were related to the conducting of a more restrictive policy regarding banking license issuance, the extension of the deposit base on which minimum reserve requirements (MRR) are calculated or the tightening of banking supervision through different prudential measures. In April 2005 BNB introduced administrative credit limits (credit ceilings), which were effective till January 2007. Banks whose quarterly credit growth exceeded the reference values set by the central bank had to hold additional minimum reserves with the BNB. Following the introduction of the credit ceilings there was an improvement in banks' balances and reduction of the credit risk in the banking system. Certain moderation of credit growth was also observed. After the administrative measures were abolished in the beginning of 2007, credit growth started accelerating again and reached 62.5% at the end of the year. Continuing to conduct a consistent counter-cyclical policy, BNB introduced an increase of the MRR ratio from 8% to 12% in September 2007.

Towards the end of 2008 and following the Lehman Brothers bankruptcy banks' behaviour changed. Parent banks reduced the availability of funds provided for market expansion. The Bulgarian banks tightened their credit standards and started to finance their activities mostly through domestic recourses. Since the end of 2008 growth of lending to the private sector slowed down significantly, reflecting the intensification of the global financial and economic crisis. The Bulgarian economy was affected through increased uncertainty on the international financial markets, lower foreign capital inflows and declining external demand. During the economic downturn, the Bulgarian National Bank continued to conduct a counter-cyclical policy, taking a number of measures in late 2008 and

2009, which aimed at providing greater liquidity management flexibility of commercial banks using liquidity buffers created in previous years. Part of the measures were related to the easing of minimum reserve requirements' regulations and included the recognition of 50% of the cash balances as reserve assets and the reduction of the MRR rate from 12% to 10%, followed by a reduction of the MRR rate to 5% for funds attracted from non-residents and to 0% for government's deposits, collateralized with government securities. After 1 January 2009 the average effective minimum reserve requirement for the banking system fell to some 7 per cent and the overall effect of these BNB measures was a release of liquidity to banks. Other measures taken by the Bulgarian National Bank as a response to the crisis concerned the easing of the loan classification and provisioning rules. These measures aimed to ease credit institutions in negotiating the credit conditions and in converging with the international practices of the more conservative approach applied so far for the classification and loan loss provisions. In this manner, more benevolent conditions were created for banks to be flexible with their viable customers who were experiencing temporary difficulties in a harsh economic situation.

4. Survey results for Bulgaria

Against the background of the banking system and credit developments before and during the financial crisis, described in the previous chapter, in this section we provide an overview of the main results of the bank lending survey for Bulgaria. The questions in the survey concern either developments in credit standards or in demand for loans.[1] First, we present these developments for the period from 2003 Q4 till 2014 Q4. Furthermore, we discuss the contributing factors put forward by the banks surveyed in more detail. Finally, we compare the results of the bank lending survey with information collected from other sources. The analysis covers lending to enterprises as well as lending to private households. Lending to enterprises is further classified into lending for short-term purposes and lending for long-term purposes while lending to households is classified into lending for house purchase and lending for consumer credit.

[1] For details concerning the structure of the bank lending survey see Annex I: Structure and implementation of the BLS, p.43.

4.1. Lending to enterprises

As the time series cover a longer period of time concerning short-term loans and long-term loans to enterprises in comparison to total lending to firms we will focus our analysis on the two types of loans separately.[2] This will enable us to include the recession years in the analysis in order to make more comprehensive conclusions. For the aims of the following analysis we define the recession period as the period from the third quarter of 2008 till the fourth quarter of 2009, when based on seasonally adjusted data on quarterly growth Bulgaria's GDP decreased. By post-recession period we mean the period from the first quarter of 2010 till present. It is important to bear in mind that the so defined recession period for Bulgaria is not identical with the period of the global financial and economic crisis as from the point of view of other countries. The first signs of the crisis were present in the USA in late 2007 and early 2008, but in Bulgaria they showed several quarters later. Bulgarian commercial banks did not have an exposure to securities tied to the US real estate market, which plummeted in 2007 damaging financial institutions globally. The crisis in Bulgaria was channeled through the real economy and was a consequence of increased uncertainty on global financial markets, which led to lower foreign capital inflows and declining external demand.

For the aims of our analysis, in the charts below, which show developments in credit standards and in demand for loans to enterprises, we explicitly indicate the so defined recession period for Bulgaria and the period in which the administrative credit limits (credit ceilings) were effective.[3]

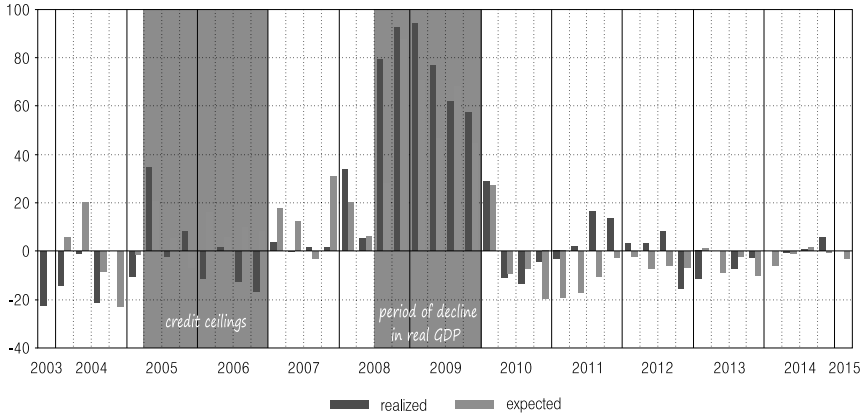
[2] For the period from the fourth quarter of 2003 to the present day, the BLS included questions on demand and credit standards separately for short-term and long-term corporate loans. The BNB has included questions on demand and credit standards for total corporate loans and consumer and housing loans to households since the first quarter of 2010.

[3] See Section 3, p. 9.

Chart 2: Changes in credit standards for loans to enterprises

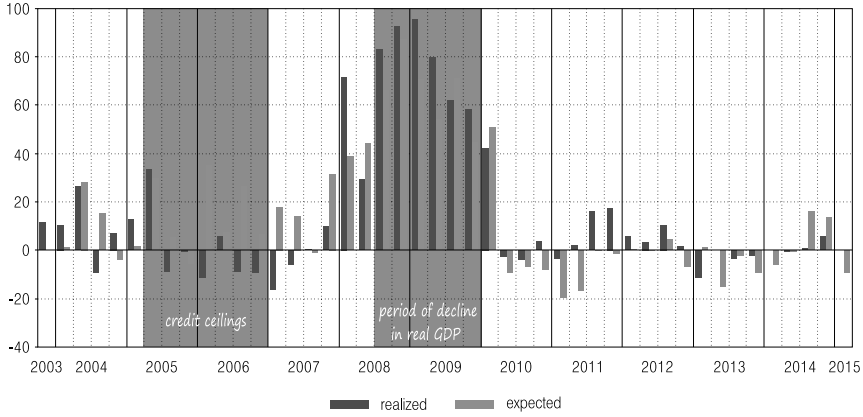
Short-term loans to enterprises

(net balance of opinions, percentage points)



Long-term loans to enterprises

(net balance of opinions, percentage points)



Note: The balance of opinions is defined as a difference in percentage points between the percentage of banks responding 'tightened' ('considerably' and 'somewhat') and the percentage of banks responding 'eased' ('considerably' and 'somewhat'). All bank responses are weighted by the bank's market share in lending to non-financial corporations for the relevant quarter.

'Realized' values refer to the period in which the survey was conducted. 'Expected' values are the net percentages calculated from the responses given by the banks in the previous survey.

Source: BNB – Bank Lending Survey.

Chart 2 shows how credit standards applied to the approval of loans to enterprises changed in the period from 2003 Q4 until 2014 Q4. In the years before the global financial and economic crisis, generally a net easing of credit standards was observed in respect to short-term loans to enterprises. Concerning long-term loans a net tightening of standards was reported in the first several rounds of the bank lending survey and a net easing afterwards. From the third quarter of 2008 till the first quarter of 2010 banks tightened strongly credit standards applied to the approval of short-term as well as long-term loans to enterprises. In the post-crisis years banks did not undertake any serious easing of standards. Easing of credit standards was observed only in respect to loan interest rates and to a lesser extent in respect to fees and commissions, which can be explained by the high competition from other banks. Concerning the maximum size of the loan, the premium on riskier loans and collateral requirements, standards remained tighter (see Chart 12 in Annex II). Expectations of banks concerning developments in their lending policy were generally in line with the actual outcomes in most of the period under consideration.

Concerning banks' responses with regard to changes in demand for loans, a net increase of loan demand from enterprises was observed till the end of 2008, followed by a net decrease in 2009 and the first quarter of 2010 (see Chart 3)[4]. In the post-recession period loan demand started growing again (more pronounced with respect to short-term loans), but growth was more slowly compared to the pre-crisis years. The certain recovery of loan demand from enterprises and the lack of considerable easing of banks' credit standards in the after-crisis years may misleadingly lead to the conclusion that low credit growth from 2010 till 2014 was supply-side driven.[5] However, it should be borne in mind, that growth of lending to enterprises concerns the stock of loans, including the maturing loans. Concurrently, when looking at the volumes of extended new loans to enterprises, they have returned close to their pre-crisis levels.[6] Expectations regarding the development of credit demand were generally in line with the actual outcomes excepting the recession period, when banks did not expect demand for loans to decrease as it did in fact.

[4] By net increase/decrease in demand for loans is meant a positive/negative value for the net percentage of banks reporting an increase in loan demand.

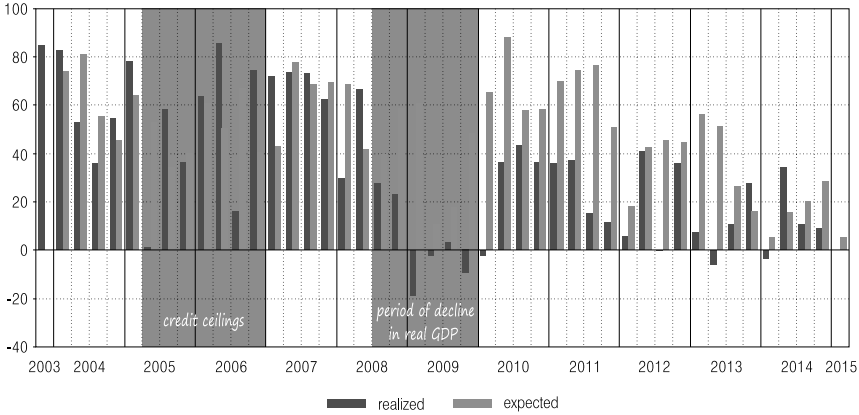
[5] The average annual growth of claims to non-financial corporations in the period 2010-2014 came to 2.8% compared to 38.6% average for the period 2003-2008.

[6] See Chart 14 in Annex II.

Chart 3: Changes in demand for loans to enterprises

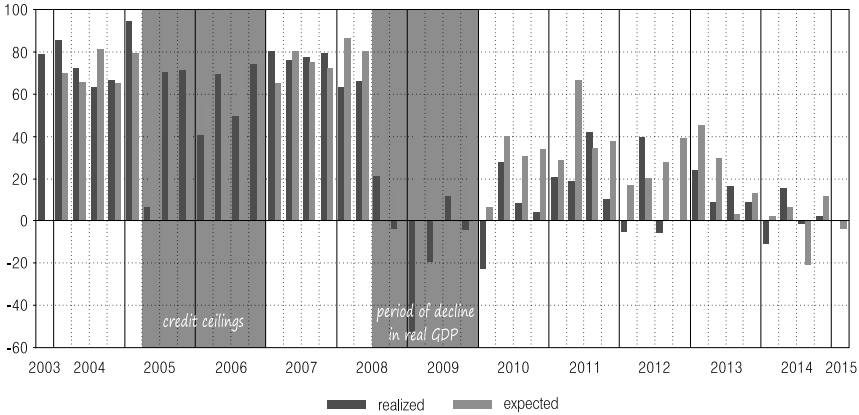
Short-term loans to enterprises

(net balance of opinions, percentage points)



Long-term loans to enterprises

(net balance of opinions, percentage points)



Note: The balance of opinions is defined as a difference in percentage points between the percentage of banks responding 'increased' ('considerably' and 'somewhat'), and the percentage of banks responding 'decreased' ('considerably' and 'somewhat'). All bank responses are weighted by the bank's market share in lending to non-financial corporations for the relevant quarter.

'Realized' values refer to the period in which the survey was conducted. 'Expected' values are the net percentages calculated from the responses given by the banks in the previous survey.

Source: BNB – Bank Lending Survey.

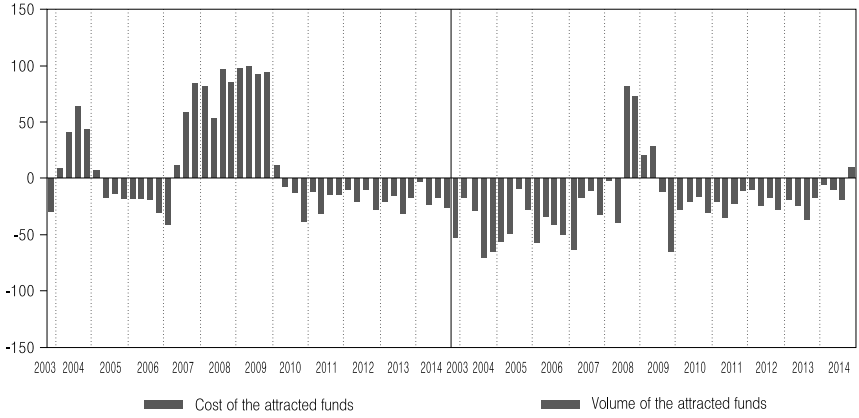
Considering the reasons behind the tightening or easing of credit standards, Chart 4 shows the factors affecting credit standards for approving loans to enterprises. In the pre-crisis years almost all of the factors, included in the bank lending survey, contributed to the easing of credit standards except the credit risk and the collateral risk. During the recession years, the main reasons behind the tightening of credit standards were linked to the increasing cost of attracted funds and the perception of risk. Against the background of heightened uncertainty related to the general economic situation, banks started competing for the attraction of funds from residents, which resulted in higher cost of financing. In the post-recession period, the factors contributing most to the easing of credit standards were related to the stronger competition from other banks, the increased volume and the declining cost of attracted funds, as banks had already accumulated enough liquidity, while perception of risk continued to play a negative role in the background of economic uncertainty.

Concerning the factors affecting demand for loans to enterprises, during the whole period under consideration demand for loans was increasing for financing needs of inventories and working capital, but at a decelerating pace. Before the crisis firms demanded loans for investment purposes, while during the recession years fixed investment was subdued and consequently credit demand decreased. In the post-recession period loan demand for investment purposes recovered slightly, but was far away from pre-crisis levels. A factor which had a positive contribution to the demand for loans to enterprises during the recession was the limited access of firms to alternative sources of finance, such as internal financing or loans from non-banking institutions.

Chart 4: Factors contributing to changes in banks' lending policies

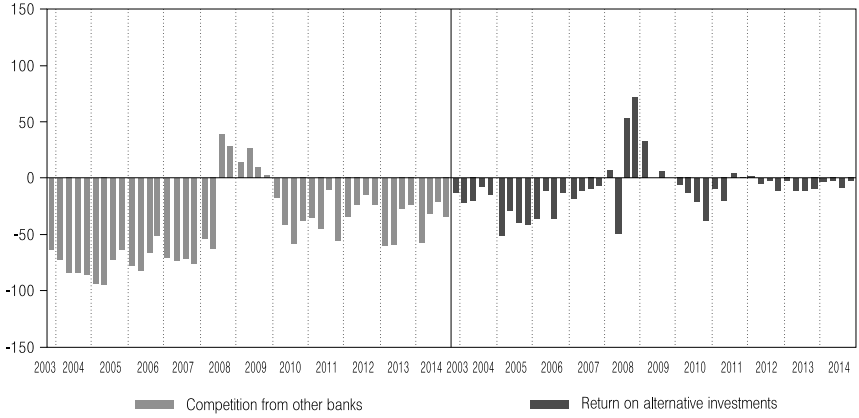
a) Cost and volume of funds

(net balance of opinions, percentage points)



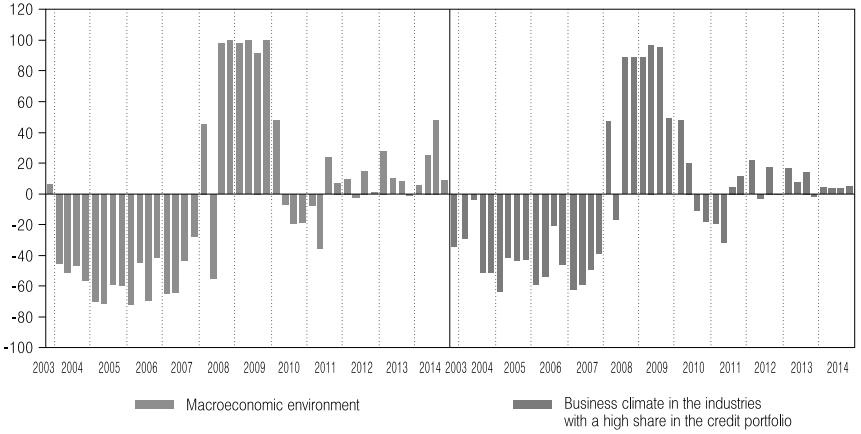
b) Competition and alternative investment

(net balance of opinions, percentage points)

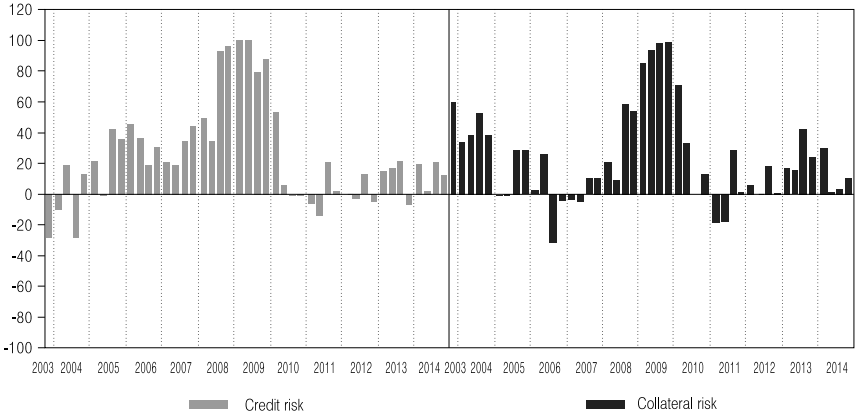


c) Perception of risk

(net balance of opinions, percentage points)



(net balance of opinions, percentage points)



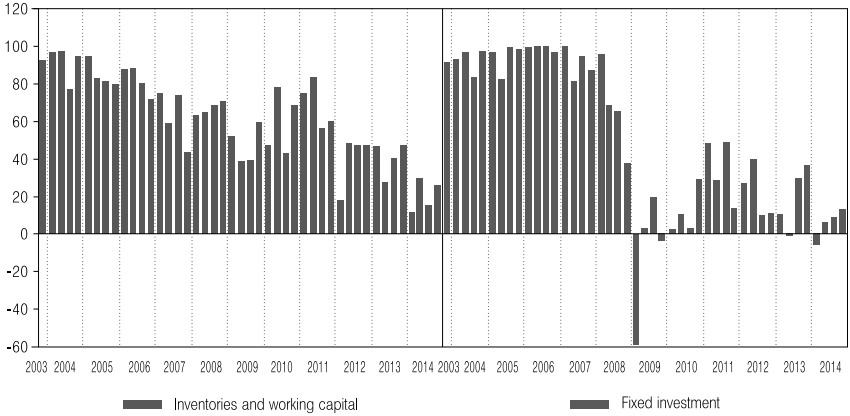
Note: The balance of opinions in responses about factors of credit standards is defined as a difference between the percentage of banks' responses for 'has contributed to tightening' ('considerably' and 'somewhat') and the percentage of banks' responses for 'has contributed to easing' ('considerably' and 'somewhat').

Source: BNB – Bank Lending Survey.

Chart 5: Factors contributing to changes in demand for loans to enterprises

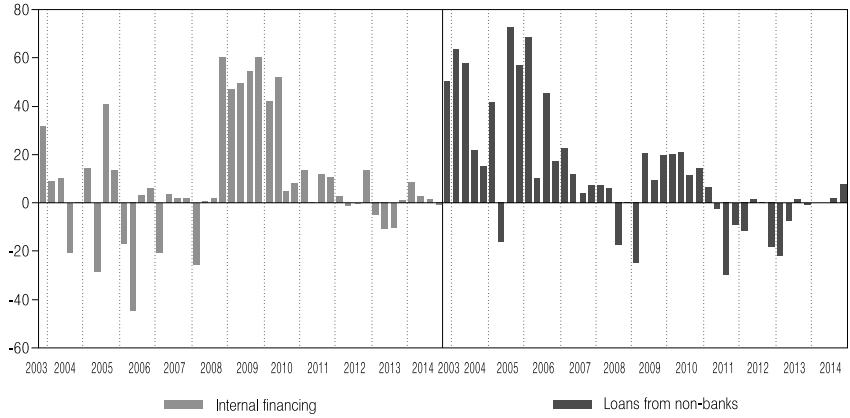
a) Financing needs

(net balance of opinions, percentage points)



b) Use of alternative sources of financing

(net balance of opinions, percentage points)



Note: The balance of opinions in responses about factors of loan demand is defined as a difference between the percentage of banks' responses for 'has contributed to growth' ('considerably' and 'somewhat') and the percentage of banks' responses for 'has contributed to a decrease' ('considerably' and 'somewhat').

Source: BNB – Bank Lending Survey.

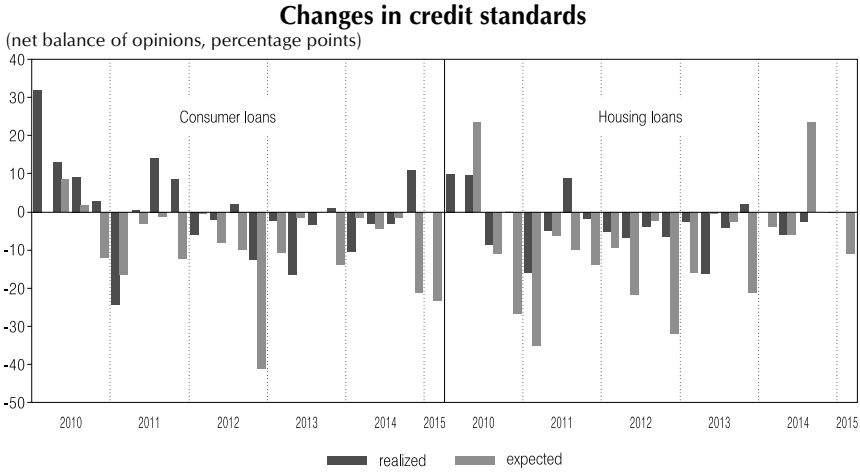
4.2. Lending to households

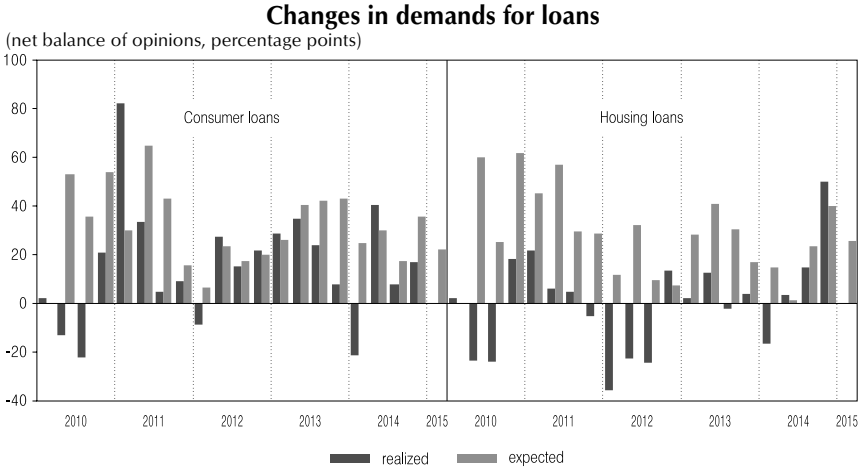
Questions concerning lending to households have been included in the bank lending survey since the first quarter of 2010.[7] Consequently, conclusions about developments in lending for consumer credit and for house purchase during the recession years cannot be drawn from the survey results. In the years after the crisis, survey results show that credit standards for approving loans to households generally eased, more pronounced concerning loans for house purposes (see Chart 6). Banks' expectations about their lending policy were generally in line with actual outcomes. Despite the easing of credit standards, demand for housing loans was decreasing from the last quarter of 2011 till the third quarter of 2012. In the quarters before and after that, changes in demand for loans for house purchase generally moved opposite to changes in credit standards. Demand for consumer loans was increasing during most of the period under consideration, whereas these developments were not always stimulated by banks' lending policy. Concurrently, banks' expectations about developments in credit demand did not always come into realization.

With regard to conditions and terms for approving loans to households, during the period under consideration banks eased their lending policy mostly in respect to loan interest rates, the interest spread and the fees and commissions for approving and managing the loans (see Chart 13 in Annex II). Furthermore, from the first quarter of 2012 banks eased credit standards in respect to the maximum size of loans for consumer credit. Standards were tightened concerning the premium on riskier loans and collateral requirements.

[7] For more detailed description of the structure of the Bank Lending Survey see Annex I: Structure and implementation of the Bank Lending Survey, p. 43.

Chart 6: Credit standards and demand for loans for consumer credit and house purchase





Note: The balance of opinions is defined as a difference in percentage points between the percentage of banks responding 'tightened/increased' ('considerably' and 'somewhat') and the percentage of banks responding 'eased/decreased' ('considerably' and 'somewhat'). All bank responses are weighted by the bank's market share in lending to households for the relevant quarter.

'Realized' values refer to the period in which the survey was conducted. 'Expected' values are the net percentages calculated from the responses given by the banks in the previous survey.

Source: BNB – Bank Lending Survey..

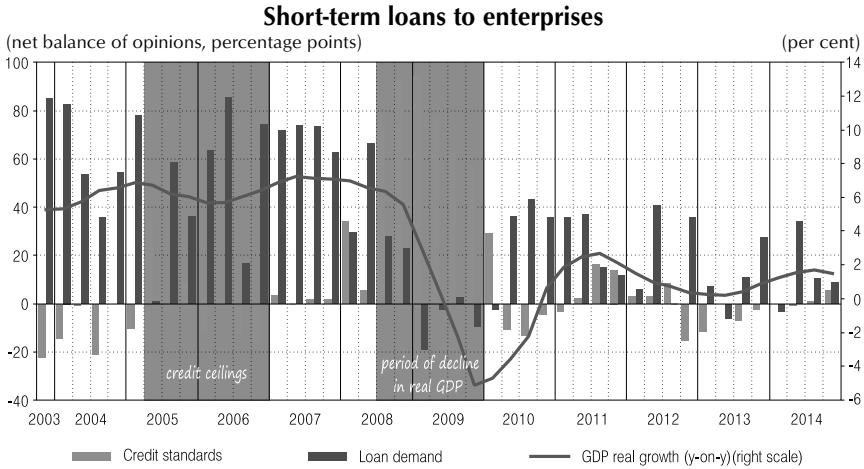
4.3. Comparison of bank lending survey data with other indicators

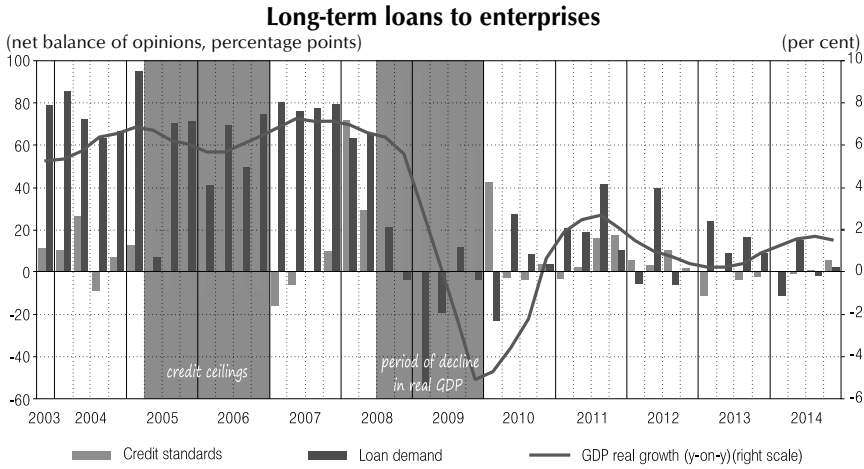
This section aims at comparing some of the reported variables in the survey with information from other sources (real GDP growth, loan growth, gross fixed capital formation and industrial confidence). The purpose of this analysis is to assess the information content of BLS results in relation to other macroeconomic and financial data.

Credit standards, among other factors such as interest rates, exchange rates, consumer or business confidence, may be linked to economic activity. To the extent that credit availability depends on lenders' standards, a tightening of banks' lending policy should cause a decline in spending by firms and households that depend on banks for credit and this in turn should lead to lower economic activity. Chart 7 presents developments in real activity alongside those in banks' credit standards and in demand for loans to enterprises.

In the years before the global financial crisis, a net easing of credit standards in respect to short-term loans to enterprises was generally observed. Concerning long-term loans to enterprises a net tightening of banks' lending policy was reported in the period 2003Q4–2005Q2 and a net easing afterwards. Indeed, taking into account the very tight initial credit standards, the cumulative effect in this period was easing of banks' lending policy towards enterprises, driven by supply factors and competition for market share. At that time banks had plain access to foreign financing. Financial resources were provided by parent banks to their subsidiary banks and their branches in Bulgaria, with the aim of boosting their market share in the region because of the significant return on investment. At the same time demand for short-term as well as for long-term loans was increasing rapidly. In line with developments in credit standards and credit demand real activity was strong, averaging 6.2% for the period 2003–2007. Banks started tightening their lending policy from the first quarter of 2008, shortly after the first signs of the global financial crisis had appeared and demand for loans started declining several quarters later. A possible explanation of these developments is the fact that banks could react more rapidly to what was happening on international financial markets and change their lending policy accordingly. At the same time, a longer period of time was needed to see a change in firms' behavior. The first signs of a slight improvement of economic activity could be observed from the first quarter of 2010 and credit demand started growing again one quarter later. Banks also started easing their lending policy from the second quarter of 2010. During the post-recession period demand for loans from

Chart 7: Comparison of BLS data on credit standards and demand for loans to enterprises and real GDP growth





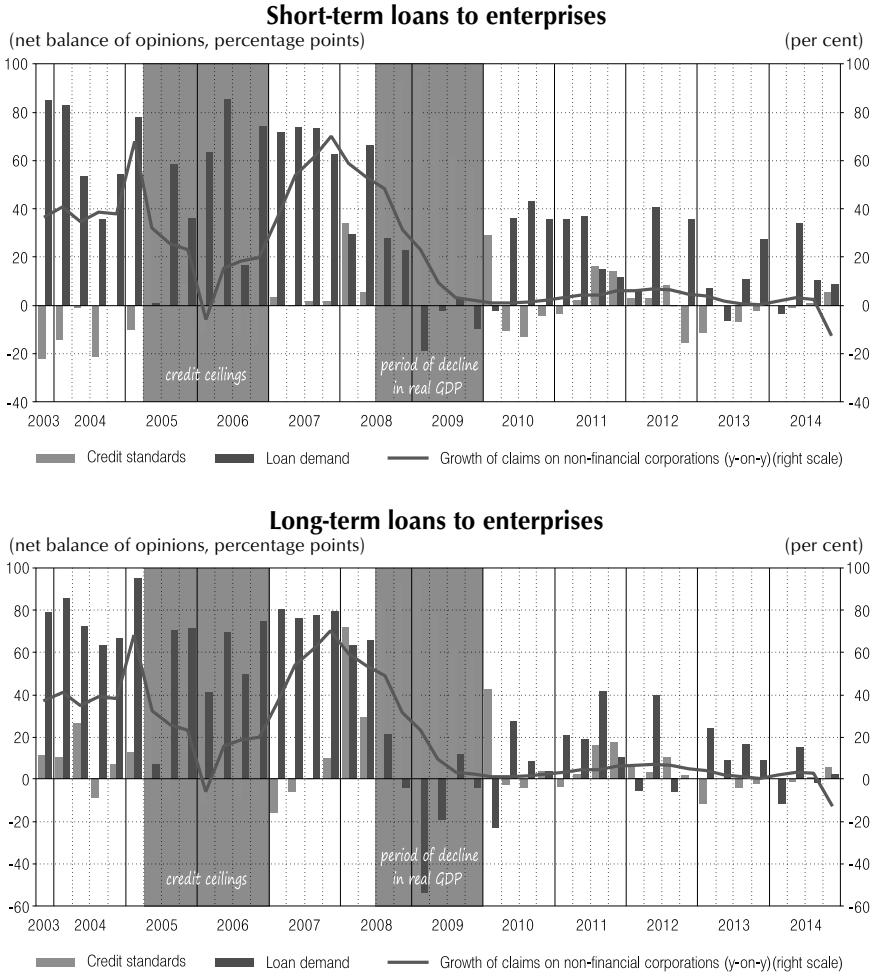
Source: BNB – Bank Lending Survey, NSI.

enterprises has been increasing in most of the time, while banks' lending policy has been not very consistent as there were periods of easier as well as of tighter lending standards.

One of the objectives of the bank lending survey is to complement information, retrieved from other sources, such as the monetary statistics. It can be expected, that a high net percentage of tightening of credit standards should be associated with low (if negative) lending growth.

In Chart 8 data from the bank lending survey is plotted together with data on the claims on non-financial corporations from the monetary statistics. As a matter of fact, in the period from the first quarter of 2008 till the first quarter of 2010 a high net tightening of credit standards was observed and at the same time the year-on-year growth of lending to non-financial corporations was posting a significant deceleration (from a peak of 70.2% in 2007 Q4 it came to around 1% at the beginning of 2010). However, the results of the bank lending survey show that the inverse relationship between tightening of credit standards and loan growth is not always apparent. For example the net tightening of standards with regard to long-term loans to enterprises over the first several rounds of the survey was associated with a net increase in demand for such loans according to banks' answers and the year-on-year growth of lending to NFC was not showing any signs of deceleration. Possible explanations for the increased loan demand from enterprises in this period are the optimistic expectations of firms for the medium-term economic outlook. With respect to short-term loans the relationship is more intuitive for the first several survey rounds. In the post-recession years there are also periods in which standards and credit growth were not moving in opposite directions. One possible reason for these results may be that banks' answers regard to short-term and long-term loans separately while the growth of lending to non-financial corporations concerns total loans to enterprises. However, if we look at banks' answers concerning total loans to enterprises, for which we have data since the first quarter of 2010, and compare them with data on lending to NFCs from the monetary statistics, the results do not show a very different picture. It is highly possible that in the after-crisis years many other factors besides credit standards, such as the uncertain economic environment, postponed investment by firms or unwillingness of enterprises to run up more debts have influenced credit growth.

Chart 8: Comparison of BLS data on credit standards and demand for loans to enterprises and growth of loans to NFCs



Note: In the fourth quarter of 2014 the year-on-year decline in claims on non-financial corporations is driven by the exclusion of Corporate Commercial Bank as a reporting unit from the monetary statistics since November 2014 after its banking license revocation.

Source: BNB – Bank Lending Survey and Monetary Statistics.

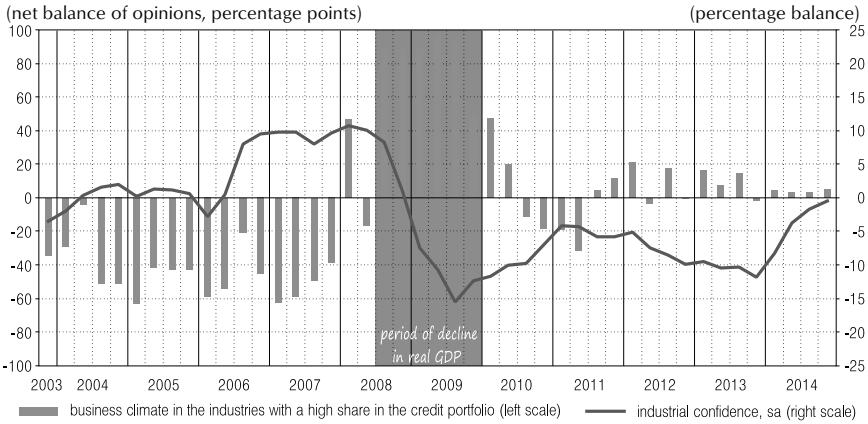
However, if we look at BLS data on demand for loans from enterprises and compare it with the growth of claims on non-financial corporations, there is much more systematic pattern in the directions they move. In the years before the global financial and economic crisis demand for loans from enterprises was high, stimulated by the favorable macroeconomic environment and high expected return on investment. At the same time rapid credit growth, as reported from the monetary statistics, was observed with the exception of the period from 2005 Q2 till 2006 Q1. The significant deceleration of growth of loans to non-financial corporations in 2005 and the beginning of 2006 was most certainly strongly affected by the introduction of the credit ceilings by the BNB and was not driven by declining loan demand.[8] During the recession years demand for loans started decreasing and credit growth was decelerating as well. In the post-recession period loan demand from enterprises returned to certain levels, while the growth of claims on non-financial corporations remained weak, but at least in positive territory and both indicators moved in the same direction.

Turning to the factors affecting credit standards, one of the reasons reported for the tightening of credit standards for loans to enterprises is the risk perception related to the business climate in the industries with a high share in the credit portfolio. Chart 9 compares the net percentage reported for the business climate with the industrial confidence indicator as reported by the European Commission's Business and Consumer Surveys.[9] In most of the period before the recession industrial confidence was positive and, at the same time, banks reported this factor as contributing to the easing of credit standards. Since the third quarter of 2008 the industrial confidence indicator started declining and even turned negative in the beginning of 2009. Along with the enhancement of risk perception banks reported a tightening of credit standards. In the post-crisis period, generally an improvement in industrial confidence, in the sense of a less negative value of the indicator, was associated with an easing of lending policy of banks, and a deterioration of the confidence indicator went along with tighter credit standards.

[8] For details see section 3.

[9] The industrial confidence indicator is the arithmetic average of the balances (in percentage points) of the answers to the questions on production expectations, order books and stocks of finished products. Balances are seasonally adjusted.

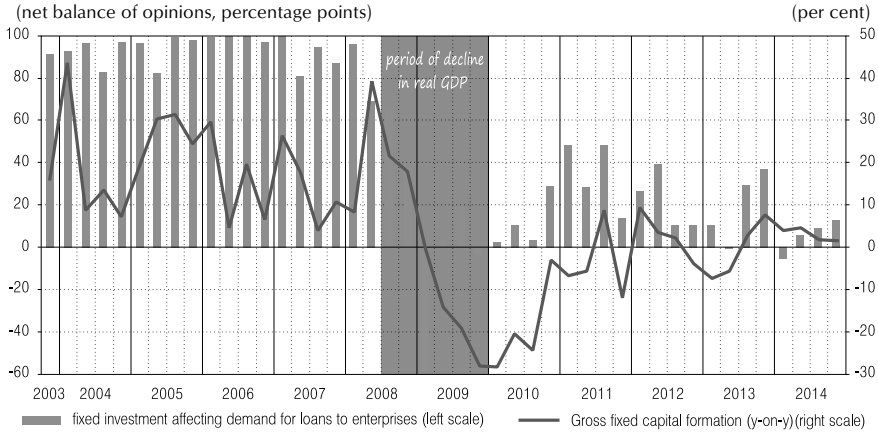
Chart 9: Comparison of BLS data and industrial confidence



Note: The balance of opinions is defined as a difference in percentage points between the percentage of banks responding "contributed considerably to tightening" and "contributed somewhat to tightening" and the percentage of banks responding "contributed somewhat to easing" and "contributed considerably to easing".

Source: BNB – Bank Lending Survey, EC.

Chart 10: Comparison of BLS data and gross fixed capital formation



Note: The balance of opinions is defined as a difference in percentage points between the percentage of banks responding "contributed considerably to higher demand" and "contributed somewhat to higher demand" and the percentage of banks responding "contributed somewhat to lower demand" and "contributed considerably to lower demand".

Source: BNB – Bank Lending Survey, NSI.

Turning to the demand side, the bank lending survey provides information on the reasons driving the demand for loans from both enterprises and households. In the pre-crisis period, almost all of the banks, participating in the bank lending survey, reported that financing needs related to fixed investment contributed to higher demand for loans from enterprises. During the recession, in the background of uncertain macroeconomic environment, demand for bank loans for financing investment opportunities declined, and recovered to certain levels in the period thereafter. Chart 10 compares this information from the bank lending survey with the growth rate of gross fixed capital formation, which is the GDP component that is mostly related to investment.

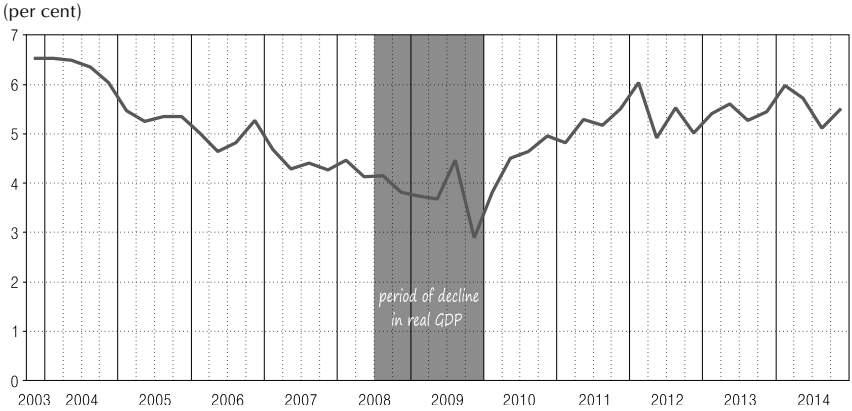
This chart shows that both indicators move in the same direction. High demand for loans from enterprises for investment purposes before the crisis was associated with comparatively high growth in gross fixed capital formation. At the same time lower credit demand for financing fixed investment, as reported in the bank lending survey, was accompanied by lower or even negative growth in gross fixed capital formation in the period from the fourth quarter of 2008 till the third quarter of 2010.

The inference from the graphical analysis displayed above is that there is comparability of data obtained from the bank lending survey with macroeconomic data, collected from other sources like GDP growth, loan growth, investment or industrial confidence. In the next section we will try to examine empirically the information content of the bank lending survey results by using them as explanatory variables for credit developments. Most certainly credit growth cannot be explained entirely by survey results. Therefore, along with survey data, we include in the empirical analysis other variables like real GDP growth, the spread between interest rates on loans and deposits of enterprises, the capital-to-asset ratio, bad and restructured loans as a share of total loans[10] and the business climate. As can be seen from Chart 11, during the recession period when the business climate was starting to deteriorate sharply, there is evidence for an increasing share of bad and restructured loans and declining banks' profit margins. Banks tried to hedge against the uncertainty and the deteriorating economic environment by increasing their capital buffers. The decrease of profit margins of banks was partly due to the significant increase of interest rates, which banks were ready to pay to attract more deposits from residents in the background of reduced access to international financial

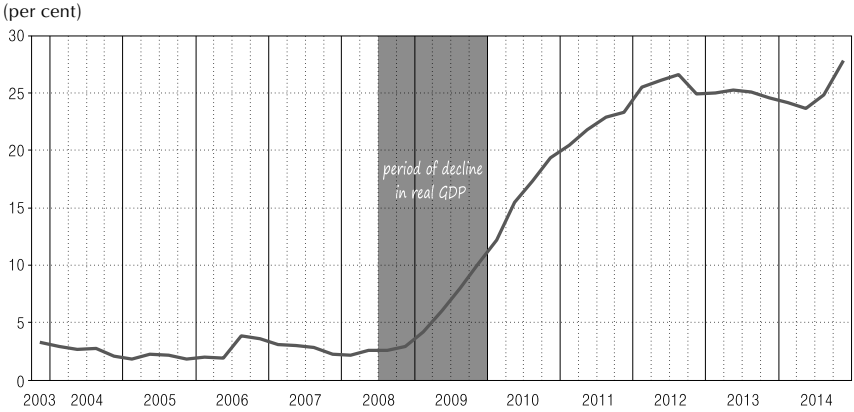
[10] Data on bad and restructured loans is taken from the monetary statistics. See also footnote 13.

Chart 11: Indicators used in the empirical analysis

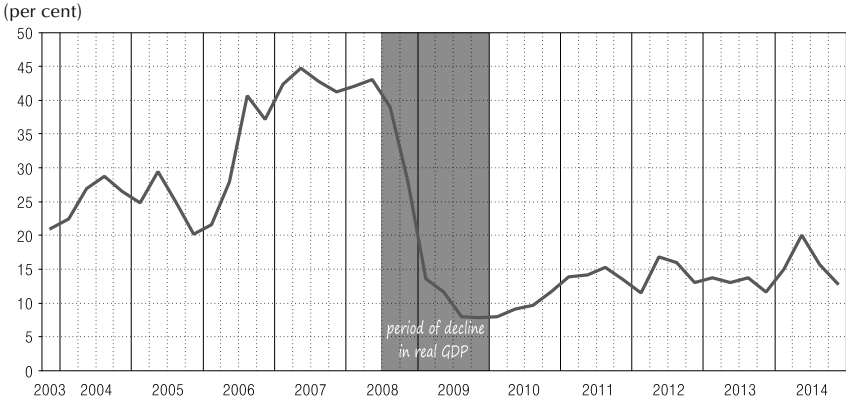
Interest rate spread



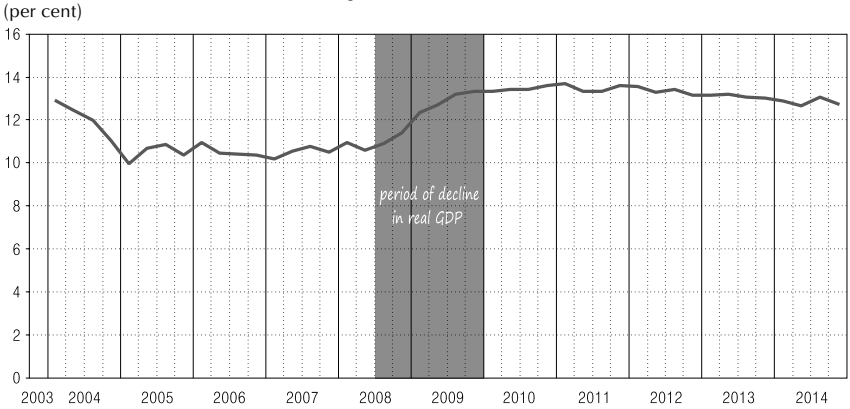
Bad and restructured loans



Business climate



Capital-to-asset ratio



The interest rate spread is defined as the spread between the average weighted interest rates on loans to non-financial corporations and the average weighted interest rates on deposits of non-financial corporations. Bad and restructured loans are defined as the share of loans to NFC with impaired performance past-due over 90 days and restructured loans in total loans to enterprises. Data on bad and restructured loan is provided by the monetary statistics. The business climate indicator is taken from the NSI tendency surveys. The capital-to-asset ratio is the ratio of bank capital to bank assets for the banking system as a whole.

Source: BNB, NSI.

markets. In the post-recession period the profit margins of commercial banks returned to certain levels, as faced to high accumulated liquidity they started decreasing deposit interest rates again. After reaching a capital-to-asset ratio of around 13% banks kept the level of capitalization close to this percentage. As a consequence of the worsened economic environment, firms started to experience difficulties in financing their investments and to repay their obligations to banks, which translated into increasing share of bad and restructured loans even after the crisis period.

Using data obtained from the bank lending survey and combining it with these additional variables, which could possibly explain changes in credit developments we will try to examine the information content of BLS results for growth of lending to enterprises. The analysis will be done first at the macro level and subsequently at the micro level using data by individual banks.

5. Empirical evidence

5.1. Macro level

As already mentioned above, definitive conclusions about what are the exact determinants of changes in bank lending cannot be drawn from the available statistics. Since there is only a limited possibility of making a clear-cut distinction between supply and demand variables using macroeconomic measurement variables, typically in loan equations are used approximation values such as GDP or investment for the demand side as well as an interest spread to capture the supply factors. In this respect, the bank lending survey can provide valuable information for a separate treatment of loan demand and loan supply as determinants in a loan equation. The net balances of banks' responses in respect to loan demand and credit standards for approving loans can be used as alternative indicators of a change in the supply of credit ($\Delta supply_t$), and of an adjustment of the demand for credit ($\Delta demand_t$), respectively.[11] In this section, using data on aggregate lending to enterprises (claims on non-financial corporations from the monetary statistics) and combining it with the results from the bank lending survey, we will try to make a distinction between loan supply-side and loan demand-side factors,

[11] Positive values for the net balances indicate an increase in demand for loans or a tightening of credit standards.

affecting the actual growth of credit. For the purpose of this analysis, we will use the following equations:[12]

$$(1) \qquad \Delta \ ln\ K_t = \beta_0 + \beta_1\ \Delta demand_t\_sh + \beta_2\ \Delta supply_t\_sh + \varepsilon_t \quad and$$

$$(2) \qquad \Delta \ ln\ K_t = \beta_0 + \beta_1\ \Delta demand_t\_lg + \beta_2\ \Delta supply_t\_lg + \varepsilon_t$$

where, the dependent variable $\Delta\ ln\ K_t$ is the growth rate of claims on non-financial corporations, $\Delta demand_t\_sh$ and $\Delta supply_t\_sh$ are the net balances of banks' responses to the BLS questions on the change in the demand and in credit standards in respect to short-term loans to enterprises, $\Delta demand_t\_lg$ and $\Delta supply_t\_lg$ are the net balances of banks' responses to the BLS questions on the change in the demand and in credit standards in respect to long-term loans to enterprises. The expected signs are positive for the coefficients β_1 and negative for the coefficients β_2. Cross correlations between the above defined BLS indicators and growth of claims on non-financial corporations at various lags (-) and leads (+) are presented in Table 4 in Annex II and tests for stationarity are reported in Table 6 in Annex II. The regression equations are estimated using the ordinary least squares method. Initially, only survey results are included in the regression, and subsequently additional explanatory variables, such as quarter-on-quarter seasonally adjusted real GDP growth (Δ ln GDP), interest spread defined as the difference between weighted average lending rates and weighted average deposit rates for non-financial corporations, the share of bad and restructured loans in the total amount of loans to non-financial corporations (Δ BRL)[13], business climate and banking system capital to assets ratio. Cross correlations between growth of claims on non-financial corporations and the additional explanatory variables at various lags (-) and leads (+) are presented in Table 5 in Annex II and tests for stationarity in Table 6 in Annex II. To deal with problems of normal distribution of the

[12] The approach we follow in this section is similar to the one used in the Monthly Report of Deutsche Bundesbank (January 2009).

[13] The regression analysis is based on monetary statistics data on loans, which are restructured and with impaired performance past-due over 90 days, due to available data time series for the whole period under review (fourth quarter of 2003 to fourth quarter of 2014). It should be stated that in monetary statistics banks provide aggregated data on these loans, because detailed data on the exposures according to their past-due periods are not collected for the purpose of these statistics. In accordance with the international practice, reporting of monetary statistics differs from supervisory reporting, including the reporting of loans, which are restructured or with impaired performance. Therefore, the aggregated data on loans which are restructured and with impaired performance past-due over 90 days represent neither the total loans with impaired performance, nor the share of loans with impaired performance past-due over 90 days.

Table 1: Dependent Variable: Δ ln Claims on Non-financial Corporations (Δ ln K_t)

Explanatory Variables	Short-term loans to corporations		Long-term loans to corporations	
Constant	0.01 (0.01)	0.01 (0.01)	0.01* (0.01)	0.00 (0.01)
Δ credit demand (-1)	0.05** (0.02)	0.03 (0.03)	0.07*** (0.02)	0.07*** (0.03)
Δ credit supply (-1)	0.02 (0.03)	0.01 (0.02)	0.01 (0.02)	-0.01 (0.02)
d_2005q1	0.20** (0.03)	0.26*** (0.03)	0.20*** (0.03)	0.25*** (0.03)
d_2005q2	-0.32** (0.04)	-0.27*** (0.04)	-0.31*** (0.04)	-0.26*** (0.04)
d_2005q4	-0.17*** (0.03)	-0.17*** (0.03)	-0.16*** (0.03)	-0.16*** (0.03)
Δ BRL (-1)		-0.55 (0.56)		-0.47 (0.52)
Δ ln GDP (-1)		1.20* (0.71)		0.34 (0.73)
Δ Business climate (-1)		0.00 (0.00)		0.00 (0.00)
Δ Capital/Assets (-1)		5.41*** (1.78)		5.46*** (1.64)
Δ Interest spread (-1)		0.58 (1.11)		0.98 (1.00)
Δ ln K_t (-1)	0.38*** (0.12)	0.43*** (0.11)	0.24** (0.12)	0.29** (0.12)
R^2	0.81	0.88	0.84	0.90
S.E. of regression	0.03	0.03	0.03	0.03
Jarque-Bera test	0.06	0.58	0.11	0.67
Breusch-Godfrey LM test	0.40	0.90	0.97	0.68
Durbin-Watson test	1.61	2.03	1.72	2.05
Breusch-Pagan-Godfrey test	0.04	0.01	0.56	0.25
Number of observations	45	42	45	42

**** indicates significance at the 1% level, ** at the 5% level,* at the 10% level, standard errors in parenthesis.*

Notes: Three dummies are included in the specifications: d_2005q1, d_2005q2 and d_2014q4 for the first and second quarters of 2005, and the fourth quarter of 2014.

The results of the following test are presented in the table: Jarque-Bera normality test for distribution of residuals with null hypothesis: normal distribution, p-value is presented; Breusch-Godfrey LM test for serial correlation with null hypothesis: a lack of serial correlation in the residuals, p-value is presented; Durbin-Watson test for serial correlation in the residuals with DW statistics presented; Breusch-Pagan-Godfrey test for heteroscedasticity with null hypothesis: a lack of heteroscedasticity, p-value is presented.

According to Jarque-Bera criterion for normality of residuals, they are normally distributed. While the tests indicate that no serial correlation in the residuals is observed, Breusch-Pagan-Godfrey test reveals problems with heteroscedasticity of residuals regarding short-term loans to corporations. When applying the White's procedure to clear heteroscedasticity, the significance of coefficients in front of explanatory variables remained unchanged. Therefore, it may be concluded that it has no effect on empirical assessment conclusions.

Source: BNB

residuals we include three dummies for 2005Q1, 2005Q2 and 2014Q4 in our specifications and to deal with problems of serial correlation we include one lag of the dependent variable. The main results of the empirical macro analysis are presented in Table 1.

The empirical analysis outcomes show that the variable recording the change in demand for loans by corporations is statistically significant for the growth of claims, both for short-term and long-term loans to corporations. These results remain unchanged, if demand significance in the current or previous period is tested (i.e. if the first lag of explanatory variable is taken into account). The inclusion of additional explanatory variables into the specifications has also no impact on the robustness of estimates. The coefficient in front of the variable recording the changes in demand for loans remains stable in the various specifications, moving within a range of 0.05 to 0.07, i.e. the 1 percentage point increase in demand for loans positively affects the growth of claims on non-financial corporations by 0.05–0.07 percentage points. Changes in credit standards have statistically insignificant effect on corporate loans dynamics. Among the additional explanatory variables, statistical significance for the growth of claims is found regarding real GDP growth and banking system capital to assets ratio. The coefficients in front of these variables have the expected positive signs and are relatively higher than those in front of the variables from the survey. The overall explanatory power of the equations is comparatively high: the explanatory variables explain between 80 and 90% of the variation of the dependent variable.

To test if our conclusions till now can change if we go down to the micro level, we will do the analysis taking into account individual banks' answers to the bank lending survey and matching them to individual volumes of loans, granted by each bank.

5.2. Individual banks

By matching BLS responses to aggregate data on lending, potential mismatch errors and inaccurate interpretations of the results could arise. To deal with this problem, we construct a panel, for which purpose we merge the individual banks' responses to the BLS-questions with individual data on lending amounts for the surveyed banks. In doing so, we guarantee that survey responses and loan data refer to the same panel of banks. Data on banks' lending amounts are drawn from the financial supervision reports and represent the end-of-quarter values of stocks. Complementary to the survey results, additional explanatory variables are added to the

panel. They comprise specific factors for each individual bank, such as interest spread between corporate loans and deposits by individual bank[14], individual bank capital to assets ratio[15], and variables that are common to all banks, as real GDP growth (quarter-on-quarter seasonally adjusted), business climate in Bulgaria and the share of bad and restructured loans in the total amount of loans to non-financial corporations[16].

The panel econometric analysis is carried out for unbalanced data panel comprising the period between the fourth quarter of 2003 and the last quarter of 2014, applying panel estimation with cross-section fixed effects to account for the unobserved variation among the banks. To examine the determinants of banks' lending to non-financial corporations, we estimate an equation of the following form:[17]

$$(3) \qquad \Delta \ln K_{i,t} = \alpha_i + \beta\,(L)\, BLS_{i,t} + \gamma\,(L)\, X_{(i)t} + \varepsilon_{i,t}$$

where, the dependent variable $\Delta \ln K_{i,t}$ is the first difference of the logarithm of loans to enterprises for bank *i* in period *t*. $BLS_{i,t}$ denotes a set of BLS indicators for loan supply and loan demand for bank *i* in period *t* and $X_{(i)t}$ is a vector with the additional macro and micro control variables mentioned above. Since the information content of the BLS indicators is of qualitative nature, they are included in our specifications as dummy variables. As regards loan demand and credit standards, two pairs of variables are designed for a decrease and an increase in loan demand by corporations and a tightening and an easing of credit standards respectively. Thus, specification equation (3) can be rewritten as:

$$(4) \qquad \Delta \ln K_{i,t} = \alpha_i + \beta_1\,(L)\, Demand\ decreased_{i,t} + \beta_2\,(L)\, Demand\ increased_{i,t} + \\ + \beta_3\,(L)\, Standards\ tightened_{i,t} + \beta_4\,(L)\, Standards\ eased_{i,t} + \gamma\,(L)\, X_{(i)t} + \varepsilon_{i,t}$$

where, for instance, the variable '*demand decreased*' takes the value 1 if bank *i* has reported a decrease in demand in period *t* (response categories

[14] Interest spread between corporate loans and deposits by individual bank is implicitly calculated, using the ratio of interest income on extended loans to average loans and the ratio of interest expenditure on attracted funds to the average amount of attracted funds.

[15] Data of the Banking Supervision Department on capital and assets of individual banks.

[16] Monetary statistics data. See also footnote 13 above.

[17] The approach used in this section is similar to the one in Blaes (2011).

'*decreased considerably*' or '*decreased somewhat*') and 0 otherwise. The variable 'standards tightened' takes the value 1 if bank i has reported a tightening of credit standards in period *t* (response categories '*tightened considerably*' or '*tightened somewhat*') and 0 otherwise. Similarly, the variables '*demand increased*' and '*standards eased*' are designed. The expected signs are negative for the coefficients β_1 and β_3, and positive for β_2 and β_4. We estimate six alternative specifications. We first estimate the impact of only BLS indicators on growth of lending to enterprises and include step by step the additional control variables afterwards. In Annex II (Table 7) we report cross correlations between loan growth and the additional macro and micro control variables at various lags (-) and leads (+). Tests for Unit Roots are presented in Table 8 in Annex II. The main results of the empirical micro analysis with respect to banks' answers concerning credit standards and demand for short-term loans are presented in Table 2, while those concerning credit standards and demand for long-term loans are shown in Table 3.

As can be seen from Table 2 (concerning short-term loans to enterprises) the constructed variables for "demand decreased" and "standards tightened" have the expected negative sign, meaning that lower credit demand or tighter banks' lending policy affect negatively the growth of loans to enterprises, while the corresponding variables for "demand increased" and "standards eased" show the expected positive sign. However, the results show that the BLS indicators, which are significant in explaining growth of lending to enterprises, are those for "demand decreased" and "standards tightened". As a matter of fact, the BLS indicator "demand decreased" is significant in all six specifications and the BLS indicator "standards tightened" is significant in two of them. The coefficients are broadly comparable among the different specifications used. In our baseline specification (1), the coefficient of "demand decreased", for example, indicates that a decrease in credit demand by 1 percentage point in period t-1 is associated with a decline of loan growth amounting to 0.09 percentage points in period t. The impact of this variable on lending remains robust, even when we include additional variables in the estimation. Regarding these additional control variables (the growth of real GDP, the interest spread by bank, the change in the capital-to-assets

Table 2: Dependent Variable: Δ ln Loans to Corporations (Δ ln K), Unbalanced Panel (OLS, Cross-Section Fixed Effects), short-term loans

Explanatory Variables	Short-term loans to corporations					
	(1)	(2)	(3)	(4)	(5)	(6)
Constant	0.05*** (0.01)	-0.01 (0.02)	-0.02 (0.02)	-0.02 (0.02)	-0.01 (0.02)	-0.01 (0.02)
Δ credit demand (-1) (decrease)	-0.09*** (0.03)	-0.09*** (0.03)	-0.08*** (0.03)	-0.08*** (0.03)	-0.08*** (0.03)	-0.08*** (0.03)
Δ credit demand (-1) (increase)	0.02 (0.02)	0.02 (0.02)	0.02 (0.02)	0.02 (0.02)	0.02 (0.02)	0.01 (0.02)
Δ credit standards (-1) (tightening)	-0.04* (0.02)	-0.04** (0.02)	-0.01 (0.02)	-0.01 (0.02)	0.00 (0.02)	0.01 (0.02)
Δ credit standards (-1) (easing)	0.03 (0.03)	0.03 (0.03)	0.02 (0.03)	0.03 (0.03)	0.03 (0.03)	0.03 (0.03)
Interest spread by bank		1.26*** (0.39)	0.86** (0.39)	0.82** (0.39)	0.83** (0.39)	1.04*** (0.39)
Δ ln real GDP			4.09*** (0.67)	4.29*** (0.67)	3.79*** (0.69)	2.49*** (0.80)
Δ Capital to assets				0.75*** (0.21)	0.73*** (0.21)	0.72*** (0.21)
Δ BRL (-1)					-2.00*** (0.69)	-2.32*** (0.69)
Δ Business climate						0.01*** (0.00)
Δ ln K (-1)	-0.22*** (1.11)	-0.23*** (1.11)	-0.24*** (1.11)	-0.24*** (1.11)	-0.24*** (1.11)	-0.24*** (1.00)
Periods	43	43	43	43	43	43
Cross-sections	41	41	41	41	41	41
Number of observations	1303	1301	1301	1301	1301	1301
R^2	0.08	0.09	0.12	0.13	0.13	0.14
DW	2.12	2.14	2.12	2.09	2.09	2.08

**** indicates significance at the 1% level, ** at the 5% level,* at the 10% level, standard errors in parenthesis.*

Notes: The results of the following test are presented in the table: Durbin Watson (DW) test for serial correlation is presented with DW statistics. The results of this test suggest that no serial correlation is observed in the residuals.

The number of cross-sections is determined by the existence of restructuring in the banking sector over the review period related to mergers at banks or emergence of new banks. In the case of mergers, individual banks before the merger and the emerged new bank thereafter are treated as separate units in the panel.

Source: BNB

Table 3: Dependent Variable: Δ ln Loans to Corporations (Δ ln K), Unbalanced Panel (OLS, Cross-Section Fixed Effects), long-term loans

Explanatory Variables	Short-term loans to corporations					
	(1)	(2)	(3)	(4)	(5)	(6)
Constant	0.05*** (0.01)	-0.01 (0.02)	-0.02 (0.02)	-0.02 (0.02)	-0.01 (0.02)	-0.01 (0.02)
Δ credit demand (-1) (decrease)	-0.07*** (0.03)	-0.07*** (0.03)	-0.05** (0.03)	-0.05** (0.03)	-0.04* (0.03)	-0.05* (0.03)
Δ credit demand (-1) (increase)	0.03 (0.02)	0.02 (0.02)	0.01 (0.02)	0.01 (0.02)	0.01 (0.02)	0.01 (0.02)
Δ credit standards (-1) (tightening)	-0.02 (0.02)	-0.02 (0.02)	0.01 (0.02)	0.01 (0.02)	0.01 (0.02)	0.02 (0.02)
Δ credit standards (-1) (easing)	0.04 (0.03)	0.03 (0.03)	0.03 (0.03)	0.03 (0.03)	0.03 (0.03)	0.03 (0.03)
Interest spread by bank		1.15*** (0.39)	0.80** (0.39)	0.76** (0.39)	0.78** (0.39)	1.00*** (0.39)
Δ ln real GDP				4.30*** (0.67)	3.80*** (0.69)	2.38*** (0.80)
Δ Capital to assets				0.76*** (0.21)	0.73*** (0.21)	0.72*** (0.21)
Δ BRL (-1)					-2.02*** (0.69)	-2.35*** (0.69)
Δ Business climate						0.01*** (0.00)
Δ ln K (-1)	-0.22*** (0.03)	-0.23*** (0.03)	-0.24*** (0.03)	-0.24*** (0.03)	-0.24*** (0.03)	-0.24*** (0.03)
Periods	43	43	43	43	43	43
Cross-sections	41	41	41	41	41	41
Number of observations	1303	1301	1301	1301	1301	1301
R^2	0.08	0.09	0.11	0.12	0.13	0.14
DW	2.13	2.14	2.12	2.09	2.09	2.07

**** indicates significance at the 1% level, ** at the 5% level,* at the 10% level, standard errors in parenthesis.*

Notes: The results of the following test are presented in the table: Durbin Watson (DW) test for serial correlation is presented with DW statistics. The results of this test suggest that no serial correlation is observed in the residuals.

The number of cross-sections is determined by the existence of restructuring in the banking sector over the review period related to mergers at banks or emergence of new banks. In the case of mergers, individual banks before the merger and the emerged new bank thereafter are treated as separate units in the panel.

Source: BNB

ratio or in bad and restructured loans), their estimated coefficients are highly significant in most specification variants. In particular, the coefficient of real GDP growth has the expected positive sign, meaning that positive developments in economic activity translate into higher growth of lending. The sign of the capital-to-assets ratio, included as a further micro variable is also positive, indicating that higher capitalization of the banking system and hence lower risk is a factor stimulating loan growth from the supply side. Non-performing loans as a macroeconomic risk variable have the expected negative impact on bank lending growth. According to the panel estimation results, an increase of NPL by one percentage point in period t-1 is associated with a decline in growth of loans to enterprises by around 2 percentage points in period t. The coefficient of the interest spread between corporate loans and deposits is positive, indicating that higher banking profits stimulate banks to increase the credit supply and translate into higher lending growth. The business climate, included as an explanation variable in the last specification, has the expected positive sign, but the coefficient is low, meaning that it does not explain much of the variance of credit growth to enterprises.

The panel estimation results with respect to banks' answers concerning long-term loans to enterprises (Table 3) show that the BLS indicators still have the expected signs, but the only variable which is statistically significant for the growth of corporate loans is the variable for "demand decreased". The coefficient in front of this BLS indicator remains broadly unchanged among the different specifications (it varies between -0.05 and -0.07). Concerning the additional macro and micro control variables their explanatory power for loan growth remains high. According to the estimation results corporate loan dynamics is positively influenced by real GDP growth, the interest rate spread and the capital to asset ratio, and negatively by the share of bad and restructured loans in total loans extended to enterprises.

In conclusion, the results of empirical micro analysis generally confirm macro analysis results. The variable recording the changes in demand for loans by corporations, particularly 'demand decreased', has the expected negative sign and is statistically significant for the growth of

corporate loans in all tested specifications. The coefficient in front of it is stable, ranging between -0.04 and -0.09. Overall, changes in credit standards have statistically insignificant effect on credit growth. These results are not affected by the inclusion of additional explanatory micro and macro variables. Besides the demand for loans by corporations, a statistical significance is found regarding real GDP growth, the share of bad and restructured loans in the total amount of loans to non-financial corporations, as well as for bank-specific factors, such as interest spread between loans and deposits and individual banks' capital to assets ratio. The coefficients in front of these variables display the expected signs: positive for real GDP growth, business climate indicator and individual banks' specific interest spread and capital to assets ratio, and negative in front of the share of bad and restructured loans in the total amount of corporate loans. The coefficients in front of these variables are relatively higher than those in front of the variables derived from the survey.

6. Conclusions

The main goal of this paper was to shed some additional light on the factors that influence credit growth on the demand as well as on the supply side in Bulgaria, with a more detailed focus on lending to non-financial corporations. Using data obtained from the regular quarterly bank lending survey, conducted by the BNB among commercial banks in Bulgaria and combining it with data from the monetary statistics, the banking supervision and with other macroeconomic variables like GDP growth or the business climate we undertook first a descriptive analyses, followed by an empirical assessment on a macro level and at the level of individual banks.

The general conclusion of the descriptive analysis suggests broadly similar trends in the change of demand and credit standards based on the survey results, on the one hand, and the growth dynamics of loans to non-financial corporations based on monetary statistics data, on the other hand. A similar conclusion can be made comparing the survey results with other macroeconomic indicators as real GDP growth, investment in fixed capital and the confidence in the industry sector based on the business situation survey.

Empirical analysis carried out on the basis of macro data and individual bank data for the 2003 to 2014 period show that changes in demand estimated by survey data have statistically significant effect on corporate

loans dynamics. The empirical research also reveals that important factors positively affecting corporate loans dynamics at both analyzing levels are real GDP growth and banks' capital to assets ratio. The analysis at individual banks level finds that statistically significant factors for the growth in corporate loans are also the improvement in business climate in Bulgaria, the decrease in the share of bad and restructured loans in the total amount of loans, and bank-specific factors such as interest spread between corporate loans and deposits.

References

Bassett, W.F., M.B. Chosak, J.C. Driscoll, and E. Zakrajsec. "Changes in bank lending standards and the macroeconomy." *Journal of Monetary Economics* 62, (2013): 23 – 40.

Berg, J., A. van Rixtel, A. Ferrando, G. de Bondt and S. Scopel. "The bank lending survey for the euro area." *ECB Occasional Paper No 23*, European Central Bank, 2005.

Blaes, B. "Bank-related loan supply factors during the crisis: an analysis based on the German bank lending survey." *Discussion Paper Series 1: Economic Studies No 31/2011*, Deutsche Bundesbank, 2011.

De Bondt, G., A. Maddaloni, J.-L. Peydro and S. Scopel. "The euro area bank lending survey matters: empirical evidence for credit and output growth." *ECB Working Paper Series No 1160*, *European Central Bank*, 2010.

Deutsche Bundesbank. "Bank Lending Survey: an interim assessment and current developments." *Monthly Report*, January (2009): 15 – 30.

Hempell, H.S. "New Bank Lending Survey of the Eurosystem: Interpretation and Use of First Results for Germany." *Journal of Business Cycle Measurement and Analysis* Vol. 1, No 3, (2004): 387 - 407.

Lown, C.S. and D.P. Morgan. "The Credit Cycle and the Business Cycle: New Findings Using the Loan Officer Opinion Survey." *Journal of Money, Credit, and Banking* Vol. 38, No 6, (2006): 1575 – 1597.

L**own, C.S., D.P. Morgan and S. Rohatgi.** "Listening to Loan Officers: The Impact of Commercial Credit Standards on Lending and Output." *FRBNY Economic Policy Review*, July 2000.

Mottiar, R. and A. Monks. "The Bank Lending Survey for Ireland." *Quarterly Bulletin No 2 2007*, *Central Bank & Financial Services Authority of Ireland*, (2007): 147 – 178.

Annex I

Structure and implementation of the Bank Lending Survey

The current questionnaire of BNB's Bank Lending Survey is consistent with the survey conducted by the ECB. It consists of 12 regular questions and is divided into two subsections. The first one concerns loans to enterprises (short-term loans and long-term loans) and the second one lending to households (consumer and housing loans). Parts of the questions are backward-looking and examine developments during the preceding three months. Changes in credit demand and the factors underlying these changes are covered by the survey. On the supply side questions concern changes in credit standards, their determinants, and changes in credit term and conditions. Furthermore, there is a forward-looking element in the survey whereby banks are asked to give an opinion on what changes they expect on both in their own lending policy and in customer demand during the next three months. By answering questions concerning changes in demand for loans and in credit standards banks have to choose between five options: 1 – decreased/tightened considerably, 2 – decreased/tightened somewhat, 3 – remained basically unchanged, 4 – increased/eased somewhat, 5 – increased/eased considerably. Regarding the factors affecting demand for loans or credit standards and terms banks are asked to attribute answers on a scale ranging from "- -", "-", "0", "+", "++" and NA.[18]

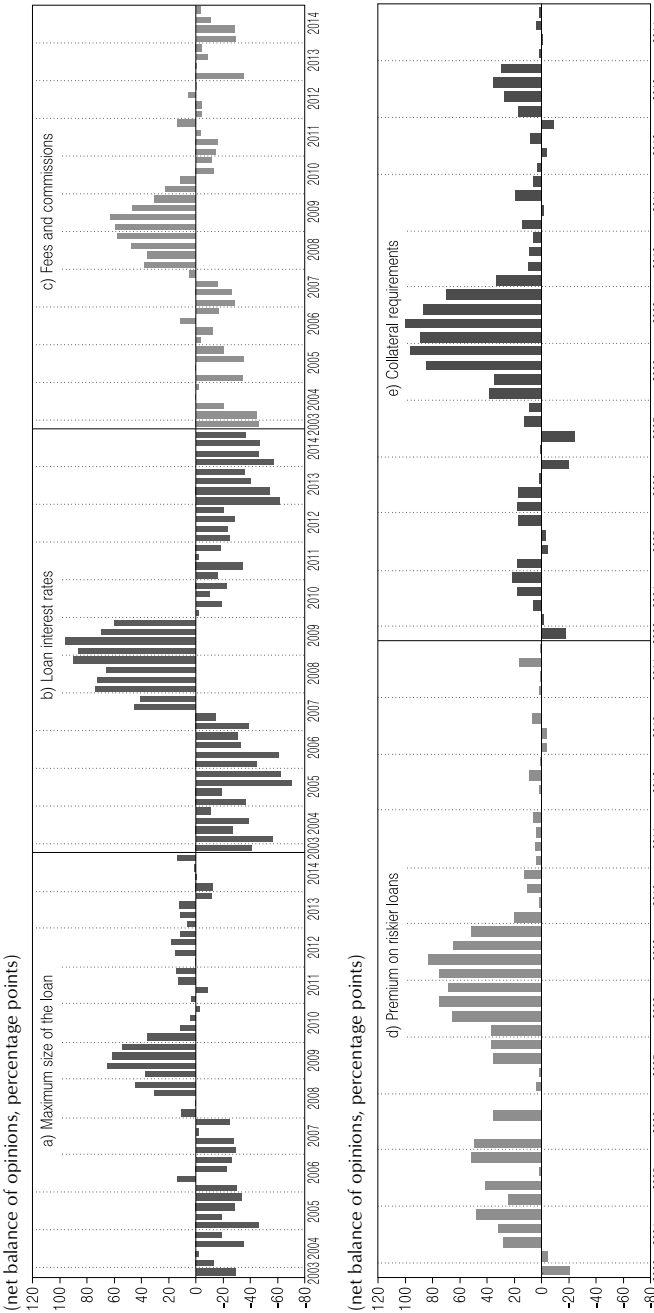
The bank lending survey is conducted in the first month of each quarter (i.e., January, April, July and October). In Bulgaria the survey is addressed to contact persons set up by the individual banks, who answer the questionnaire electronically. All 27 commercial banks operating in Bulgaria at present have taken part in the survey. After all of the participating banks have passed on their answers to the BNB, the central bank undertakes an aggregation of the results on the basis of individual banks weights. These weights are calculated as a ratio of the amounts of loans to enterprises, consumer and housing loans, allowed by each bank to the total amount of the respective loans, allowed by the banking system as a whole. Each quarter the results of the survey are sent back to the participating banks and are also published in the quarterly economic review of the BNB.

[18] „ - -" (contributed considerably to lower demand/tightening of credit standards), "-" (contributed somewhat to lower demand/tightening of credit standards), "0" (contributed to basically unchanged demand/credit standards), "+" (contributed somewhat to higher demand/easing of credit standards), "+ +" (contributed considerably to higher demand/easing of credit standards), NA (not applicable).

The bank lending survey in Bulgaria has been conducted since the fourth quarter of 2003, i.e., the data set covers a period of forty five quarters. However, it should be noted that the data set covers the whole period from 2003 Q4 to 2014 Q4 only in respect to short-term and long-term loans to enterprises. Regarding total lending to firms, housing and consumer loans data is available only from 2010 Q1 till present. When interpreting the survey findings the qualitative nature of the results should be borne in mind. They are not objective, quantitative data such as precise figures on credit volume but reflect tendency estimates recorded on a five-point scale. Furthermore, the survey is concerned only with identifying changes in respect to the previous quarter. As a result, information on levels (such as the degree of restriction imposed by a bank's current lending policy) cannot be automatically derived from the survey data. In order to be able to interpret and analyze the results the net balance of responses in percentage terms is calculated. For questions related to the supply side of lending, this net percentage is the difference between the percentage share of responses in the restrictive range (i.e., reporting a tightening of credit standards) less the percentage share of responses in the expansionary range (i.e., reporting an easing of credit standards). This means that a positive value suggests a restrictive tendency while a negative value indicates an expansionary tendency. Regarding loan demand, the net percentage is the difference between the percentages reporting an increase and a decrease in demand.

Annex II

Chart 12: Conditions and terms for approving loans and enterprises



Note: The balance of opinions is defined as a difference in percentage points between the percentage of banks responding "tightened considerably" and "tightened somewhat" and the percentage of banks responding "eased somewhat" and "eased considerably".

Source: BNB – Bank Lending Survey.

Chart 13: Conditions and terms for loans to huseholds

Chart 13.1: Loans for consumer credit

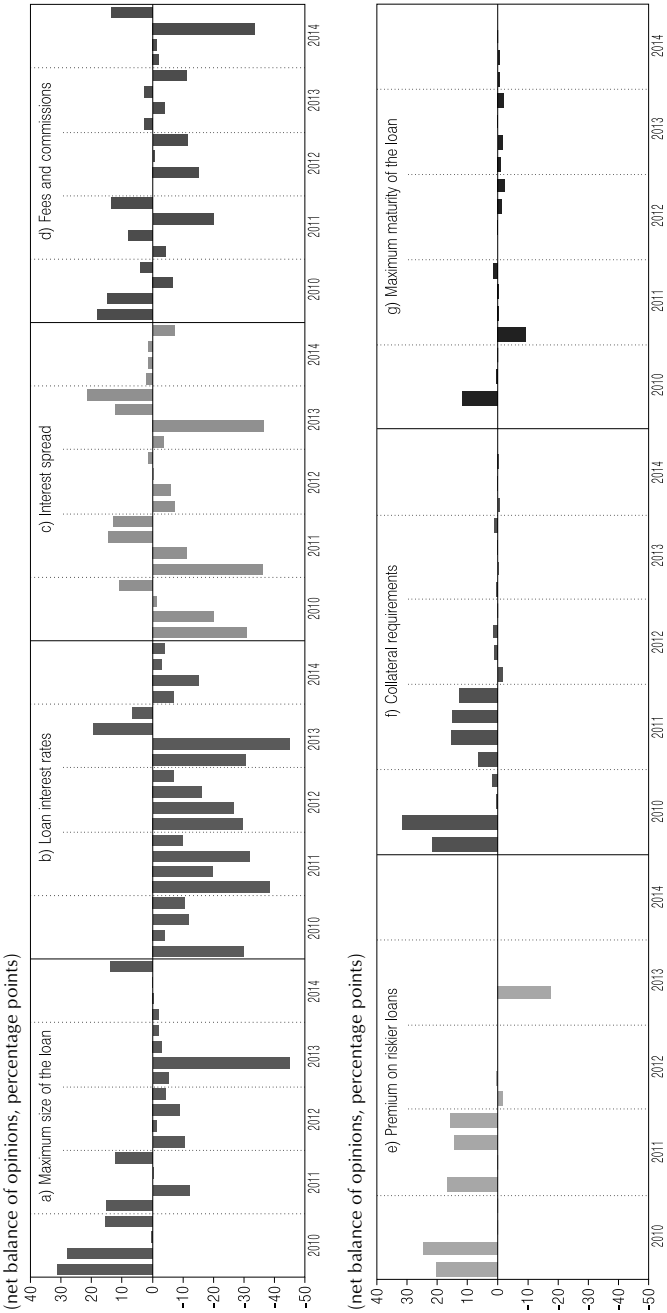
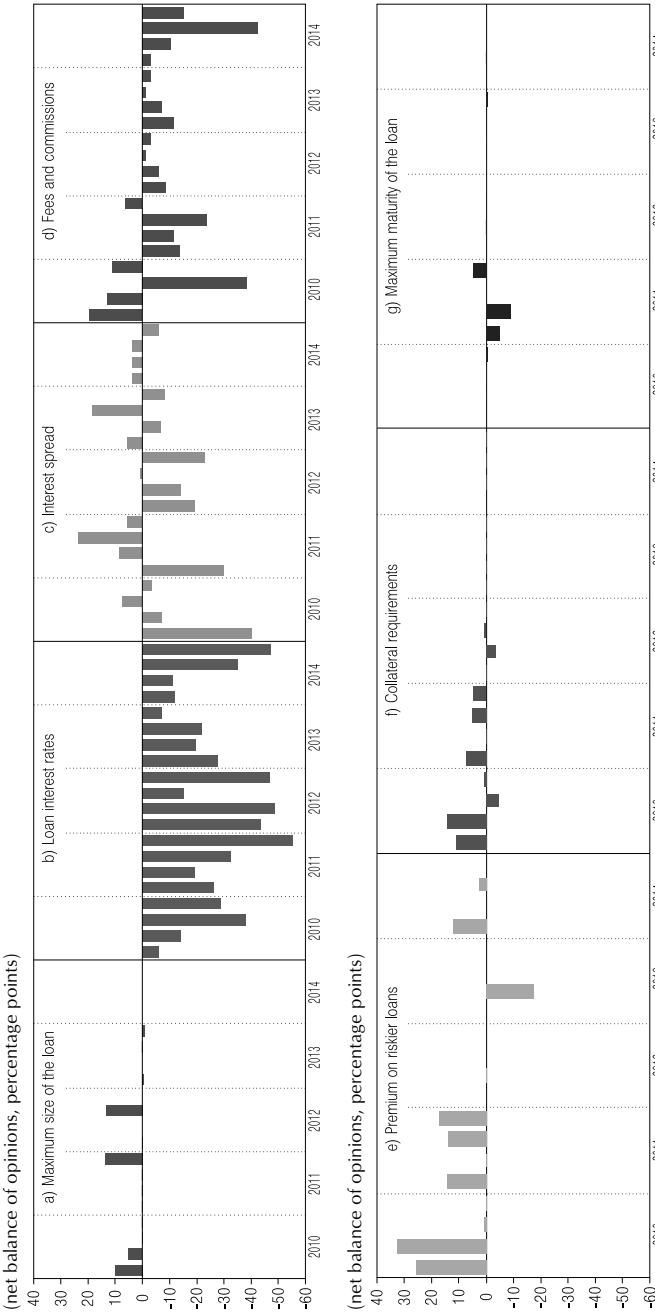


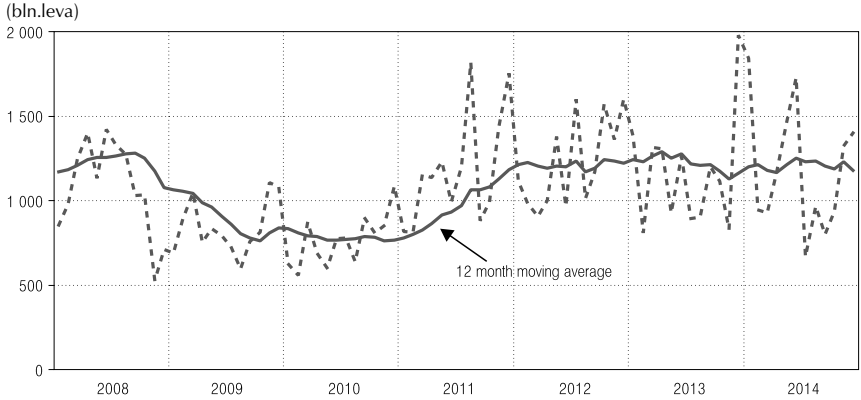
Chart 13.2: Loans for house purchase



Note: The balance of opinions is defined as a difference in percentage points between the percentage of banks responding "tightened considerably" and "tightened somewhat" and the percentage of banks responding "eased somewhat" and "eased considerably".

Source: BNB – Bank Lending Survey.

Chart 14: Volume of new-business loans to non-financial corporations



Source: BNB

Table 4: Cross correlation between growth of claims on non-financial corporations (in period t=0) and BLS indicators at various lags (t-k) and leads (t+k), k=1...4. Macro level.

Δ ln Kt	Δ demand_sh	Δ supply_sh	Δ demand_lg	Δ supply_lg
-4	0.44**	-0.03	0.39**	-0.02
-3	0.49**	0.09	0.48**	0.19
-2	0.30**	-0.08	0.41**	-0.07
-1	0.39**	0.13	0.46**	0.08
0	0.66**	-0.10	0.70**	0.04
1	0.21	0.49**	0.28	0.45**
2	0.34**	-0.03	0.32**	-0.04
3	0.05	0.27	0.27	0.23
4	0.09	-0.02	0.12	0.02

*** indicates significance at the 5 % level*

Table 5: Cross correlations between growth of claims on non-financial corporations (in period t=0) and additional explanatory variables at various lags (t-k) and leads (t+k), k=1...4. Macro-level.

Δ ln Kt	Δ ln real GDP	Δ Interest spread	Δ BRL	Δ Business climate	Δ Capital/ Assets
-4	0.29	-0.08	-0.13	0.15	0.03
-3	0.40**	-0.16	-0.15	0.23	-0.14
-2	0.47**	-0.01	-0.08	0.21	0.06
-1	0.35**	0.05	-0.27	0.19	-0.13
0	0.35**	-0.23	-0.45**	0.03	-0.44**
1	0.33**	-0.16	-0.26	0.14	0.13
2	0.17	-0.13	-0.23	-0.1	0.06
3	0.12	-0.03	-0.18	-0.22	-0.08
4	-0.03	-0.07	-0.15	-0.30	0.34**

*** indicates significance at the 5 % level*

Table 6: Unit-Root Tests

H0: Variable has a Unit Root	ADF Test t-statistics			Phillips-Perron Test t-statistics		
	Level	First difference	First difference of the log	Level	First difference	First difference of the log
Claims on non-financial corporations	-1.82	-2.63*	-4.85***	-1.80	-2.52	-5.24***
Demand for loans (BLS): short-term loans	-4.30***			-4.65***		
Credit standards (BLS):short-term loans	-2.39	-8.60***		-2.42	-8.45***	
Demand for loans (BLS): long-term loans	-3.06**			-3.21**		
Credit standards (BLS): long-term loans	-2.35	-9.05***		-2.38	-8.80***	
BRL	0.27	-2.96**		0.20	-2.62*	
Real GDP_sa	-2.47	-2.92**	-2.93**	-2.64*	-2.96**	-2.94**
Business climate	-1.68	-4.74***		-1.49	-4.76***	
Capital to asset ratio	-1.06	-5.60***		-1.32	-5.62***	
Interest rate spread	-2.50	-9.36***		-2.58	-9.48***	

**** indicates significance at the 1 % level; ** at the 5% level;* at the 10% level*

Table 7: Cross correlations between growth of loans to corporations (in period t=0) and additional explanatory variables at various lags (t-k) and leads (t+k), k=1...4. Micro-level.

Δ ln Kt	Δ ln of real GDP	Interest spread by bank	Δ Capital/ Assets	Δ BRL	Δ Business climate
-4	0.10 **	0.00	0.05**	-0.06**	0.07**
-3	0.13 **	0.01	0.09**	-0.02	0.16**
-2	0.12 **	0.09 **	0.00	-0.04	0.06**
-1	0.05 **	0.08 **	-0.07**	-0.11**	-0.06**
0	0.17 **	0.07 **	0.08**	-0.07**	0.13**
1	0.06 **	0.05	0.01	-0.04	0.07**
2	0.07 **	0.10 **	-0.15**	-0.09**	-0.1**
3	-0.03	0.13 **	0.05**	-0.05	-0.17**
4	0.02	0.07 **	0.04	-0.19**	-0.19**

*** indicates significance at the 5 % level*

Table 8: Panel Unit-Root Tests

H0: Variable has a Unit Root	Im, Pesaran and Shin W-statistics		
	Level	First difference	First difference of the log
Loans to corporations	0.65	-15.69***	-18.58***
Real GDP_sa	-1.14	-2.73***	-2.25***
Interest rate spread by individual banks	-2.62***		
Capital to asset ratio	-3.96***		
BRL	11.57	-3.69***	
Business climate	-0.90	-17.46***	

**** indicates significance at the 1 % level; ** at the 5% level;* at the 10% level*

DP/103/2017

Analysis of the Housing Market Developments and the Underlying Macroeconomic Fundamentals in Bulgaria

Petra Kotseva, Mihail Yanchev

Abstract. The current paper aims to identify the main factors driving housing prices in Bulgaria in the last decade and formulate a model that would help determine whether housing prices were in line with the fundamental macroeconomic factors of supply and demand and if not measure and evaluate the magnitude and persistence of the misalignments. Our econometric methodology is a vector error correction model that exploits a cointegration relationship between the variables and is designed to deal with the dynamic fluctuations of housing prices around their long-run trend. Our results show that the dynamics in the housing market in the last decade were caused by objective factors driving the demand and consecutively the supply of housing. On the other hand the accumulation of a housing market bubble driven by subjective factors has started just few quarters prior the crises. The correction of house prices during the global crisis was also sufficient and at present property prices are in accordance with income, interest rates and foreign demand dynamics.

Резюме. Настоящото изследване има за цел да установи основните фактори, променящи цените на жилищата в България през последното десетилетие и да формулира модел, който би могъл да определи дали цените на жилищата съответстват на основните макроикономически фактори на търсенето и предлагането, а ако ли не, да измери и оцени амплитудата и устойчивостта на съществуващото несъответствие. Иконометричната ни методология представлява векторен модел с корекция на грешките, който използва коинтеграционните зависимости между променливите и е конструиран да се справи с динамичните флуктуации в цените на жилищата около техния дългосрочен тренд. Резултатите ни показват, че динамиката на жилищния пазар през последното десетилетие е производна от обективни фактори, променящи търсенето, а след това и предлагането на жилища. От друга страна, набъбването на „балона“ на пазара на жилища – продукт на субективни фактори, започва точно няколко тримесечия преди кризата. Приспособяването на жилищния пазар по времето на глобалната криза също беше достатъчно и понастоящем цените на имотите съответстват на доходите, лихвените проценти и динамиката на външното търсене.

Petra Kotseva, Economic Research and Forecasting Directorate,
Bulgarian National Bank, rogleva.p@bnbank.org

Mihail Yanchev, Economic Research and Forecasting Directorate,
Bulgarian National Bank (worked in the period 2012–2015),
yanchev.mihail@gmail.com

Acknowledgment:

We thank our colleagues from the Economic Research and Forecasting Directorate for the support and helpful discussions and especially Daniel Kasabov for his overall contribution to this paper.

Contents

1. Introduction .....

2. Concepts of Housing Market Bubbles and Approaches to Estimation of Overvaluation.....

3. Construction and Housing Market in Bulgaria .....
   3.1. Demand Side Factors.....
   3.2. Supply Side Factors .....
   3.3. Preliminary Conclusions.....

4. Data and Methodology for Estimation of Overvaluation.....

5. Results .....

6. Conclusions.....

References .....

Appendix.....
   A1. Reconstructed Data on Interest Rates and Loans .....
   A2. Population.....
   A3. Hypothetical Borrowing Value .....
   A4. The Model.....

1. Introduction

Like most countries across Europe and all new European Union member states (NMS), in the period before the 2008 – 2009 global crisis Bulgaria experienced a strong growth in construction activity, real estate related activity and housing and land prices. Angello and Schuknecht (2009), Min Zhu (2014), among many, agree that the integration of the global financial markets were the main driver of housing market developments across the world and that these unprecedented, temporary spikes in housing prices observed in many advanced and developing economies (coupled with the strong mortgage credit growth) became one of the fundamental reasons behind the global crisis. The sharp increase of construction activity and house prices prior to the 2008 – 2009 global crisis and the significant correction in the following years demonstrated the unsustainability of the observed dynamics in the housing market in the boom years. Indications for slight revival of the sector at the end of 2013 and the beginning of 2014 also suggest that the bottom has already been reached and recovery in construction and housing market activity can be expected. It is imperative to understand whether the housing market was in line with fundamentals in the period before the 2008 – 2009 global crisis and determine where the current price level stands compared to its long-run path, in order to formulate appropriate expectations for the future developments. Accurate assessment of the current and future developments, on the basis of which relevant and timely policy measures can be determined is of great importance, since housing market imbalances can have a significant impact on the overall macroeconomic stability.

Various factors affected housing market dynamics in the CEE countries in the period before the 2008 – 2009 global crisis. Egert and Mihaljek (2007) study the determinants of the house prices dynamics in CEE in the period from 1998 to 2005. They argue that along with the fundamental factors such as GDP per capita, real interest rates, housing credit and demographic factors important drivers in the region were also some (as they call them) transition-specific factors. Such factors, particularly relevant for Bulgaria were weak housing market institutions (housing market was fully regulated up until 1989), almost nonexistent housing finance, high level of home-ownership, limited supply of new homes, improvements in housing quality (higher growth due to composition effect – the share of higher-quality and higher-priced housing increased) and house price misalignment (initial undershooting).

Both the sluggish transition of the economy and the severe domestic financial crisis in 1997 in Bulgaria had a strong negative impact on the construction sector and the housing market. As a consequence of the prevailing macroeconomic instability and lacking restructuring of the economy up to 1997 both the banking sector remained predominantly state-owned and underdeveloped and the interest rates were persistently high – legacy that contributed to the prolonged period of suppressed construction activity even in the years of growth acceleration. These country specific developments made housing market in Bulgaria hard to fit in an international comparison and difficult to be investigated by methods that are already traditional for most advanced economies. Therefore, it is necessary to base the analysis of the housing market development in a historical context.

This paper aims to analyze the main driving factors of the housing market and determine whether housing prices were in line with the macroeconomic fundamentals. A detailed review of the literature is used to determine a framework of concepts and definitions of a misalignment and a bubble. Furthermore, using established methods in the literature and applying them to the context of the Bulgarian housing market, an econometric model is constructed to measure the misalignments of the observed housing prices from the trend/level that would be suggested by the fundamental macroeconomic factors. Understanding of the macroeconomic developments related to the housing market, which took place prior and after the 2008-2009 global crisis is essential to identify the main factors that determined housing price dynamics. Once these are pinpointed one has to evaluate the available macroeconomic indicators that would most effectively account for these fundamental factors. Traditional indicators of supply and demand for housing often suggested in the literature like GDP, household income, interest rates on mortgage lending, FDI inflows, construction activity and others are analyzed. An additional demand-side indicator (hypothetical borrowing value) is constructed for Bulgaria using a method outlined by Addison-Smyth, McQuinn and O'Reilly (2009), which is valuable from an analytical standpoint apart from its use in the context of an econometric model in this current paper. The evaluation of indicators of consumer sentiments and expectations also provide evidence whether there are periods of accumulation of a housing bubble or not, since rapid price increases alone do not provide enough evidence for the formation of a price bubble. In the end, the selection of the most relevant and informative indicators is crucial for the construction of a robust economic model, as only a relatively short sample of data is available.

In the next section we summarize the main concepts in housing market literature and some of the approaches recently used for estimating market overvaluation. Then we review some of the developments in the economy which we identify as having a major impact on the housing market starting with short overview of the transition period and its legacy. The estimated model, together with a short discussion on some of the possible alternative specifications is presented in section 4. In section 5 the results of the model are interpreted in the context of the observations previously made in the paper. The last section concludes.

2. Concepts of Housing Market Bubbles and Approaches to Estimation of Overvaluation

The literature on the housing market business cycle has reached an agreement over several facts. First, income is the main determinant of housing demand. It is either the case that richer agents can afford to spend more on housing or that higher income would relax credit constraints, which also raises the housing affordability for the particular agent (Poterba (1984), Case and Schiller (1990), Poterba (1991), Englund and Ioannides (1997)). Case and Schiller and later Englund and Ioannides use different econometric technique, but reach a similar finding that lagged real GDP growth and real interest rates have a high predictive power for real housing prices.

Second, housing supply is inelastic and adjusts to demand over long periods of time (Glaeser and Gyourko (2007), Ortalo-Magné and Prat (2007), Glaeser, Gyourko and Saiz (2008)). It follows that the business cycle of the housing market spans over longer periods of time. In fact Kearl (1979), Poterba (1984), Topel and Rosen (1988) develop a framework where the housing market is actually presented as consisting of two separate markets: one for the existing housing stock and one for the flow of new construction, which determines the level of new investment. Shocks to either of these markets can affect house prices.

Third, agents can invest in housing for speculative purposes and not only because they demand a dwelling to inhabit. That means that the desire to buy housing is strongly influenced by the expectation of reselling at higher prices in the future (Case and Shiller (1988), Case and Shiller (2004)). Moreover, such speculation is based on extrapolative expectations which are formed on the basis of past housing price movements, rather than any knowledge of fundamentals (Case and Shiller (1988)). Thus, housing price booms tend to persist as home buyers become “destabilizing speculators.”

Lastly, individuals hold heterogeneous expectations about housing prices and assessment of the situation on the housing market, meaning that people hold various beliefs about the future developments of the market. It is not possible to short-sell housing, which can be a reason for why prices are affected only by agents with positive expectations. Elaborating on the last point, it is the people with inflated expectations that make the decisions to buy when the slump is imminent that eventually go into foreclosure or are forced to sell at a loss (Case and Shiller (2004)).

With respect to the definition of an asset price bubble it is precisely that if investors base their decisions to purchase an asset on the belief that they will be compensated by a rapid further price increase in the future and not on some knowledge of fundamentals, then there is a bubble (Stiglitz (1990)). More specifically, if home-buyers decide that a home they would normally consider as too expensive is an acceptable purchase, because they will be compensated by further housing price increases, then there is a housing bubble (Case and Schiller (2004)). On the other hand home-buyers expecting a rapid price increase might decide to buy now rather than later when they will not be able to afford it. Therefore, if expectations of rapid price increases are important motivating factors for home-buyers, then "housing prices are inherently instable" (Case and Schiller (2004)).

In respect to measuring misalignments of housing prices from the long-run level suggested by the macroeconomic fundamentals, there is a broadly accepted methodology. According to Detken and Smets (2004) and Angello and Schuknecht (2009), housing prices can be compared to their long-run trends in order to define the persistence and magnitude of periods of housing overvaluation or undervaluation. Defining boom and bust phases can be done by identifying the periods when housing prices are respectively above or below their long-run trend. Detken and Smets define a "boom cycle" as a period during which asset prices are 10% above their trend. That threshold, however, seems quite arbitrarily chosen and seems to fit the case of the US housing market. According to Angello and Schuknecht (2009) it is not only the magnitude of an overpricing that matters in analyzing real or financial asset prices. If housing prices are overpriced for a longer period of time, consumers and investors may adjust their expectations upwards, which would have an effect on domestic demand and demand for credit as well. Therefore, persistence is as important as the actual magnitude, since people's expectations can fuel the continuation of a boom period and in effect produce a deeper bust.

The methodology of Angello and Schuknecht (2009) is also followed by the European Commission (2012) in their "housing imbalance toolkit."

Additionally, in their article on "Assessing the dynamics of house prices in the euro area" (2012) it is pointed out that house price cycle analysis can be supplemented with affordability and dividends ratios such as respectively price-to-income and price-to-rent ratios. Housing price booms should be analyzed in the context of affordability pressures for the homebuyers. An increase in households' real disposable income can potentially accommodate rising house prices due to demand pressures and a slower reaction of the supply. On the other hand, prolonged and rapid increases in the price-to-income ratio or even deviations from its long-term average could be interpreted as a sign of overvaluation. Housing price dynamics can also be assessed against the cost of renting. Looking at housing as any other asset, house price changes are expected to be driven by changes in expected capital gains or in income from future housing services (income from renting). In equilibrium, agents should be indifferent between buying/selling and renting. Thus, the dynamics in the price-to-rental ratio could be interpreted as a sign of overheating or cooling, respectively when the ratio is higher or it is lower. In "normal times" the higher the ratio goes, the more potential buyers would chose to rent rather than buy, which should exert downward pressure on prices and vice versa.

However, as pointed out by the European Comission (2012), comparing house prices or other derived indicators to their long-run averages or assuming that their fundamental value is their historical average can bias the analysis, since it makes the simplifying assumption that the series are stationary. Therefore it is more appropriate to use other methods to derive equilibrium values of housing prices that are based on fundamental factors. In a VAR/VECM framework one can simulate the dynamic interplay between housing prices and other factors on the demand and supply side.

Gattini and Hiebert (2010) use a VECM to model for purpose of forecasting and analyzing euro area house prices and their interplay with the macroeconomic indicators. To arrive at an equilibrium level of Euro Area housing prices they use both demand and supply side factors as their explanatory variables: housing investment, real disposable income per capita and a mixed maturity measure of the real interest rate. Their results suggest that housing demand and financing cost shocks appear to have contributed strongly to the dynamics of euro area house prices.

In another paper Addison-Smyth, McQuinn and O'Reilly (2009) apply a model of mortgage credit and examine the relationship between Irish house prices and changes in lending patterns. In their model they use fundamental factors as mortgage credit, income, interest rates and housing stock. Their results suggest that post 2003 a significant amount of the

increase in Irish housing prices was determined by developments in international finance, which enabled Irish institutions, in particular, to secure alternative sources of lending funds. A useful contribution of their paper is an indicator they construct (hypothetical borrowing volume), which represents in a more realistic way the role of the interest rates on the agent's behavior and in particular on the amount an individual can borrow. They formulate the indicator as the present value of an annuity, where the annuity is a given fraction of the current disposable income discounted at the current mortgage interest rate for a horizon equal to the term of the mortgage.

3. Construction and Housing Market in Bulgaria

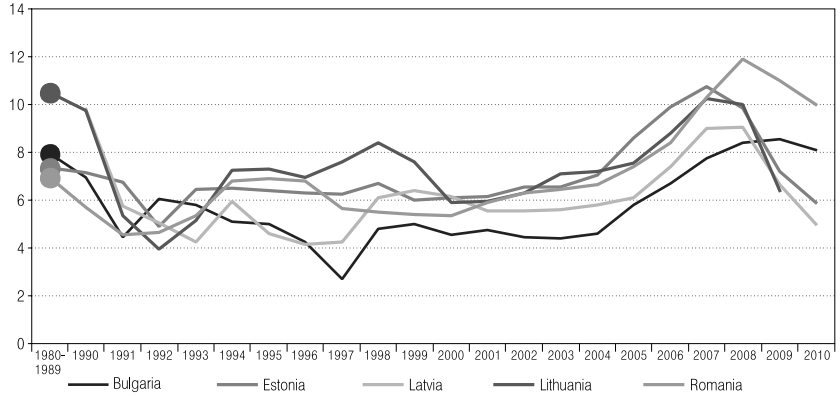
The housing stock and respectively the supply of housing evolve slowly over time given the long life of housing and the small additions to the housing stock and thus the current housing market situation is significantly affected by the economic developments decades ago. Due to the turbulent period of the transition from a command to market based economy in Bulgaria the construction sector and the housing market underwent a prolonged period of depressed supply.

Construction activity in Bulgaria remained subdued until 2004 compared to the overall economic growth in Bulgaria, which started to pick up after the two consecutive domestic crises in 1991 and 1996-1997. The share of value added in the sector to the total value added in the economy is an indicative measure of the supply of construction in the economy. According to the United Nations Statistical Division data, the share of the construction sector in Bulgaria dropped from average 8.3% in 1970-1989 to 4.5% in 1991 and then further to 2.7% in 1997 (Figure 1). By the end of the 1980s, a higher share of construction in the total value added was common for most centrally planned economies in Europe, and in the beginning of the transition period there was a slump in the construction activity. However for most NMS the construction activity recovered to around 6-7% of total value added quickly after the beginning of the transition and remained at these levels until the end of the 1990s. While in most countries the share of construction in total value added approaches 6% on average in the long run, Bulgaria remains in the first quartile in the distribution of the 28 EU member states according to this indicator together with Germany and Belgium at stable levels about 5% from 2000 to 2004[1] (Figure 2). As opposed to Germany and Belgium, in Bulgaria the process of restructuring of the economy, the underdeveloped infrastructure and the obsolete hous-

[1] Missing data for significant part of EU before 2000.

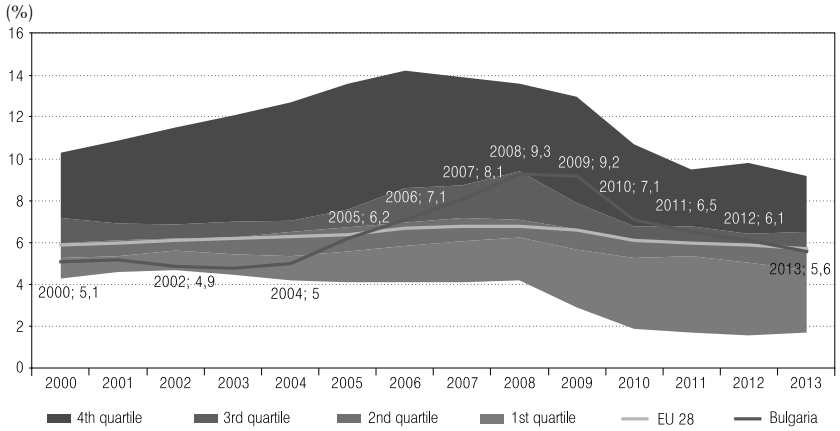
ing stock suggested outpacing growth rates in construction as it was the case in most NMS shortly after the beginning of their transition period. This prolonged period of depressed supply is one of the significant differences between the construction and housing market developments in Bulgaria and the other CEE countries.

Figure 1: **Over a Longer Horizon – Share of Construction in Total Value Added in Selected CES Countries**



Source: United Nations Statistical Division

Figure 2: **Share of Construction in Total Value Added**



Source: Eurostat

The high rate of owner occupied dwellings in Bulgaria was also a legacy of the socialist regime (see Table A4 in the Appendix for Distribution of population by tenure status). In 1990 Bulgaria was among the few countries in Eastern Europe, where the housing was predominantly private but the state dominated all related activities. A major part of the dwellings were state-built or cooperative and especially in urban areas[2]. The public sector was the largest developer of new properties and private construction was represented typically by dwellings in the rural areas constructed by their owners.

The efforts at the beginning of the transition to transform the housing sector from state domination into a free market included legislation changes. The housing market deregulation of 1990 eliminated the involvement of the state or municipal authorities in the market operations. A change in the Property Law allowed each citizen in Bulgaria to possess unlimited amount of real estate. Another inevitable step in the process of transformation was the halt in public investment in dwellings which was driven by shortage of financing which imposed significant government budget expenditure cuts.

The economic crisis, the lack of housing finance and the withdrawal of the state from the development of housing led up to a collapse of residential construction which was common for Eastern Europe (Palacin and Shelburne 2005). In Bulgaria the number of newly built dwellings declined continually from around 65 000 in 1985 and 26 000 in 1990 to less than 7 000 in 1995 (NSI Statistical Yearbook 1985, 1990, 1995).

The development of a functioning real estate market in Bulgaria began well after the deregulation of the economy and much later than in the other CEE countries. Factors for the delay were the low income of households and the lack of a developed banking sector, which was associated with high cost of financing in both relative and absolute terms. Despite the highly worn-out housing stock and the deficient infrastructure in the beginning of the 2000s it wasn't until 2004 that the share of construction started to rise from its level of about 5% of total value added. By that time the outstanding amounts of loans for house purchases had been growing with rates over 100% from initially very low levels while the interest rates were following a steady downward trend at the background of sustained macroeconomic stability.

[2] According to the NSI Statistical Yearbook 1985, 56% of the newly build homes in 1975 were state-build, 26% - co-operative housing and 18% - privately build. In 1985 these shares are respectively 46%, 26% and 28%.

Figure 3: **House Prices**

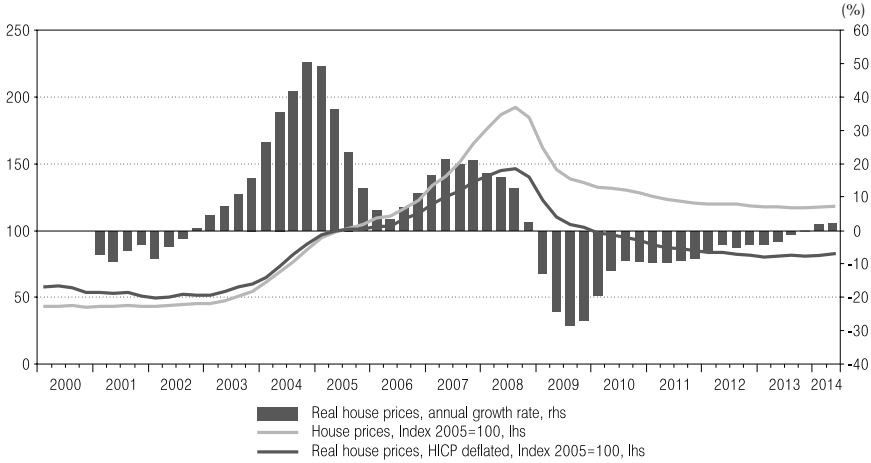
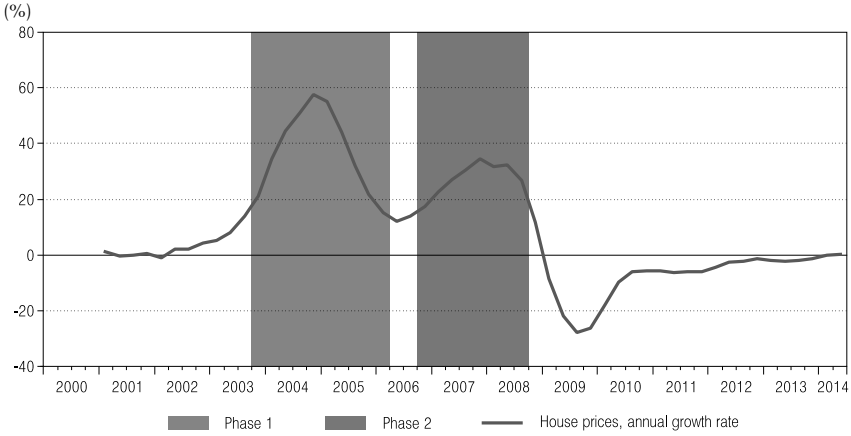


Figure 4: **House Prices, Annual Growth Rate**



Thus, although somewhat later, Bulgaria joined the general trend in CEE countries of revival of the property market, with a sharp increase in construction activity and rapidly growing housing prices in the period before the 2008 – 2009 global crisis. Only in two years (2004-2005) house prices doubled and in 2007-2008 grew by another 50%.

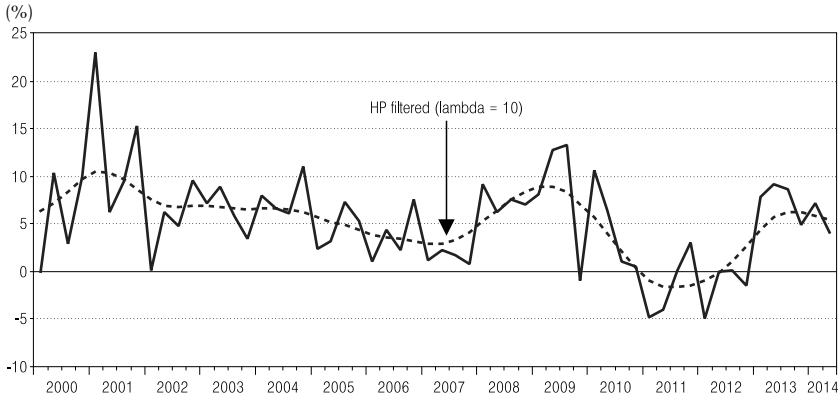
In what follows we aim at identifying the possible demand and supply factors driving house prices. Additionally, we try to summarize the main changes which have occurred in 2009 due to the global crisis. A remarkable deceleration in prices in 2006 also attracts attention. For this reason

we separate the housing boom in the period before the 2008 – 2009 global crisis into two phases covering periods with annual growth rates exceeding an arbitrarily chosen threshold of 15% (see Figure 4). The first phase spans between 2003 and end 2005 when prices start growing from very low levels and reach annual growth of 50% and a second phase of growth between the end of 2006 and the end of 2008 with a slightly weaker magnitude.

3.1. Demand Side Factors

The growth in disposable income and the fast development of the banking sector, which was associated with an improved access to credit and the decline in interest rates, increased the demand for housing. Disposable income of households was growing at a high rate and often at double digit rates in the period before the 2008 – 2009 global crisis. The main contribution to the growth of disposable income came from the growth in compensations of employees. Government transfers to the households and the current transfers to the private sector are the other components of disposable income which have a significant contribution to its growth in the period before the 2008 – 2009 global crisis.

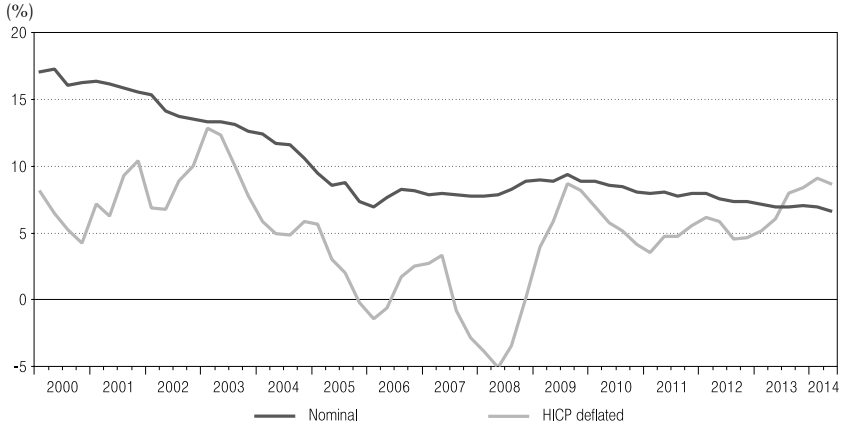
Figure 5: **Households Real Disposable Income, Annual Growth Rate (Private Consumption Deflated)**



Mortgage interest rates, were well above 15% in the beginning of 2001 and fell down to below 7% on average in 2006. They were generally following the dynamics of international interest rates as of during the period before the 2008 – 2009 global crisis, and were strongly influenced by the increased access to foreign financing and the expansion of the financial sector. Together with the high growth of all claims on non-financial cor-

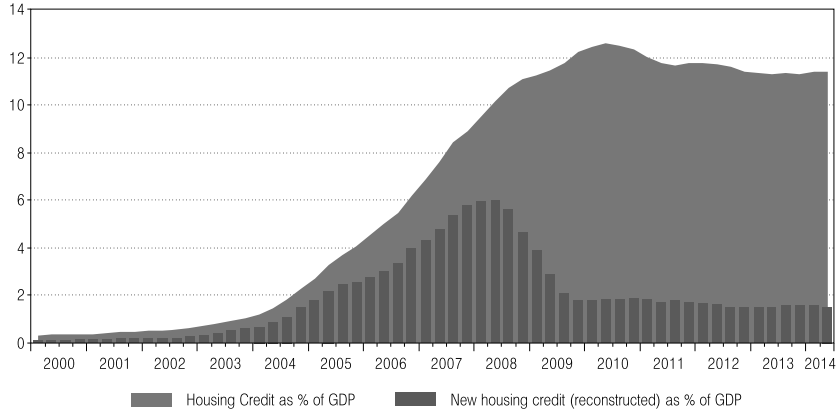
porations and households, mortgage credit volumes also started to rise rapidly in 2003 and 2004. In response, the BNB implemented a number of counter-cyclical measures[3] to moderate and stabilize the rate of growth of private sector credit to levels sustainable for the medium run and not threatening the financial stability. From the beginning of 2005, implicit credit ceilings were introduced for banks where the growth rate of credit exceeded a certain limit. The banks which violated the limit had to maintain additional minimum required reserves. Thus, lending was made more costly for the banks where credit growth was exceeding the limit. Moreover, from the beginning of 2006 (but in effect from May 2006) BNB introduced a progressive scale for determining additional required reserves. Special attention was paid to the rapid expansion of mortgage credit and particular measures in this respect were seen as necessary. Such a measure was a change in the risk weights of mortgage loans, which were to be used in calculating the bank's capital adequacy ratio. In the second half of 2006 the BNB gradually lifted all these administrative measures. Although the exact effect of these measures is hard to be determined, a stabilization of mortgage credit growth in the end of 2005 and the beginning of 2006 is visible, which corresponds with BNB's objectives. Thus, these measures could have affected the demand for housing through retention of the amount of the credit supply.

Figure 6: **Interest Rate on New Housing Loans (Reconstructed)**



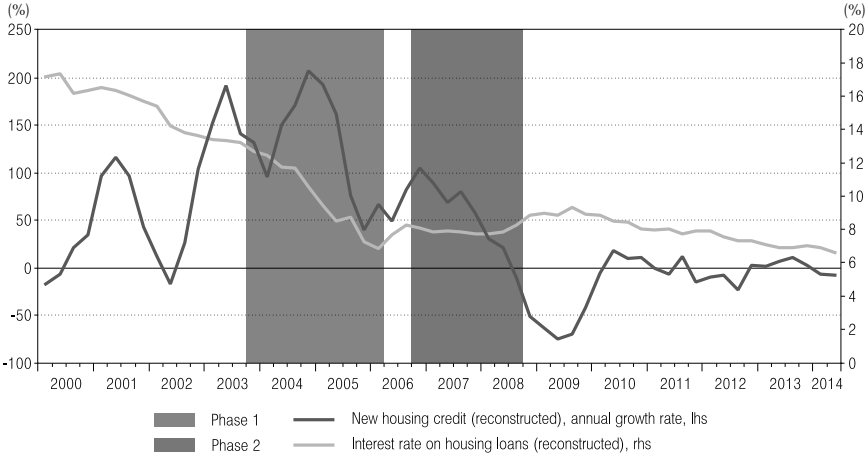
[3] For detailed information on the counter-cyclical measures of the BNB see BNB Annual Reports 2004 – 2006.

Figure 7: **Housing Credit Volume as a Share of GDP**



Note: *Information about the reconstructed series is provided in the Appendix.*

Figure 8: **Interest Rates and Growth in New Credit for Housing**



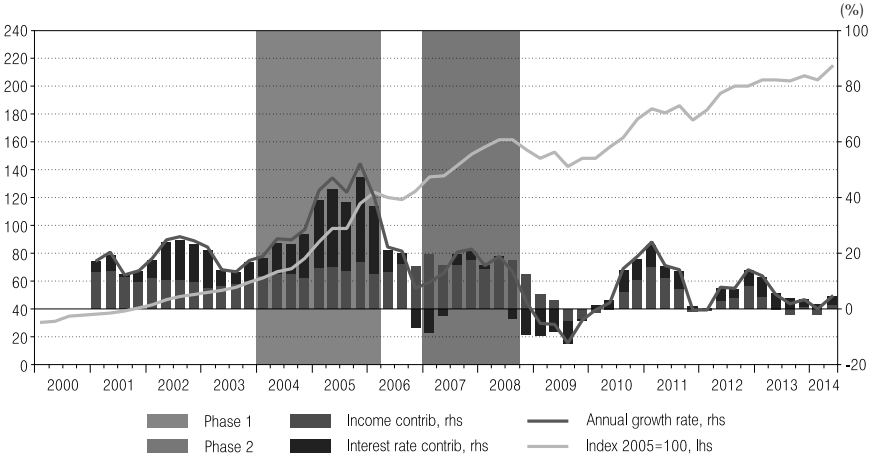
The access to financing and in particular the supply of credit and the interest rates are important determining factors for housing demand. However, the effect of interest rates comes through its influence upon the amount that prospective house buyer can borrow from financial institution and thus the effects on the demand for housing can be far from linear. The typical amount of a mortgage offered by a bank is usually based on the present value of an annuity, where the annuity is a given fraction of the current disposable income, which is used for mortgage repayments and is discounted

at the current mortgage interest rate for a horizon equal to the term of the mortgage. Following the example of Addison-Smyth, McQuinn and O'Reily (2009) we construct a hypothetical borrowing volume indicator B_t :

$$B_t = kY_t \frac{1-(1+R_t)^{-\tau}}{R_t}$$

where Y_t is the nominal GDP as an approximation of the current disposable income of households, k is the fraction of that income that goes towards repayments of the mortgage, R_t is the current mortgage interest rate and τ is the duration of the mortgage.[4] Constructed in this way this indicator should better resemble the considerations of consumers who want to purchase a home. Additionally we can decompose the growth of the hypothetical borrowing volume indicator into growth of income and change in interest rates. The overall indicator increased rapidly in the period before the 2008 – 2009 global crisis with steady contribution on the side of the income and more pronounced acceleration during the first phase, driven by the decline in mortgage interest rates. The growth in hypothetical borrowing volume notably decelerated starting from the end of the first phase and came to a halt in the beginning of the second when interest rates discontinued their decline.

Figure 9: **Hypothetical Borrowing Volume**

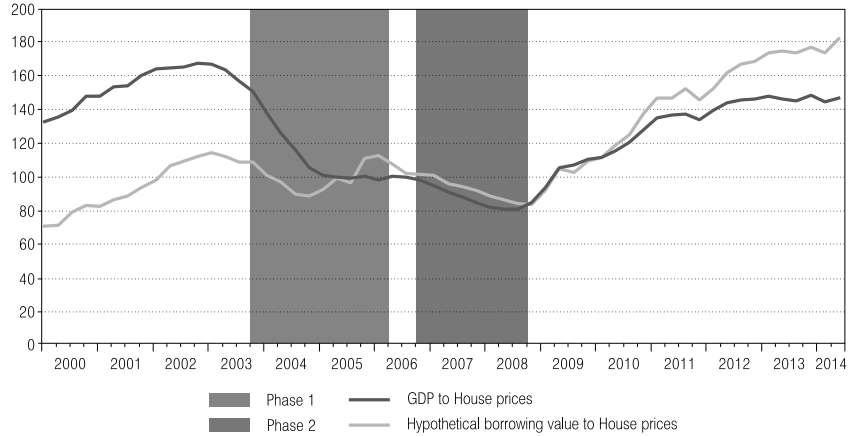


[4] More information about the construction of the indicator is provided in the Appendix.

The increase in income and in the amount that can possibly be borrowed made real estate more affordable for households prior to the first phase of house prices acceleration as the ratios of GDP to house prices and hypothetical borrowing volume to house prices indicates (see Figure 10). Although incomes increased and interest rates decreased during the housing price boom, house prices growth outpaced the growth in GDP and hypothetical borrowing volume and affordability deteriorated steeply during the first phase of the housing price boom and remained low throughout the second phase as well.

Figure 10: **Affordability**

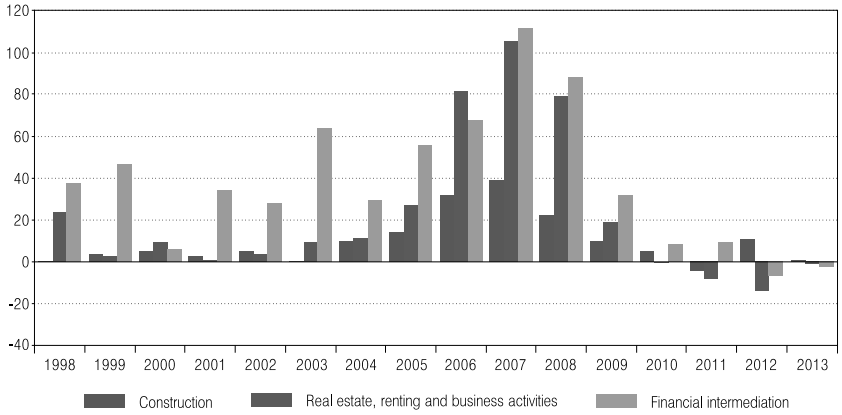
(index 2005=100)



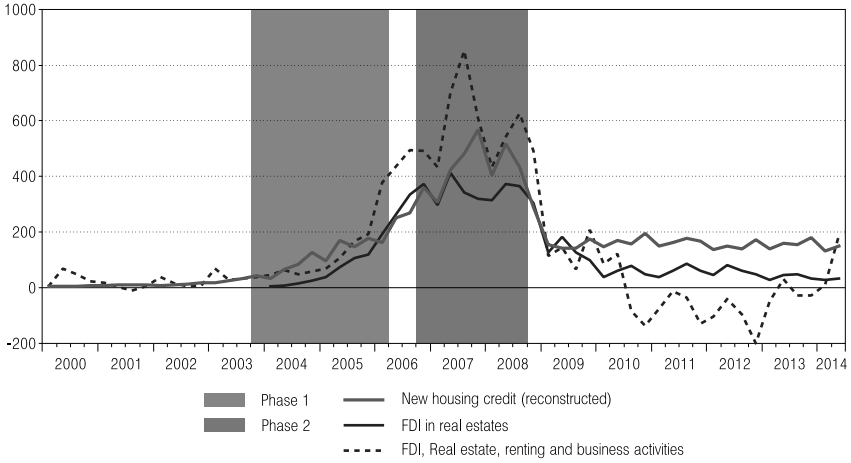
An important factor affecting all aspects of the economic development in the period before the 2008 – 2009 global crisis was the high foreign capital inflow into the country. The housing market in particular was affected through various channels. The supply of new housing was boosted through foreign direct investment in construction and real estate activities and in the banking sector, increasing the access for credit for non-financial corporations. The FDI in the banking sector had an effect on the demand side through the mortgage lending channel but probably the major effect of FDI on the housing market was the indirect one through the overall increase in income at that time. However, purchases of real estates by foreigners had direct effect on house prices. As a preferred tourist destination for many Europeans and with its relatively low housing prices, Bulgaria attracted large FDI inflows in real estates since 2004. The FDI inflows increased as percent of GDP after the initial rise in housing prices and the beginning of

the construction boom, when the quality of the housing stock was quickly improving.

Figure 11: **FDI Inflows by Sectors as % Value Added in the Sector**



FDI and New Housing Credit, EUR mln.



Consumer sentiment in the period before the 2008 – 2009 global crisis seemed to be in line with the fundamentals at least during the first phase. Their expectations for the financial situation were broadly consistent with the dynamics of the disposable income. Intent of households to purchase or build a home in the following year was upward sloping during that period which corresponds to a gradually increasing demand for dwellings.

According to this indicator, the peak in demand was in the middle of 2007, which roughly coincides with the peak in the growth of housing prices. On the other hand, consumer expectations for their own financial situation and the overall economic situation in the country were on a downward trend since the middle of 2005 with the exception of an outlying peak in the middle of 2007 and did not deteriorate much further as the global crisis struck. Therefore, probably consumer sentiment was a factor during the second phase (2007 – 2008) of housing price growth and speculative demand also had a role in the housing boom as individuals bought housing not as a consumption good, but rather an investment good as the prices were steadily climbing.

Figure 12: **Consumer Expectations Balance of Opinions**

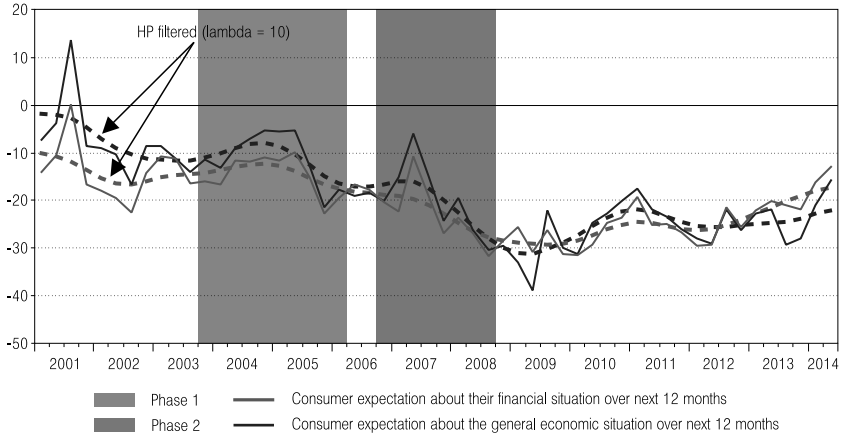
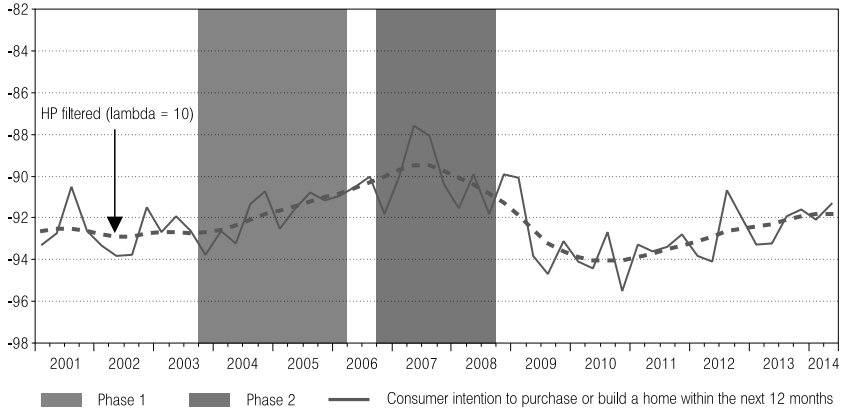


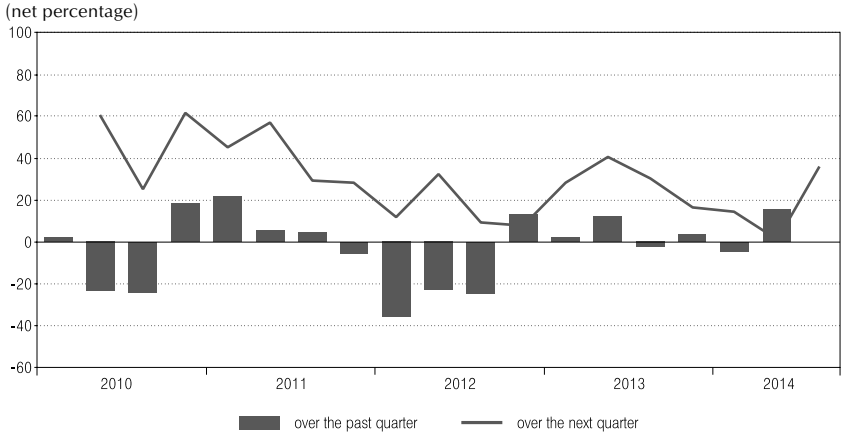
Figure 13: **Consumer Sentiments Balance of Opinions**



Since the beginning of the 2008-2009 global crisis house prices corrected with 32% from their peak in Q3 2008 only in 2 years and continued to decline till the end of 2013. The response of disposable income was milder and more gradual as compensations continued to grow during the global crisis, but the reduction in employment caused consumer confidence to drop. Interest rates increased temporarily in 2009 and 2010, but the effect was much more pronounced in real terms as inflation slowed down significantly. As a whole the effect on house prices was much stronger than on income and interest rates and affordability increased to its levels before the boom in house prices. Thus the prevailing uncertainty in regards to future income and employment prospects and the expectations for future house price reductions had a major effect on demand. Households shifted their preferences towards mortgage loan repayments, as is evident from the broadly stable total amount of outstanding loans for house purchase and the constant low level of new housing loans that is observed since the middle of 2009.

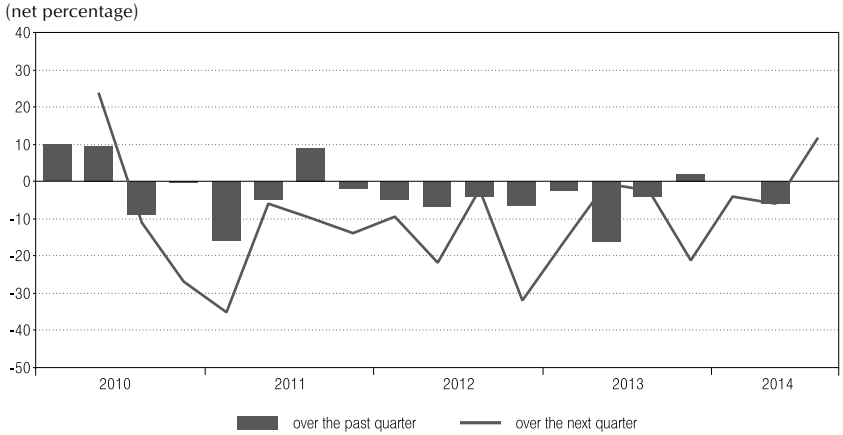
Data from the Bank Lending Survey, conducted each quarter by the BNB, point that the demand for housing credit is weak, even though interest rates have been coming down and overall credit standards for housing loans have been loosening since the middle of 2010.

Figure 14: **Change in Demand for Housing Loans**



Source: BNB, Bank Lending Survey

Change in Credit Standards as Applied to the Approval of Housing Loans



Source: BNB, Bank Lending Survey

Note: *Net percentage balances are calculated by weighting together the responses of those banks who answered the question. The bars show the responses over the previous quarter. The lines show the expectations over the next quarter. Expectations balances have been moved forward one quarter to be compared with the actual outturns in the following quarter. A positive balance indicates that more secured credit is available.*

However, Registry Agency data show a slight upward trend in the number of registered sales of properties since 2010, which is a sign of weak, but strengthening demand, not necessarily financed through mortgage loans.

Figure 15: **Number of Registered Sales of Properties, Seasonally Adjusted Data**

Source: *Bulgarian Registry Agency*

3.2. Supply Side Factors

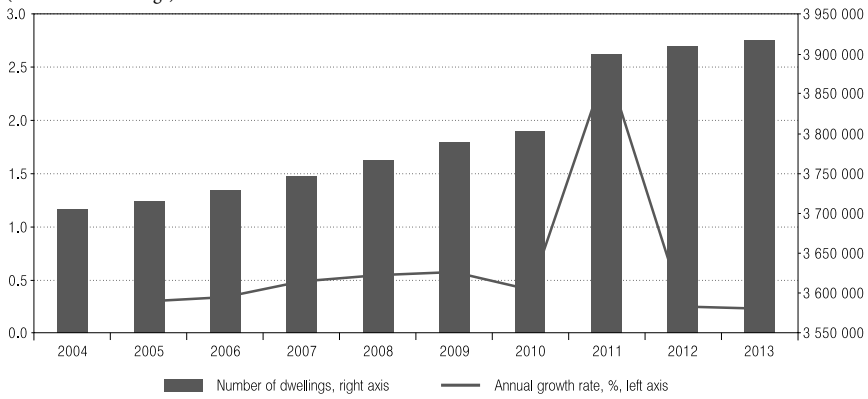
With respect to the supply of housing, limited data availability forced us to gather information from a broad range of indicators. Due to a short sample for the housing stock we analyze short term statistics on construction production, national accounts data on value added in construction and data on residential investment that provide some broad measures of supply and show how it reacted to demand. Firm side indicators like profits and costs shed more light on the misalignment between supply and demand. On the other hand, information from business climate surveys in construction provides insight about how expectations and sentiments of firms' managers as well as competition between firms in the sector evolved throughout the analyzed period.

There are two separate estimations of the housing stock from the 2001 and the 2011 census. However, they were declared unreliable by the National Statistical Institute itself in an ex-post evaluation of the last census. The data for the housing stock available on an annual frequency since 2004, shows that both the number and useful floor space of dwellings increased gradually until 2010. An abrupt jump in the data follows in 2011 when data was most probably corrected with additional information from the national census. The overall increase of housing stock in 2010 compared with 2004 is 2.7% in numbers of dwellings and 3.3% in useful floor space. Newly built dwellings data show gradual decrease from 2010 on in number of dwellings and from 2009 on in useful floor space, which corresponds

with the decrease in newly issued permits[5] for residential buildings starting from 2008.

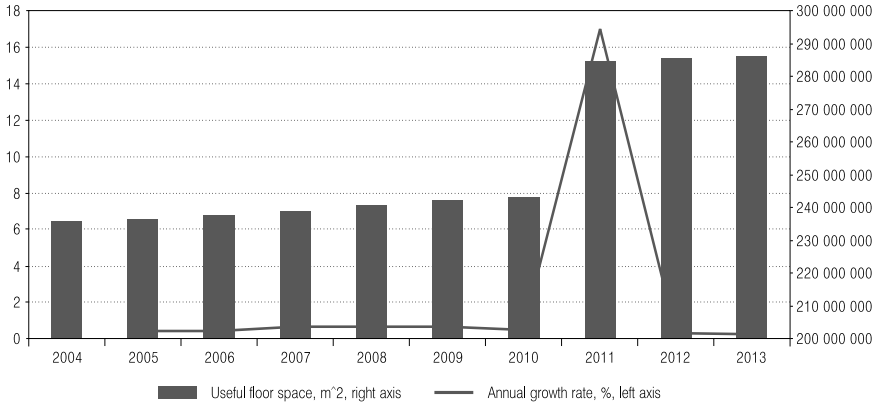
Figure 16: **Housing Stock**

(number of dwellings)



Housing Stock

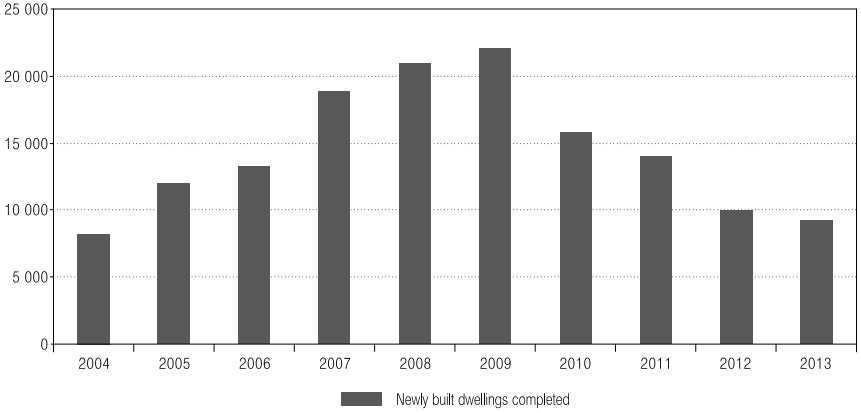
(useful floor space)



[5] A building permit is an authorization to start work on a building project, and as such is the final stage of authorization prior to the start of work. The objective of building permits indices is to assess the development of construction activity. Source: NSI

Newly Built Dwellings Completed

(number of dwellings)



Newly Built Dwellings Completed

(useful floor space)

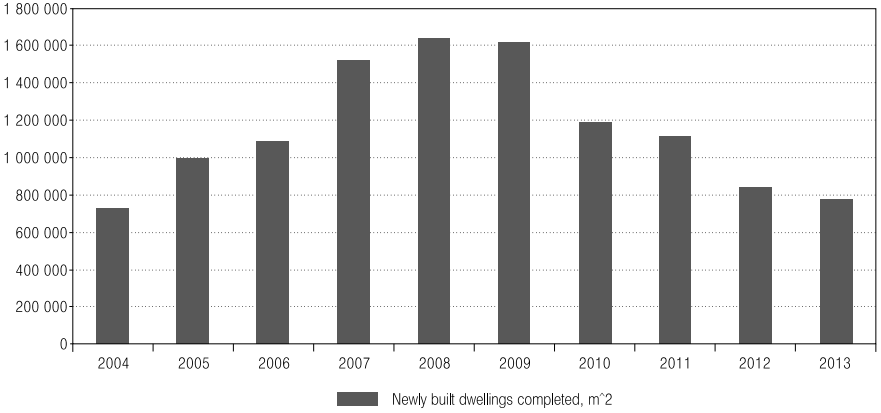
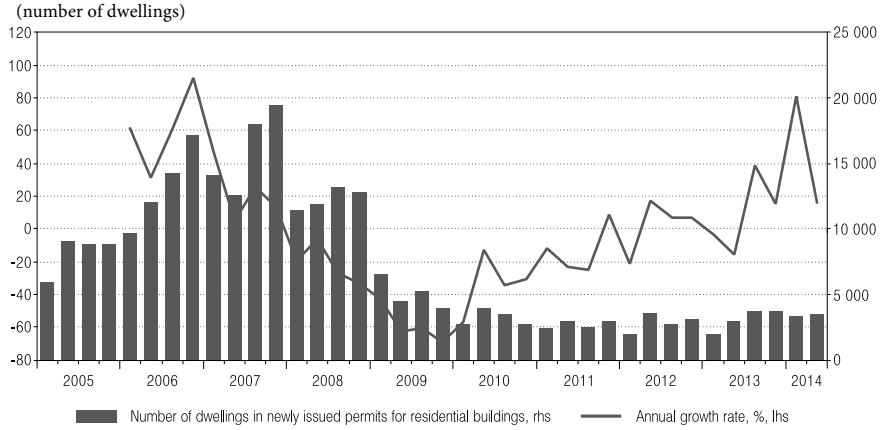
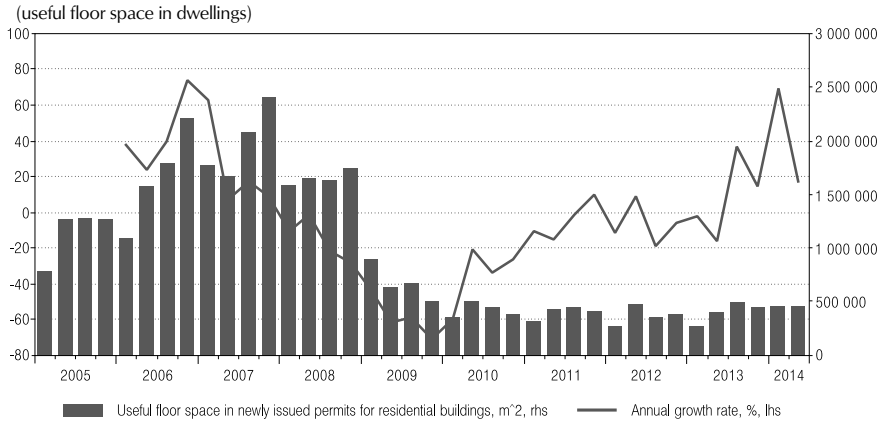


Figure 17: **Newly Issued Permits for Residential Buildings**



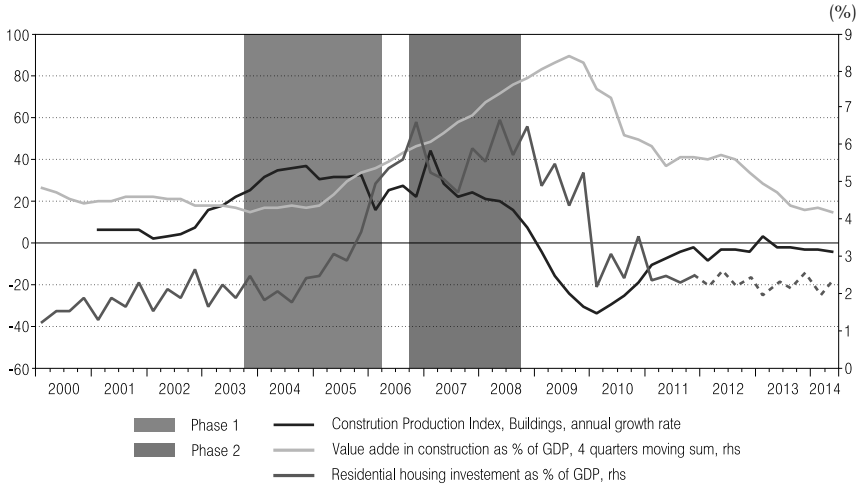
Newly Issued Permits for Residential Buildings



Another indirect indicator for the supply of housing is the data for cosntruction activity from the short term business statistics and the national accounts (Figure 18). This construction production index covers both construction of residential buildings and civil engineering. According to this indicator, supply of new buildings maintans constant growth rates in the period before the 2008 – 2009 global crisis. However, National Accounts data (which also takes into account civil engineering) shows acceleration in supply from the end of the first phase on. We considered also a more imperfect measure of housing supply, by constructing quarterly series for residential investment from the final use method of the national accounts

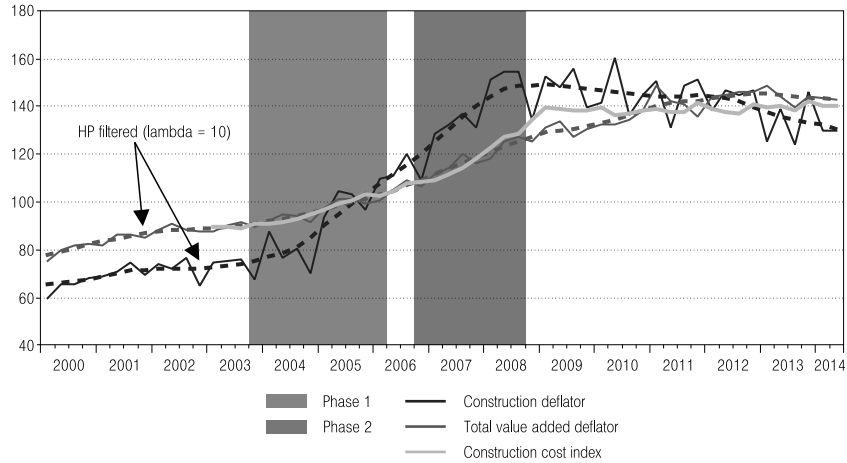
through interpolation of annual data. In general, data confirms that supply of housing reacts with a lag to changes in demand and price dynamics are leading for the supply. Nevertheless, in between the two phases a feedback effect possibly had a downward pressure on prices. On the other hand, it is worth noting that the construction production index started declining about a year earlier, around Q1 2009, compared to the other two supply-side indicators, which points out that the short-term statistics are leading the national accounts data.

Figure 18: **Construction Production and Housing Investment**



As supply was not able to react as quickly on the rise in demand, profitability in the construction sector increased and the value added deflator for construction was outpacing both the value added deflator for the whole economy and the construction cost index. The latter encompasses both expenditures on labour and input materials and in the period before the 2008 – 2009 global crisis commoves together with the total economy deflator.

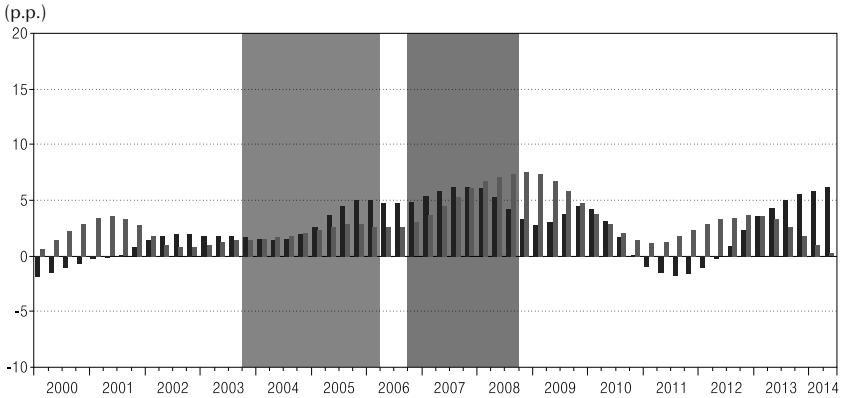
Figure 19: **Value Added Deflator and Construction Cost,**
(index 2005=100)



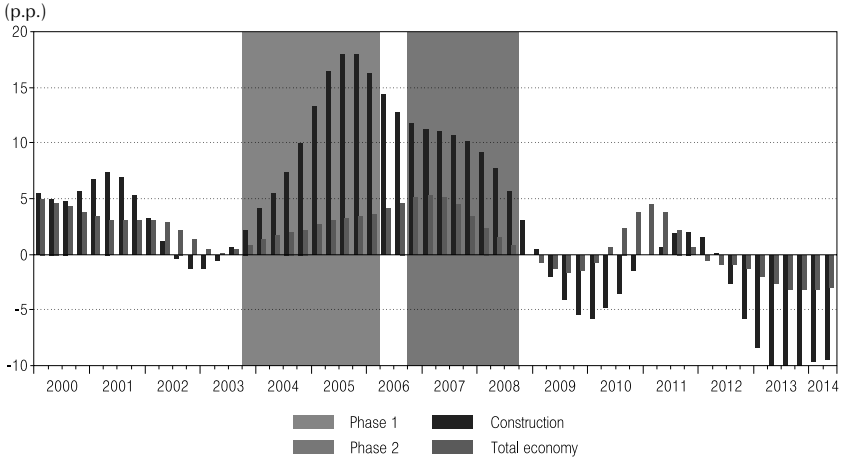
Decomposing the annual growth rate of the value added deflator in construction into unit labor costs, unit profit and unit (net) tax reveals that the misalignment between supply and demand was quite significant during the first phase of the housing price boom. It is evident that the strong growth of the value added deflator in construction was due to the high contribution of the unit profit component rather than the unit labor cost. The latter was more aligned to the one observed for the whole economy, while the former points out to an excess demand. This misalignment is much more pronounced in the first phase of the housing boom, but persists until the beginning of 2009.

Figure 20: **Contribution to Value Added Deflator Annual Growth HP Filtered** (Lambda=10)

Unit Labour Cost

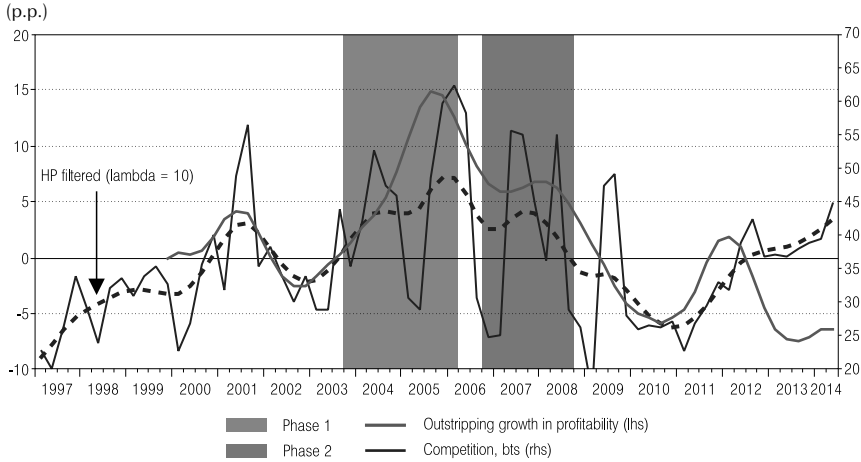


Unit Profit



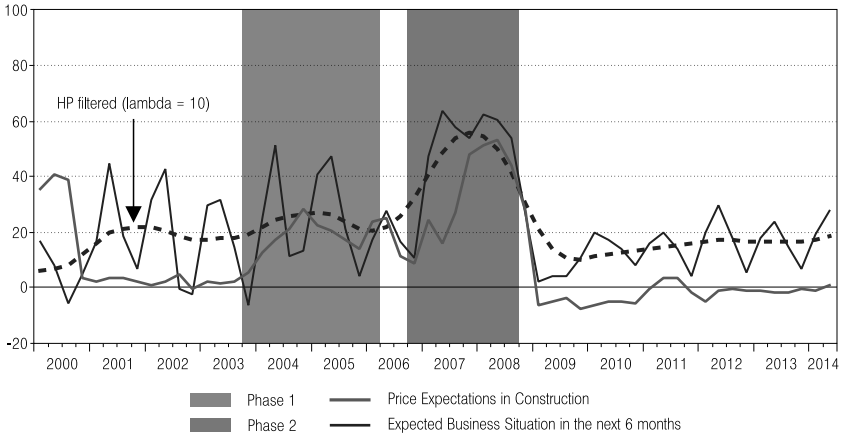
Moreover, as profit margins in the sector increased relative to the ones for the whole economy, managers of surveyed construction firms pointed that higher competition is a significant factor limiting the business activity. According to this indicator, competition in the sector increases when growth rate of profitability in the sector accelerates above the average for the economy levels.

Figure 21: **Profitability and Competition in Construction**



Lastly, it is important to point out that during the second phase of the housing boom, up until the beginning of the global crisis, the business activity surveys for the firms in construction show overly optimistic expectations about the future business activity and housing prices. This leads to the conclusion that expectations of economic agents both on the demand and supply side were an important factor for the second phase of the housing boom and much less so for the first phase. Therefore, it is plausible to speak of a housing bubble, since the decisions of firms were based on unrealistic expectations that the housing boom will continue indefinitely, despite the fact that in 2008 profit margins were already quickly shrinking and mortgage interest rates were starting to go up. Some of these expectations might have been based on the assumption that the global crisis would not affect Bulgaria as much as other EU member states, but they were proven wrong as decline in external demand and a deceleration in capital inflows quickly followed the collapse of Lehman Brothers in September 2008 and domestic economic activity declined as well.

Figure 22: **Expectations of Firms in Construction**



The business climate expectations play an important role in the formation of housing market trends. Since the outbreak of the global crisis, there is a pronounced downward shift in the overall confidence indicator of the construction sector with a weak trend of improvement since the second The business climate expectations play an important role in the formation of housing market trends. Since the outbreak of the global crisis, there is a pronounced downward shift in the overall confidence indicator of the construction sector with a weak trend of improvement since the second quarter of 2012. Price expectations over the next 3 months have remained stable in 2013 pointing to an expected slight decrease in prices. Building activity in construction sector was assessed by managers as continuing to decline, though at a decelerating rate.

Since the beginning of 2011, the number of issued building permits has stabilized on a low level, after a period of decline between the end of 2007 and the end of 2010. Newly completed residential buildings experienced a slight recovery in the course of 2011, but continued falling on an annual basis since Q1 2012. The construction production index generally follows with a lag the building permits dynamics and has been declining on an annual basis since end-2008 until present, with the exception of a temporary stabilization in 2012. These developments point out to a stagnant supply.

Profits in construction also corrected downward after the beginning of 2009. The decline in the value added deflator for the construction sector during the 2008 – 2009 global crisis was higher than for the total economy.

Firms in the sector maintained lower increase in labor costs through larger reduction in employment and decreased significantly the gross operating surplus. The number of firms in the sector had decreased by 20% from 2009 to 2013 according to preliminary data from the structural business statistics. In this period value added declined by 31% in current prices and 22% in constant prices. Reduction in employment started even a year earlier and from 2008 to 2013 40% of employed were dismissed.

3.3. Preliminary Conclusions

Several conclusions can be drawn on the basis of the gathered data. During the first phase the increase in income and reduction in interest rates stimulated demand for housing; the high demand in turn, combined with the insufficient supply (before the first phase) allowed firms to increase significantly their profit margins. Most likely, the slowdown in property price growth in 2006 cannot be attributed to factors on the side of domestic or external demand, but rather to the dynamics of credit supply, mortgage interest rates and housing supply. During the second phase of the boom external demand for real estate (FDI) also had a significant effect on prices. An improvement in expectations was observed during the second phase, pointing to possible subjective factors pushing house prices above their fundamental values and causing a bubble on the real estate market.

With the outbreak of the global crisis house prices corrected rapidly under the declining demand, driven by households' sentiments and the halt in external demand. On the supply side, the continuation of on-going projects postponed the decline in production for several quarters, but firms started to limit labour costs and observe a decline in profit margins in 2008 with the deterioration in confidence in the sector.

As of 2014 the weak real economic growth, the high level of unemployment and the uncertainty in regards to future income can be viewed as some of the major demand-side factors for the continuing decline in housing prices since the onset of the global crisis. Furthermore, there is a prevailing preference for mortgage loan repayment by households, as is evident from the broadly stable total amount of outstanding loans for house purchases and the constant low level of new housing loans that is observed since the middle of 2009, even though interest rates have been coming down and overall credit standards for housing loans have been loosening since the middle of 2010.

4. Data and Methodology for Estimation of Overvaluation

The preferred approach for estimation of house price misalignment is to find a stable long-run relationship between house prices and some fundamental driving factors, around which house prices fluctuate in the short run. We choose to use a vector error correction model, designed to deal with the dynamic fluctuations of housing prices around their long-run trend because of the non-stationarity of the variables of interest and the existence of a cointegration relationship between housing prices and the domestic demand indicator - hypothetical borrowing value. We arrive at a stable long-run relationship after testing numerous specifications, which reveals the size and persistence of the misalignments of observed housing prices from their fundamental level. It also provides enough information to make a guess about the expected future developments on the housing market.

Based on the investigation of the historical development of the Bulgarian housing market we arrived at three main factors affecting the house prices, for which we have statistical indicators. These factors are domestic demand for housing, foreign demand for real estate and the supply of housing.

On the demand side we have taken two measures for income: gross domestic product and disposable income. Disposable income is calculated as the sum of the wage bill of employed (wages of self-employed persons are approximated by wages of employees), imputed rent of owner occupied dwellings, government transfers to the households, net current transfers of the private sectors, compensations of employees paid by non-residents and net of personal income tax and social security contributions paid by the employees. The credit market indicators that we used were the reconstructed nominal interest rate on housing loans and the reconstructed new business housing loans (see the appendix) and the quarterly change of the outstanding volume of housing loans as a measure of the credit activity of banks, alternative to the one reconstructed by us. Theoretically, population is one of the main factors, driving fundamental house prices and is also one of the considered series. In the case of Bulgaria, population decline should have had negative effect on property prices, but it is unclear whether such an effect could be identified in such a short and volatile sample such as ours.

Additionally, as a demand side indicator the hypothetical borrowing value constructed following the example of Addison-Smyth, McQuinn and O'Reily (2009) was used. The aim of including this indicator was twofold.

First, to account better for the non-linear effects of interest rates on the amount that an individual can afford to pay on a house. This indicator should capture in a better way the decision making considerations of consumers, compared to the indicators used for its construction taken separately. The second purpose was to arrive at a more parsimonious specification of the model combining the income and interest rate effect in one variable which has been proven to be crucial in our case after using more than three dependent variables in our ECM/VECM specifications did not yield robust results.

Our external demand proxy is foreign direct investment in real estate taken from the balance of payments statistics. Foreign direct investment in real estate data is available since 2004 when foreigners were allowed to purchase real estates in Bulgaria. We extrapolated the series for the period 2000-2003 on the basis of the data for FDI in real estate, renting and business activities, which were very low at that period.

On the supply side our main indicator was the construction production index for construction of buildings. The share of value added in construction in the total value added for the economy was also used in several specifications, even though we understand that it encompasses not only the construction of residential buildings, but civil engineering as well. We also constructed a measure for residential investment (although this is more of a demand side indicator) by taking the available data on annual shares of residential investment to gross fixed capital formation and applying them to the quarterly series for gross fixed capital formation. Thus, we interpolated annual data to get quarterly series using the gross fixed capital formation quarterly profile. From the business climate survey in construction we also used the expectations of managers of firms in construction for the future economic activity and selling prices in the next 12 months.

Models for house prices were tested both in nominal and in real terms. To construct real terms indicators the following deflation procedure was applied. House prices were deflated by HICP index, GDP by GDP deflator, mortgage interest rates by the annual rate of HICP inflation, FDI in real estates and new housing loans by the house price index. Additionally, income was considered both in absolute value and per capita. All variables were seasonally adjusted and taken in natural logarithms. Most series start in Q1 2000 and end Q2 2014, which provided us with 58 observations. As we aim at finding cointegration relationship, all variables were first tested

for unit roots (see the appendix). Phillips-Perron test suggests that all indicators are first order integrated.[6]

By testing different specifications we have arrived at an error correction model of house prices (HP), hypothetical borrowing value (HBV) and foreign direct investment in real estate (FDI). Factors on the supply side rarely entered significantly or with the expected negative sign in any of the specifications tested. Granger causality tests between housing prices and supply-side factors confirmed that the supply is always lagging behind the movement of prices. Therefore, this in a way confirms that if the supply was a driving factor for the housing market, it was only so for a short part of the sample between the two identified phases of housing price growth.

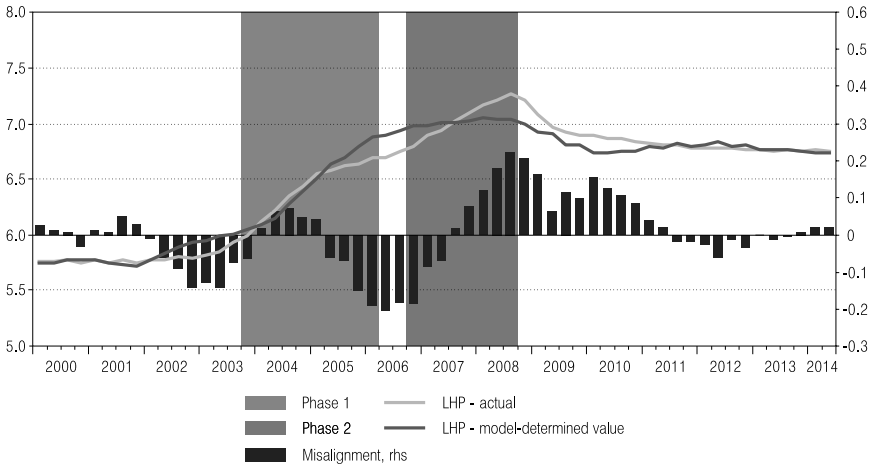
Johansen method for cointegration testing suggested one cointegration relationship between the variables with deterministic trend in the data, which is significant only for HBV. Tests for weak exogeneity further show weak exogeneity of HBV and FDI, which we also find plausible, given that both were driven by external factors. Income, interest rates and FDI were determined by the overall economic development in the country and, especially prior to the global crisis, by the international economic conjuncture. Estimated cointegration relationship and further details for the model are given in the appendix.

5. Results

The developed VECM model estimates house price levels which are in line with the factors included in the model – income, mortgage interest rates and foreign demand for real estate. These levels are usually referred to as fundamental or long-term house prices, but with the following caveats. They cannot be called equilibrium prices as effects on the side of supply could be nontrivial and part of their dynamics is also cyclical. However, they are informative and useful as they show (keeping in mind the possible limitations of the model) to what extent the dynamics in house prices are caused by the factors included (income, interest rates and FDI).

[6] Three types of unit root tests were used in this analysis: the Augmented Dickey-Fuller Test, the Kwiatkowski-Phillips-Schmidt-Shin Test and the Phillips-Perron test. All three types of tests gave comparable results, which led to the decision to report only the Phillips-Perron results in full.

Figure 23: **Estimated House Prices Misalignment**



The estimated model suggests that during the first phase of rapid growth property prices were broadly in line with the increase in income and the reduction of interest rates. Between the two phases of housing price appreciation that we identified there was a housing price undervaluation, which is not explained by the fundamental factors in the model, i.e. household income and mortgage interest rates. Most likely it was driven by the increased investment in construction which was prompted by the outpacing growth in profitability in the construction sector during the first phase. The constraining measures for credit growth, which limited the growth of mortgage credit are most probably another factor for the undervaluation in 2006.

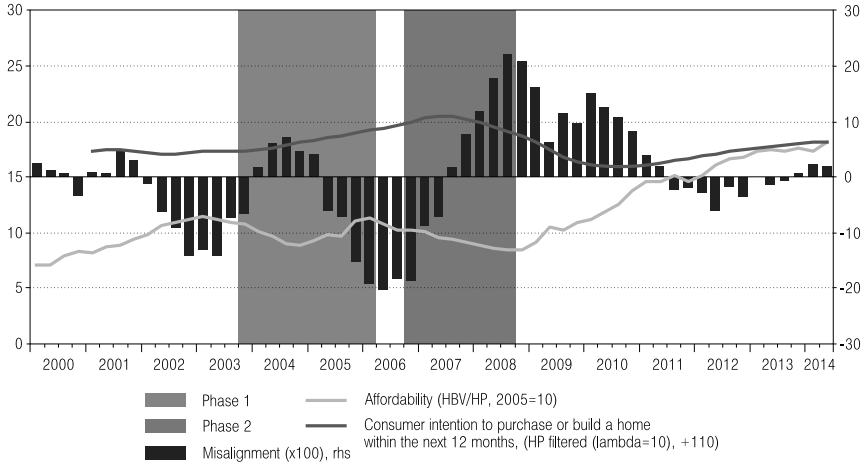
During the second phase the withdrawal of the administrative measure limiting the credit growth and the large FDI inflows allowed once again acceleration in house prices. Around the end of the second phase there was a significant overvaluation of about 20 percent, which can be attributed to the inflated expectations of economic agents for housing prices and future income. Therefore, there is evidence for a housing price bubble during the second phase of housing price growth, since the overvaluation in housing prices during that period cannot be explained by fundamentals alone.

Moreover, the model suggests that during the global crisis, the fundamental house prices declined and in the first quarter of 2010 they were around 30 per cent below their peak value from Q1 2008. The reduction in the

model-determined house prices was driven by the slowdown in the growth of income and, subsequently, its decline and the increase in interest rates in combination with a rapid decline in foreign demand for real estate. Agents quickly corrected their expectations and actual house prices corrected as well. As of 2014 housing prices were very close to their fundamental value, which means that there was no significant over- or undervaluation.

Interestingly, the juxtaposition of the estimated house prices misalignment and the indicator for consumer intention to purchase or build a home shows that indeed households' intentions were growing in the period when house prices were undervalued from the middle of the first phase of house prices acceleration until the middle of the second phase (Figure 24). Consumers' intentions started to decline in the middle of 2007, well before the outbreak of the global crisis, suggesting that consumers started to assess housing as overvalued. The indicator has started to improve since the beginning of 2011 as house prices came closer to and even beneath their fundamental values, probably partly driven by the continued improvement in housing affordability.

Figure 24: **House Price Misalignment, Consumer Sentiments and Affordability**



6. Conclusions

Like most countries in Europe and all NMS, Bulgaria experienced construction boom prior the global crisis and a sharp correction in house prices and construction activity after that. Given the limitations of data available for the housing market in Bulgaria, a reasonable explanation of house price dynamics can be achieved by the model proposed. It exploits the cointegration relationship between house prices, hypothetical borrowing volume, which is constructed on the basis of income and mortgage interest rate, and FDI in real estate, which is our indicator for foreign demand for housing. We argue that after a prolonged period of suppressed housing market the first phase of housing market boom, from the end of 2003 to the end of 2005, was driven by higher income and decreasing interest rates. The limited supply of new housing created conditions for demand side pressures on housing prices, since the response of supply occurs with a lag. On the other hand, the significant increase in the profitability in the construction sector stimulated construction activity in the following years. This, together with the measures aiming at constraining the fast growth in mortgage credit helped to cool off the housing market temporarily in 2006 before a second phase of house price acceleration. According to our model, during this period housing was undervalued for the observed levels of income, interest rates and foreign demand. Households also continued to improve their assessment for the appropriateness of house purchases. The second phase continued until the end of 2008 and was caused by the high foreign demand and overly optimistic expectations of economic agents for the future housing prices. Thus, there is evidence of a housing bubble during the second phase of housing price growth, since the fundamentals alone cannot explain the housing price dynamics. By the end of 2008 house prices were overvalued by about 20% and started correcting downwards rapidly with the deterioration of business sentiments. Although the affordability of housing improved after the 2008 – 2009 global crisis compared with the period of housing boom, high unemployment and the uncertainty about future income and economic developments reduced housing demand and households shifted their preferences towards mortgage loans repayment. Firms in construction sector also reduced their profit margins and shed significant part of their employment. As of 2014 house prices are estimated as being aligned with income, interest rates and foreign demand.

If the upward trend in household income and the downward trend in interest rates on mortgage loans continue, the demand for housing can be expected to recover further. Construction activity will probably remain

depressed until a more substantial rise in prices is observed. Therefore, the increase in demand for housing can be expected to bring an increase in housing prices in the medium term, in the absence of increase in supply. On the other hand, the low yield and interest rates environment and the expectations for an increase in housing prices and rents, can lead to increase in speculative demand for housing as well. In the long term a correction from the supply side can be expected, due increase in prices and profit margins. This expected upturn will be more tamed in the absence of FDI inflows, compared to the housing boom of 2003-2008. In light of these expectations, no significant overvaluation can be anticipated in the medium term.

References

Addison-Smyth, D., K. McQuinn, G. O'Reilly (2009), "*Modelling Credit in the Irish Mortgage Market*", Central Bank of Ireland Research Technical Paper, 9/RT

Angello, L., L. Schuknecht (2009), "*Booms and Busts in Housing Markets: Determinants and Implications*" ECB Working Paper No. 1071

Case, K., R. Schiller (1988), "*The Behavior of Home Buyers in Boom and Post-Boom Markets*", NBER Working Paper No. 2748

Case, K., R. Schiller (1990), "*Forecasting Prices and Excess Returns in the Housing Market*", AREUEA Journal, 18, 253-273

Case, K., R. Schiller (2004), "*Is There A Bubble in The Housing Market?*", Cowles Foundation Paper No. 1089

Detken, C., F. Smets (2004), "*Asset Price Booms and Monetary Policy*", European Central Bank Working Papers Series, No 364, May

Egert, B., D. Mihaljek (2007), "*Determinants of House Prices in Central and Eastern Europe*", BIS Working Papers, No 236, September

Englund, P., Y. Ioannides (1997), "*House Price Dynamics: An International Empirical Perspective*", Journal of Housing Economics, 6

European Commission (2012), Quarterly Report on The Euro Area, Volume 11, #4, 2012, Focus: "*Assessing the dynamics of house prices in the euro area*"

Gattini and Hiebert (2010), "*Forecasting and assessing euro area house prices through the lens of key fundamentals*", European Central Bank Working Papers Series, No 1249

Glaeser, E., J. Gyourko (2007), "*Arbitrage in Housing Markets,*" NBER Working Paper 13704

Glaeser, E., J. Gyourko, A. Saiz (2008), "*Housing Supply and Housing Bubbles*", Harvard Institute of Economic Research, Discussion Paper No. 2158

Hamilton, J. D. (1994), "*Time Series Analysis*", Princeton University Press.

Juselius, K., D. Hendry (2000), "Explaining Cointegration Analysis: Part II", Discussion Papers 00-20, University of Copenhagen. Department of Economics.

Kearl, J. (1979), "*Inflation, mortgages, and housing*", Journal of Political Economy, 87

National Statistical Institute, Statistical Yearbook 1985, 1990, 1995

Ortalo-Magné and Prat (2007), "*The Political Economy of Housing Supply: Homeowners, Workers and Voters*," STICERD Theoretical Economics Discussion Paper No. 514

Palacin, J., R. Shelburne (2005), "*The Private Housing Market in Eastern Europe and the CIS",* ECE Discussion Papers Series 2005_5, UNECE.

Poterba, J. (1984), "*Tax Subsidies to Owner-Occupied Housing: An Asset-Market Approach*", The Quarterly Journal of Economics, Vol. 99, No. 4 (Nov., 1984), pp. 729-752

Poterba, J. (1991), "*House Price Dynamics: The Role of Tax Policy and Demography*", Brookings Papers on Economic Activity, 2:1991

Rahbek, A., E. Hansen, J.G. Dennis (2002), *"ARCH innovations and their impact on cointegration rank testing*", Preprint no. 12, 1998, Department of Theoretical Statistics, Working paper no. 22, Centre for Analytical Finance.

Stiglitz, J. (1990),"*Symposium on Bubbles*", The Journal of Economic Perspectives, Vol. 4, No. 2. (Spring, 1990), pp. 13-18

Topel, R., S. Rosen (1988), "Housing Investment in the United States", Journal of Political Economy, Vol. 96, No. 4 (Aug., 1988), pp. 718-740

Zhu, M. (2014), "Housing Markets, Financial Stability and the Economy", Opening Remarks at the Bundesbank/German Research Foundation/IMF Conference

Appendix

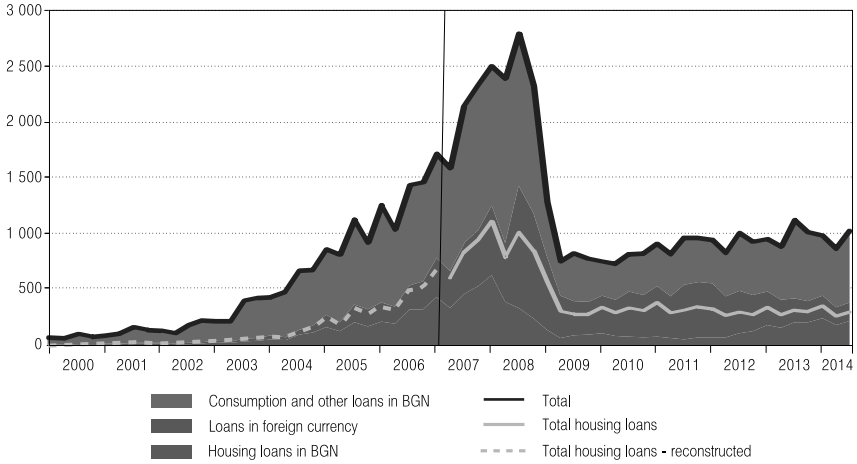
A1. Reconstructed Data on Interest Rates and Loans

A break in interest rates statistics in Bulgaria in the beginning of 2007 causes additional difficulties in analyzing and modeling housing market developments. Prior to 2007 data for interest rates on new business housing loans is available only for loans in BGN, while in other currencies data is available on an aggregated level – for all long-term loans of households. Long-term loans are defined as loans with maturity one year and longer. We should also note that in 2007 the definition of new business loans was changed.

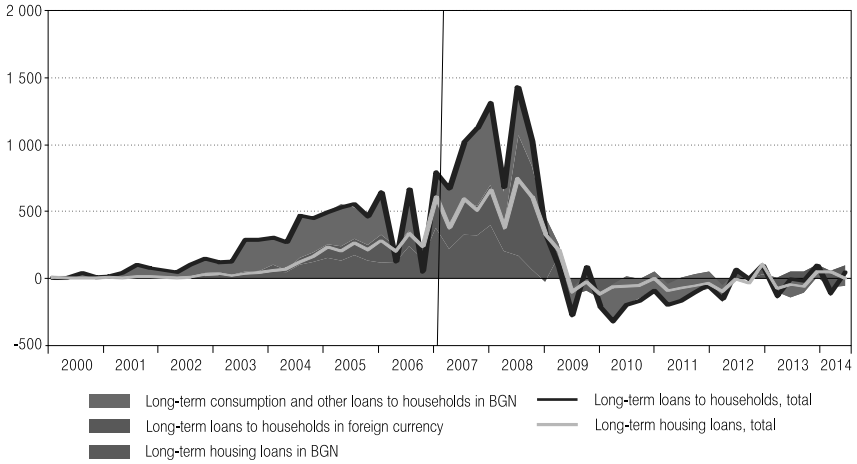
We reconstructed the data for interest rates on housing loans weighted for all currencies before 2007 based on some necessary and somewhat arbitrary assumptions. To weight the interest rates we need to construct time series for new business long-term housing loans in foreign currency on the basis of the available data for total long-term new business loans to households in foreign currency. We assume a constant share of 80% of housing loans in all long-term loans to households in foreign currency as this is their share for 2007 (the first year for which data is available). Monetary statistics data for the change in the outstanding volumes of loans also confirms that housing loans were the major part of all long-term loans to households in foreign currency.

Next, we make an assumption for the spread between interest rates on housing loans in foreign currency and consumption and other long-term loans to households in foreign currency. We again assume that the spread observed in 2007, 1.37 p.p., was the one in the period from 2000 to 2006. Having data on the weighted interest rate on long-term loans for households prior to 2007 and the reconstructed new business loans we construct series for interest rates on housing loans in foreign currency and the overall weighted interest rate on housing loans.

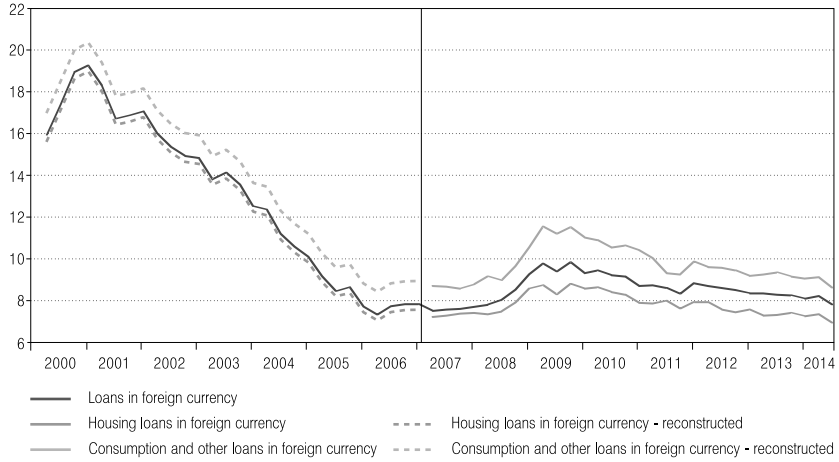
Interest Rate Statistics Data – Long-term Loans to Households – New Business, Quarterly Data, BGN mln.



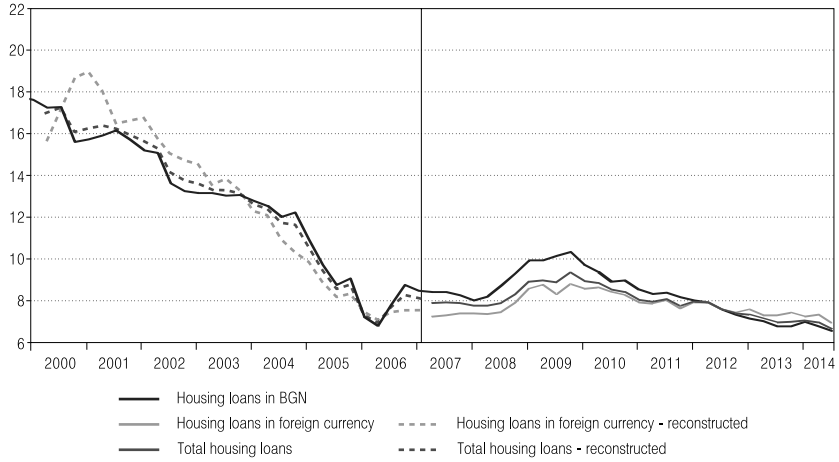
Monetary Statistics Data – Long-term Loans to Households – Change in Outstanding Stock, Quarterly Data, BGN mln.



Interest Rate Statistics Data – Interest Rates on Long-run Loans for Households in Foreign Currency for New Business, Quarterly Weighted



Interest Rate Statistics Data – Interest Rates on Long-run Housing Loans for New Business, Quarterly Weighted

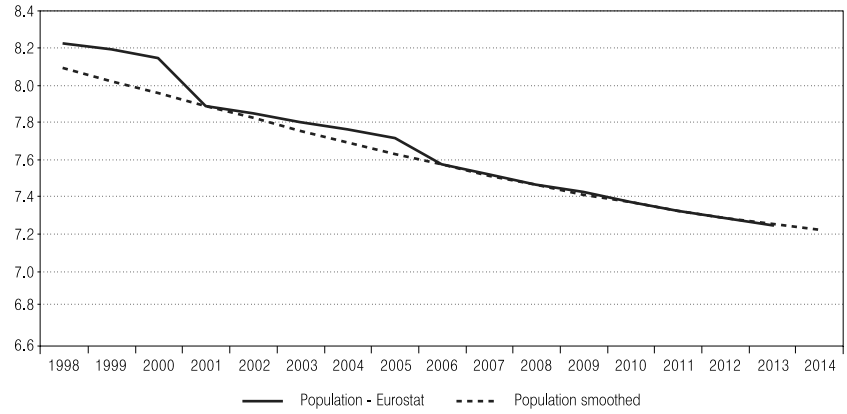


A2. Population

To calculate indicators like income per capita, a population data with quarterly frequency was needed. The official statistics on population, however, show drastic changes in demographics in the years of national census, which maybe is an indication that the standard means for yearly estimation of the population, used by the statistical office, cannot resemble the fast shrinkage of the number of individuals. One possible reason for this might be the unregistered emigration outflows. In our research we preferred to rely solely on census data and the available forecast for 2015 and to interpolate the data for the years in between. This procedure gave us smoother population data, without drastic changes in demographic dynamics, as one would typically expect to be observed. As a next step, the smoothed population data was additionally interpolated to a quarterly frequency using the Boot-Feibes-Lisman method.

Population

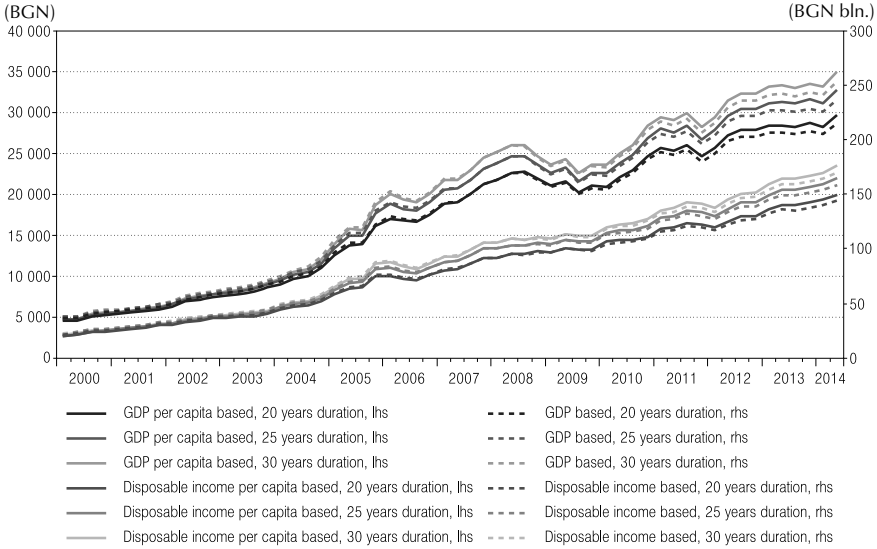
(end of period; mln.)



A3. Hypothetical Borrowing Value

Several specifications of the hypothetical borrowing value indicator have been tested. As a measure of personal income we have used both GDP and disposable income, both total and per capita, calculated with the smoothed population series (see appendix A2). Three different durations were also assumed for the mortgage loans – 20, 25 and 30 years. In terms of dynamics all tested specifications show similar results independent of the duration. GDP based indicators have significantly more pronounced reaction during the 2008-2009 crisis, and per capita indicators give slightly higher growth rates on average. Our preferred measure at the end is the hypothetical borrowing value indicator based on GDP per capita with duration of the loan of 25 years.

Hypthetical Borrowing Value



Hypthetical Borrowing Value – Annual Growth Rates

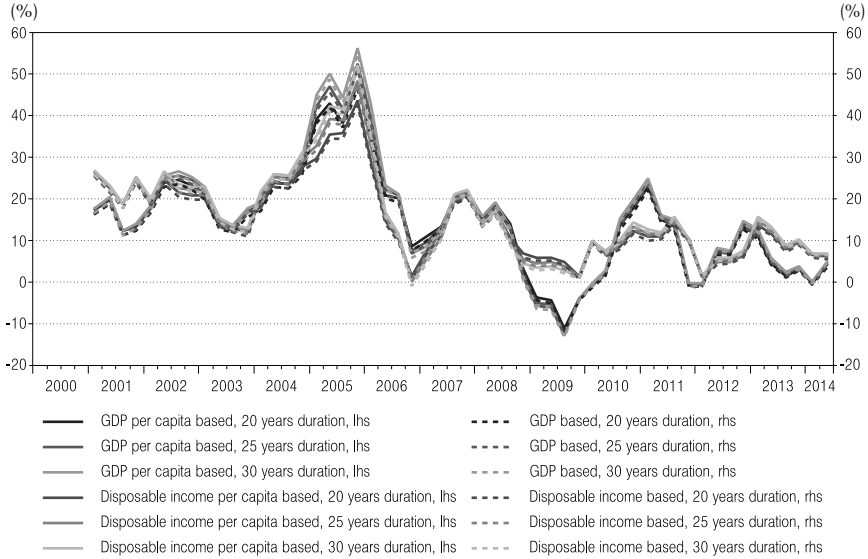


Table A1: **Unit Root Phillips–Perron Test**

H0: Variable has a unit root	Adj. t-statistics			
	Level		First diff.	
	Sample 2000q1-2014q2	Sample 2004q1-2014q2	Sample 2000q1-2014q2	Sample 2004q1-2014q2
House prices	-1.47		-2.94***	
Real house prices	-1.33		-3.48 ***	
Hypothetical borrowing value	-2.38		-4.34 ***	
FDI in real estates	-1.45	-2.38	-3.70 ***	-3.15 ***
FDI in real estates, house prices deflated	-1.44	-2.08	-4.49 ***	-3.89 ***
Gross domestic product	-2.46		-3.79 ***	
Gross domestic product per capita	-2.53		-3.53 ***	
Real gross domestic product	-2.34		-2.65 ***	
Real gross domestic product per capita	-2.46		-2.40 **	
Housing interest rate	-1.96		-3.89 ***	
Real housing interest rate	-2.09		-4.69 ***	
New housing loans	-2.52		-3.79 ***	
New housing loans, house prices deflated	-2.88 *		-4.67 ***	
Construction production	-1.69		-2.44 ***	
Share of value added in construction	-1.67		-11.05 ***	
Test critical values:				
*** 1% level	-3.55	-3.60	-2.61	-2.62
** 5% level	-2.91	-2.93	-1.95	-1.95
* 10% level	-2.59	-2.60	-1.61	-1.61

Note: *Unit root tests on variables are done with automatic lag length selection based on SIC, for variables in levels constant is included in the equation. The Hypothetical borrowing value was also tested under the assumption of deterministic linear trend and the test again suggested that the series is I(1). FDI in real estates was additionally tested for unit roots on the shorter sample 2004-2014q2 as for 2000-2003 there are no official data and the series was extrapolated.*

A4. The Model

We have followed the Hendry and Juselius (2000) procedure to construct our model. Both Schwarz information criterion and Hannan-Quinn information criterion suggested inclusion of 3 lags in a VAR model of variables in levels (HP, HBV, FDI). Under the assumption of linear deterministic trend in the data trace test and maximum eigenvalue test indicate one cointegration relationship (Table A2). Tests on significance of the variables in the relationship and weak exogeneity suggest that all three variables are significant and only house prices correct to the estimated long-run relationship (with error-correction term of -0.23). Residual test for normality was passed. White heteroscedasticity test rejected residual homoscedasticity at 0.05 critical value but Rahbek et al. (2002) have shown that the cointegration rank tests are robust against moderate residual ARCH effects. LM autocorrelation test also showed first order serial correlation in the residuals. As OLS estimate of the cointegration vector should still be consistent even in the presence of serial correlation (Hamilton (1994) we find our model appropriate for illustrative purposes. The estimated cointegration equation is given in Table A3.

Table A2: **Cointegration Tests**

Hypothesized No. of CE(s)	Eigenvalue	Trace Statistic	0.05 Critical Value	Prob.**	Max-Eigen Statistic	0.05 Critical Value	Prob.**
None *	0.383	36.982	29.797	0.006	26.589	21.132	0.008
At most 1	0.141	10.393	15.495	0.252	8.366	14.265	0.343
At most 2	0.036	2.028	3.841	0.155	2.028	3.841	0.155

Table A3: **Cointegration Vector**

Cointegration equation	
HP	1
HBV	-0.31 (-0.05)
FDI	-0.16 (-0.02)
Constant	-1.91

Note: *Standard errors in parenthesis*

Table A4:

Distribution of population by tenure status								
	2005	2006	2007	2008	2009	2010	2011	2012
Owner	85.4	85.4	87.6	87.1	86.8	86.9	87.2	87.4
Owner, with mortgage or loan	2.3	2.3	2.5	9.9	9.3	1.8	1.5	2.0
Owner, no outstanding mortgage or housing loan	83.1	83.1	85.1	77.3	77.5	85.2	85.7	85.3
Tenant	14.6	14.6	12.4	12.9	13.2	13.1	12.8	12.6
Tenant, rent at market price	3.0	3.0	2.3	1.6	2.0	2.1	1.7	1.3
Tenant, rent at reduced price or free	11.6	11.6	10.1	11.2	11.2	10.9	11.1	11.3

Source: *Eurostat*

DP/104/2017

Relationship between Inflation, Potential Output and Structural Unemployment in Bulgaria

Daniel Kasabov, Petra Kotseva, Andrey Vassilev, Mihail Yanchev

Abstract. The paper estimates a Phillips curve type relationship for the Bulgarian economy in a setup that takes into account the simultaneous interaction between inflation, output and unemployment. Using a state space model estimated with Bayesian methods on quarterly data from 1999 to 2015, we assess the cyclical developments in unemployment and output, and identify the main drivers of inflation over the period. Compared to other methods such as univariate filtering or production function approaches, our methodology enables the combination of information from different time series and admits a structural interpretation of the results in terms of the underlying shocks.

Резюме. Изследването оценява зависимост от вида на Кривата на Филипс за българската икономика в набор, който взема предвид едновременното взаимодействие между инфлацията, производството и безработицата. Използвайки пространствен модел на състоянието, оценен с Байесови методи въз основа на тримесечни данни за периода от 1999 г. до 2015 г., ние определяме цикличната динамика на безработицата и производството и установяваме основните движещи сили на инфлацията през този период. В сравнение с други подходи, като едномерното филтриране или производствената функция, нашата методика дава възможност да се комбинира информацията от различни динамични редове и допуска структурна интерпретация на резултатите, що се отнася до основните шокове.

Daniel Kasabov, Economic Research and Forecasting Directorate, Bulgarian National Bank, kasabov.d@bnbank.org
Petra Kotseva, Economic Research and Forecasting Directorate, Bulgarian National Bank, rogleva.p@bnbank.org
Andrey Vassilev, Economic Research and Forecasting Directorate, Bulgarian National Bank, Corresponding author, vassilev.a@bnbank.org
Mihail Yanchev, Economic Research and Forecasting Directorate, Bulgarian National Bank (worked in the period 2012–2015), yanchev.mihail@gmail.com

We thank two anonymous referees, our colleagues from the Economic Research and Forecasting Directorate for the support and helpful discussions and especially Mariella Nenova for detailed and valuable comments on the paper.

Contents

Introduction .....

Literature Overview .....

Identifying Relationships .....

Laws of Motion for Equilibrium Variables.....

Methodology and Data.....

Model Specification.....

Parameter Calibration .....

Estimation Strategy and Data .....

Results .....

Conclusions .....

References .....

Appendix 1. Model Specification and Parameters .....

Appendix 2. Estimates of Unobservable Variables and Additional Model Output .....

Introduction

The historical analysis of cyclical developments in the Bulgarian economy faces challenges if traditional methods like univariate filters or a production function approach are used. Due to country-specific factors like the transition to a market economy and the consequences of bad economic policies in the mid-1990s, Bulgaria went through a period of relatively high inflation coupled with high unemployment and, likely, an increase in structural unemployment in the period 1999–2003. This suggests that at the time there was a slowdown in potential output growth and a significant negative output gap. However, results from conventional univariate filters contradict this view and simple Phillips curve models focused solely on the relationship between output and inflation cannot explain the developments in this period and fit them with the more recent history. This underscores the importance of modelling the interaction between inflation, output and unemployment simultaneously to obtain a unified picture of the observed developments.

This paper aims to contribute to filling the above gap by analysing the factors that influence inflation dynamics for the Bulgarian economy and model the relationship between inflation and measures of business cycle fluctuations in the framework of a Phillips curve type relationship. This in effect allows us to obtain estimates for the output gap and the natural level of unemployment, using other relevant economic information. For that purpose we use a state space model with Kalman filter, based on the methodology of Benes et al. (2004, 2010). The model is estimated using Bayesian methods to allow for the inclusion of prior information and to alleviate the potential drawbacks from using a relatively short sample. The inclusion of external supply-side shocks to affect directly inflation and the use of model-consistent inflation expectations and hard anchoring are important features of the model. More generally, our methodology enables a structural interpretation of the results in terms of the underlying shocks.

The model results are compared to the output of the simple HP filter. It is expected that the state space model will give more reliable estimates of the output gap, especially in regards to the period of 2000–2003, when high inflation is coupled with high unemployment in the context of a post-recession recovery. Furthermore, the state space model can be employed in medium and long-term growth forecasting, in addition to uses related to the analysis of the cyclical position of the economy. Results from such forecasting exercises can also be used in order to formulate appropriate

expectations about indicators relevant to economic policy like the cyclically-adjusted budget deficit and the gross government debt.

We find a significant negative output gap during the period 2000–2003, in line with the high unemployment at the time, but despite high inflation, which would traditionally be associated with an overheating phase of the business cycle. In contrast, HP filter estimates point to a negative output gap close to zero during the same period. During the period 2003–2008, as well as during the recession period after 2009, estimates of the output gap are similar to the ones obtained using the HP filter. In regards to the characteristics of the Phillips curve relationship, we obtain a low estimate of the elasticity of inflation to demand. According to our results, shocks in oil prices and other international commodity prices are more important for inflation dynamics than demand. The historical decomposition of inflation dynamics reveals that other cost push shocks, not explicitly included in the model, matter the most for inflation dynamics during periods of high inflation. Such cost push shocks are difficult to specify precisely but they probably combine spill-overs from administrative price changes to core inflation, the dynamics of labour costs and dynamics of housing prices and their effects on consumer price inflation. Our results seem to be consistent with other empirical studies on the drivers of inflation done in the recent years (IMF (2015), Iossifov and Podpiera (2014)).

The paper is structured as follows. In the next section we provide a concise overview of the literature and recently used empirical methods. Following that, we justify the selection of a specific approach and describe the details of the methodology and the dataset used for the estimation. The subsequent section presents and discusses the results obtained. Finally, conclusions arising from the work are formulated.

Literature Overview

The relationship between measures of inflation and real variables at the macroeconomic level has always been of special interest in economics. While traditionally such relationships are referred to as *Phillips curves,* the notion has long surpassed the original Phillips (1958) study of money wages and unemployment to include real marginal costs, output gaps or unemployment gaps on the real side and various inflation indicators on the price side.

Starting from the seminal Phillips (1958) paper, the thinking on the Phillips curve-type mechanisms has undergone several transformations (see Fuhrer et al. (2009) for a more detailed account of this evolution). The initial interpretation of the Phillips curve as embodying a permanent trade-off between inflation and unemployment and hence implying that monetary policy could have long-term effects was criticized in the 1960s. In several influential papers Samuelson, Solow, Friedman and Phelps (Samuelson and Solow (1960), Friedman (1968), Phelps (1978)) argued that such a trade-off was valid only in the short run, thus limiting the usefulness of monetary policy measures to the near term only.

The advent of the rational expectations paradigm in the 1970s brought about a further strengthening of the initial criticism. A classical rational expectations model would imply not only a short-term role for monetary policy but would additionally require policy actions to surprise economic agents in order to have any effect at all. Results of this type are discouraging when taken at face value but they critically hinge on the assumption of price level flexibility, which is dubious from an empirical standpoint.

The need to square the assumptions of rational expectations models with empirical evidence spawned a body of literature trying to incorporate various forms of price inertia in a rational expectations framework. Introducing such rigidities ultimately makes the models more plausible and also reasserts a (temporary) role for monetary policy. The contributions in this strand of the literature are too numerous to review here but two studies are noteworthy as they play an important part in microfounded models with a Phillips curve to date. Calvo (1983) introduced a mechanism of time-dependent pricing under which firms in each period are constrained to stick to their current price with a certain probability. Rotemberg (1982, 1983) pursued the idea of state-dependent pricing and developed a mechanism under which firms pay an adjustment cost dependent on the extent of price change whenever they update prices. Both models have practically identical implications at the macro level and in a sense may be seen as only differing in the interpretability of their microfoundations by virtue of the different choice of modelling the price rigidities.

Starting from the second half of the 1980s and continuing through the 1990s, the literature on the Phillips curve took essentially its modern-day form by branching out in two directions. Gordon (2009) provides an excellent account of these developments and colourfully refers to the two branches of the bifurcation as the “left and right fork in the road”.

One branch followed up on the rational expectations construct with rigidities and strove to deepen further its microfoundations, ultimately incorporating the Phillips curve into a full-fledged DSGE-type model. The resulting Phillips curve specification – whether embedded in a DSGE model or taken as a standalone object for estimation – is commonly referred to as the New Keynesian Phillips Curve (NKPC). This is Gordon's "right fork in the road".

The other branch ("the left fork in the road") took a more empirical lean from the start and postulated a Phillips curve formulation comprising three sets of factors: inertia, demand factors and supply factors. Inertia is traditionally captured by lagged values of the inflation rate, demand factors are taken on board by means of current and lagged values of a gap-type measure (the deviation of unemployment from its "natural" level) and supply factors include current and lagged values of variables such as oil and food prices, real exchange rates or price control dummies. This approach is typically referred to as "the triangle model" (see Gordon (1998) for a typical implementation).

There is no consensus on the issue of which branch provides a more appropriate specification in terms of empirical performance. Indeed, the literature seems to have settled on the view that the NKPC and the triangle model perform differently in different historical episodes and hence both have their place in empirical applications. As a general guideline, however, the NKPC can be viewed as an appropriately restricted version of the triangle model and hence can be employed when there is evidence that these restrictions are plausible for the particular case in point or when the ultimate goal is to develop a larger microfounded model. In contrast, the triangle model seems more appropriate for situations where there is no convincing prior information justifying the imposition of particular restrictions and hence a more flexible specification is warranted.

More recent empirical work based on the "right fork" approach uses several extensions of the specification of the NKPC to bridge the gap between the theory and the data. In their influential paper, Gali and Gertler (1999) extend the theoretical framework by allowing for a fraction of firms to use a backward-looking rule of thumb to set prices. In this way they developed a "hybrid" NKPC model that accounts for inflation persistence, which is observed empirically. Other studies, including Gali and López-Salido (2001), Balakrishnan and López-Salido (2002) and Batini, Brian and Stephen (2005) augment the production function by including an imported intermediate goods input, which takes into account supply-side shocks related to international prices of commodities.

Due to the assumption of rational expectations in microfounded models, estimating the NKPC poses the problem of endogeneity, which arises from possible correlation of the errors from the model with the regressors. Since in this case the ordinary least squares (OLS) estimates of the parameters in the estimated equation might be inconsistent, the standard approach is to use instrumental variables to eliminate the correlation. There are many different approaches to using instruments to eliminate the effect of variable and residual correlation. A common practice in empirical research, including Gali and Gertler, is to estimate the NKPC model with the General Method of Moments (GMM) estimator.

When following the "left fork" approach a crucial step in modelling a Phillips curve relationship is estimating the unobservable variables that capture the state of the business cycle like the output gap or the unemployment gap. The most common method for estimation that has been used in the past decade is the univariate Hodrick-Prescott filter, which however has some significant limitations like the end-point bias and the fact that the estimation in the long-run is not robust to the introduction of new observations. Additionally, univariate filters do not take into account other relevant information, except the one contained in the series that is being filtered, which is another considerable setback.

In order to overcome these disadvantages, Benes et al. (2004, 2010) employed a methodology of measuring potential output and the output gap that makes use of the empirical relationships between relevant variables like the actual and potential GDP, unemployment, inflation and capacity utilization in the industry within the framework of a simple macroeconomic model. The model provides a multivariate filter, which can be adapted to many countries and is flexible in a way that it allows for the framework to be refined and more information added.

The most basic version of the model (Benes et al. 2010) is described and discussed below. The model equations are presented in two groups: Identifying relationships and laws of motion for equilibrium variables.

Identifying Relationships

$$\pi 4_t = \pi 4_{t-1} + \beta y_t + \Omega(y_t - y_{t-1}) + \varepsilon_t^{\pi 4} \tag{1}$$

$$u_t = \phi_1 u_{t-1} + \phi_2 y_t + \varepsilon_t^u \tag{2}$$

$$c_t = \mathrm{k}_1 c_{t-1} + \mathrm{k}_2 y_t + \varepsilon_t^c \tag{3}$$

Laws of Motion for Equilibrium Variables

Equilibrium unemployment or NAIRU

$$\bar{U}_t = \bar{U}_{t-1} + G_t^{\bar{U}} - \frac{\omega}{100} y_{t-1} - \frac{\lambda}{100}(\bar{U}_t - \bar{U}^{ss}) + \varepsilon_t^{\bar{U}} \tag{4}$$

where0 $G_t^{\bar{U}} = (1-\alpha)G_{t-1}^{\bar{U}} + \varepsilon_t^{G^{\bar{U}}}$ (5)

Potential output

$$\bar{Y}_t = \bar{Y}_{t-1} - \theta(\bar{U}_t - \bar{U}_{t-1}) - \frac{(1-\theta)(\bar{U}_{t-1} - \bar{U}_{t-20})}{19} + \frac{G_t^{\bar{Y}}}{4} + \varepsilon_t^{\bar{Y}} \tag{6}$$

where $G_t^{\bar{Y}} = \tau G_{ss}^{\bar{Y}} + (1-\tau)G_{t-1}^{\bar{Y}} + \varepsilon_t^{G^{\bar{Y}}}$ (7)

Equilibrium capacity utilization

$$\bar{C}_t = \bar{C}_{t-1} + G_t^{\bar{C}} + \varepsilon_t^{\bar{C}} \tag{8}$$

where $G_t^{\bar{C}} = (1-\delta)G_{t-1}^{\bar{C}} + \varepsilon_t^{G^{\bar{C}}}$ (9)

Perceived long-term inflation objectives

$$\pi 4_t^{LTE} = \pi 4_{t-1}^{LTE} + \varepsilon_t^{\pi 4^{LTE}} \tag{10}$$

Output gap equation

$$y_t = \rho_1 y_{t-1} - \frac{\rho_2}{100}(\pi 4_{t-1} - \pi 4_{t-1}^{LTE}) + \varepsilon_t^{y} \tag{11}$$

The group of identifying relationships comprises three equations: a simplified Phillips curve type relationship describing the short-run trade-off between the output gap (y_t) and inflation ($\pi 4_t$), adjusted for staggered prices and speed limit effects – equation (1), Okun's law type equation describing the interplay between unemployment gap (u_t) and the output gap – equation (2) and a similar link between the capacity utilization in manufacturing gap (c_t) and the output gap – equation (3). The three gaps are defined as deviation of the actual from its equilibrium level.

The laws of motion for the equilibrium variables are presented in equations (4) through (11). The equilibrium unemployment rate ($\bar{U}_t$), which corresponds to the notion of NAIRU, is allowed to deviate from its fixed steady-state level ($\bar{U}^{ss}$) and follows a stochastic process with both transitory ($\varepsilon_t^{\bar{U}}$) and more persistent shocks ($G_t^{\bar{U}}$) and it also depends on the output gap, which introduces a partial hysteresis effect, as captured in equation (4). In equilibrium the potential output ($\bar{Y}_t$) depends on changes in the underlying trend growth of potential ($G_t^{\bar{Y}}$) and on changes in the NAIRU – see equation (6). The inclusion of the latter causes short-term

and medium-term potential growth to vary from the underlying trend growth of potential in order to capture the impact of changes in the equilibrium level of employment and the effect of the induced changes in the capital stock via a Cobb-Douglas production function, in which θ is the labour share. On the other hand the equilibrium trend growth of potential output is not constant and follows serially correlated deviations from its steady-state growth rate ($G_{ss}^{\bar{Y}}$) – equation (7). Equilibrium capacity utilization ($\bar{C}_t$), similarly to the unemployment equation, includes both transitory level shocks ($\varepsilon_t^{\bar{C}}$) as well as more persistent shocks in the growth rate ($G_t^{\bar{C}}$), see equation (8). Perceived long-term inflation expectations ($\pi 4_t^{LTE}$), which follow an adaptive process with revisions to their level of the previous period – equation (10), enter the equilibrium output gap equation so that there is a negative effect on demand from inflation deviating from its target, as presented in equation (11). This recognizes the idea that monetary policy exerts influence on core inflation through control over the output gap, which is generally consistent with a variety of monetary regimes.

To estimate the model Benes et al. employ a combination of the Kalman filter and the regularized maximum likelihood method following Ljung (1999). This allows for the definition of prior distributions, which can prevent parameters from wandering into illogical directions, but is also important for the estimation of parameters for which actual data can give little insight. The prior constraints imposed by the method are useful in the cases where the data set is too limited to provide enough information for the estimation of the particular variables.

Benes et al. apply their method to a number of countries and conclude that the multivariate filter estimates of the output gap are much more robust to introduction of a newer data set with longer time series, as compared to the univariate Hodrick-Prescott filter. Moreover, the multivariate filter estimates confirm that the growth rate of potential output can vary substantially over time, in contrast to the more conventional production approach, which assumes smoothly changing productivity growth. Several studies have followed the approach proposed by Benes et al., among which Fabiani and Mestre (2001), Sramkova (2010), Bokan and Ravnik (2012), where the abovementioned conclusions were confirmed.

Methodology and Data

Since the aims of our study are not only to analyse the factors that influence inflation dynamics and model the relationship between inflation and measures of business cycle fluctuations, but also to obtain reliable estimates for the output gap and the natural level of unemployment, using other relevant economic information, we believe the Benes et al. approach is preferable and more appropriate. Specifically, one of the useful outcomes of the model is a potential output estimate alternative to the ones derived by a Hodrick-Prescott filter or the production function approach, which are considered somewhat inferior and lacking in the literature on the topic (Benes and N'Diaye (2004)).

Model Specification

Several adjustments to the methodology of Benes et al. were made in order to fit the case of a small open economy with a currency board arrangement, without independent monetary policy, in this framework. The effort was directed towards arriving at a parsimonious specification, which, however, would not disregard any important prior theoretical considerations. First, we introduce external supply-side shocks to directly affect inflation. The ones that we identify as the most important for Bulgarian inflation dynamics are the international oil and other import prices and the real effective exchange rate. These are incorporated in the model similarly to Apel et al. (1999) by transforming them into shocks relative to inflation by deflating and standardizing them. In order to accommodate this addition we reformulate the Phillips curve equation in terms of quarterly annualized changes instead of annual changes as it is done in the original model by Benes et al. Secondly we introduce model-consistent inflation expectations and a hard anchoring element to the Phillips curve relationship. Using only model consistent expectations resulted in an explosive behaviour of the model, despite any combination of priors on parameters or residuals. Adding hard anchoring improved the model stability significantly and allowed to keep the model-consistent expectations. This improved the characteristics of the model compared to a version where some combination between consumer survey-based expectations and hard anchoring was used. With these modifications our Phillips curve relationship took this form:

$$\pi_t = \eta_1 \pi_{t-1} + \eta_2 \pi_{t+1}^E + (1 - \eta_1 - \eta_2)\pi_{ss} + \beta y_t + \Omega(y_{t-1} - y_{t-2}) + \sum_{k=0}^{2} \gamma_k z_{t-k} + \varepsilon_t^{\pi}, \quad (12)$$

where $z_t = (z_t^l, z_t^m, z_t^s)'$ is the vector of the supply side shocks of respectively the oil price, import prices and the real effective exchange rate, and the corresponding coefficients are given by $\gamma_k = (\gamma_k^l, \gamma_k^m, \gamma_k^s)$ for k = 0 to 2. The variables π_{t+1}^E and π_{ss} are respectively model-consistent expectation for one period ahead and the long-run steady state for inflation.

The effect of monetary policy on output gap through the difference between actual and expected inflation was taken out, since Bulgaria does not have independent monetary policy and it is rather the ECB monetary policy and policies of other non-Euro area central banks that (indirectly) affect the economy of Bulgaria. Thus, the block of equations for the potential and actual output and output gap was formulated in the following way, broadly following the Benes et al. construct:

Output gap identity:

$$y_t = Y_t - \overline{Y}_t \tag{13}$$

Potential output:

$$\overline{Y}_t = \frac{G_t^{\overline{Y}}}{4} + \overline{Y}_{t-1} + \varepsilon_t^{\overline{Y}} \tag{14}$$

Growth of potential output:

$$G_t^{\overline{Y}} = \tau G_{ss}^{\overline{Y}} + (1-\tau) G_{t-1}^{\overline{Y}} + \varepsilon_t^{G^{\overline{Y}}} \tag{15}$$

Output gap:

$$y_t = \rho y_{t-1} + \varepsilon_t^y \tag{16}$$

The unemployment block of the model is formulated in the following way:

Unemployment gap identity:

$$u_t = \overline{U}_t - U_t \tag{17}$$

Equilibrium unemployment:

$$\overline{U}_t = \overline{U}_{t-1} + G_t^{\overline{U}} - \frac{\omega}{100} y_{t-1} - \frac{\lambda}{100}(\overline{U}_{t-1} - \overline{U}^{ss}) + \varepsilon_t^{\overline{U}} \tag{18}$$

Growth of equilibrium unemployment:

$$G_t^{\overline{U}} = (1-\alpha) G_{t-1}^{\overline{U}} + \varepsilon_t^{G^{\overline{U}}} \tag{19}$$

No capacity utilization block was included in our model. Our experiments showed that the capacity utilization series introduced more noise than valuable information for the estimation of the unobservables, probably because of its survey-based nature.

Additional modifications were tested like the inclusion of industrial production, instead of capacity utilization and the current account balance, similarly to Bokan and Ravnik (2012). However, the benefit of complicating the model further in this direction seemed little to none in terms of model stability and quality of the estimates of the unobservable variables. Therefore, similarly to the capacity utilization, these indicators were not included in the system.

Parameter Calibration

The properties of the model depend on the values of the parameters representing the long-run steady states for inflation (π_{ss}), growth rate of potential output ($G_{ss}^{\overline{Y}}$) and the equilibrium unemployment rate ($\overline{U}^{ss}$), which are typically calibrated. The challenge in setting these values lies in striking the right balance between empirical performance, which implies reliance on historical data, and theoretical consistency, which requires an assessment of the expected future development of the economy.

The calibration of the steady state for inflation is direct. Based on the premise that inflation is ultimately a monetary phenomenon, and taking into account both the present currency board framework in Bulgaria and the obligation to join the euro area in the future, we set the steady state inflation rate to equal 2%. The primary objective of the ECB is to maintain price stability which is defined in quantitative terms as 'inflation rates below, but close to, 2% over the medium term'. Although the inflation objective of ECB is defined in terms of a narrow range, solving the model requires a precise parameter value of the steady-state. Therefore, the choice of 2% for the steady state inflation rate is considered to be appropriate as any other value would be arbitrary and not necessarily in line with the perception of the ECB. This approach is often adopted in empirical research (Bulíř et al 2008).

To calibrate the value of the steady state growth of potential output, we took annual World Bank data on real GDP growth rates for a set of countries and country groups[1] over the period 1960–2015. For this sample, we computed the average annual GDP growth for each country or country group. Assuming that business cycle fluctuations cancel out sufficiently well over a long period of time, average GDP growth should provide a reasonable approximation to long-run potential output growth. Additionally, one would be interested in economies that are expected to be at or rela-

[1] The precise country and country group coverage can be retrieved from the metadata for the series NY.GDP.MKTP.KD.ZG from the World Bank Databank.

tively close to the steady state to avoid picking up the effects of transition dynamics for less developed economies. For this reason we concentrated on economies at an advanced stage of economic development, as defined below.

In order to define more precisely the set of advanced economies, we proxied the state of development by using World Bank data on GDP per capita in PPP, measured in constant 2011 international dollars, for a set of countries and country groups[2] over the period 1960–2015. Both the dataset on GDP per capita in PPP and the dataset on real GDP growth include approximately 240 countries and country groups each. Thus, the two datasets are virtually identical in terms of country coverage. As the aim was to obtain an up-to-date picture, we computed the average GDP per capita for the period from 2010 to the end of the sample for each country. This approach allows the use of the most recent information while reducing the risk of arbitrary fluctuations in the outcomes that might exist if one were to use only the most recent observation for each country or region. We then sorted the data according to the computed period-average GDP per capita and took the subsample covering the top 85 to 95 percent. This procedure was applied in order to exclude some outliers in the sample, while capturing the relevant subset of advanced countries. The results include a subset of 24 countries and country groups, containing countries like the Netherlands, Austria, Germany and the UK, as well as groups like North America and the *High income* countries. Upon comparing the respective average real GDP growth rates, we chose to use the growth rate for the *High income* group as representative of what potential output growth for Bulgaria might eventually be. Thus, we calibrated the steady state growth of potential output at 3.1%.

The calibration of the steady state value for equilibrium unemployment is the most complicated one, as labour market outcomes are substantially affected by a number of factors, including demographic developments, institutional arrangements and the structure of the economy in general. While a detailed and comprehensive analysis of such issues is obviously beyond the scope of a calibration exercise, the preceding considerations suggest that an approach transcending a standalone analysis of unemployment rates is called for. Our implementation of such an approach is based on a combination of two elements – a "historical" and a "forward-looking" estimate – to produce a final calibrated value for steady state equilibrium

[2] The precise country and country group coverage can be retrieved from the metadata for the series NY.GDP.PCAP.PP.KD from the World Bank Databank.

unemployment. Both estimates are constructed using data for the European economies.

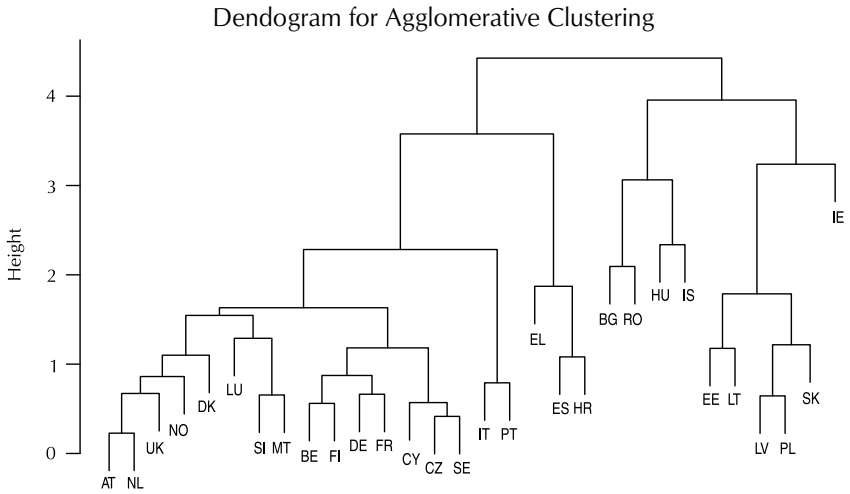
The choice of the European economies as the starting set for computing the estimate of steady state unemployment can be questioned on grounds of consistency with the data used to calibrate the steady state growth of potential output. The most straightforward approach would be to use unemployment data for the same economies that formed the potential output group. We chose to depart from that option in our implementation. The motivation for that choice is grounded in the view that, even though the evolution of many macroeconomic variables is history-dependent and institution-specific, this is particularly true for labour market outcomes. Whereas long-term growth will at best be constrained by the expansion of the technology frontier, hence the choice of a group which is arguably on or close to that frontier, this can be achieved under a variety of modes of functioning of the labour market. For this reason, we find it preferable to benchmark possible unemployment outcomes against economies that are – to the extent possible – historically and institutionally similar to Bulgaria. While this goal inevitably entails a trade-off between a larger but more heterogeneous sample and a smaller sample of close counterparts, we believe the European economies provide an acceptable compromise in that respect.

To implement the first part of our approach, we identified a set of European economies that are similar to Bulgaria in terms of historical developments in inflation, growth and unemployment. To this end, we applied clustering procedures to data for average growth, inflation (based on the overall harmonized index of consumer prices excluding energy, food, alcohol and tobacco) and unemployment for the European economies over the period 1996–2015. Clustering refers to a rich set of unsupervised learning procedures that aim to discover groups in a dataset. Groups should include "similar" data points where similarity (or, equivalently, dissimilarity) can be formally defined on the basis of a chosen metric. Such metrics commonly include proximity and distance measures. An accessible overview of the main ideas can be found in James et al. (2013), chapter 10, while Gan et al. (2007) and Kaufman and Rousseeuw (2005) offer specialized treatments of the subject.

Proximity in our case was defined via the Euclidean distance without weighting, on standardised data. To test the robustness of our results, we

used both an agglomerative[3] procedure and a divisive[4] one. Both procedures are representatives of the class of hierarchical clustering methods. In essence, hierarchical clustering methods do not produce a single partition of the data into groups but instead yield a set of nested partitions constructed for varying degrees of strictness in the application of the similarity criterion. Depending on whether the partitioning is carried out bottom-up (starting from individual observations and successively binding them together in larger groups) or top-down (starting from one group and successively subdividing it into smaller groups) we speak of agglomerative or divisive procedures, respectively. The results of both methods indicate that Bulgaria can be grouped together mostly with Eastern European countries. These results can be presented graphically via a tree-like structure called a dendrogram (see Figure 1 for the results of the agglomerative procedure and Figure 2 for the results of the divisive one), which shows the various groups as a function of the similarity criterion. The coverage of the group obviously depends on the strictness of the similarity definition and in this case we opted for a broader one comprising 10 countries[5] in order to get more representative results.

Figure 1: **Cluster Analysis Results for Agglomerative Clustering**

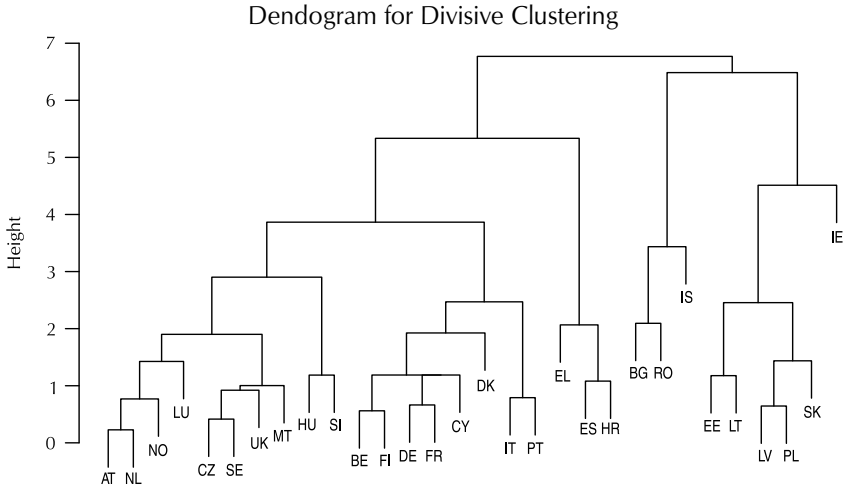


[3] Algorithm AGNES as implemented in R package *cluster*. See also Kaufman and Rousseeuw (2005), chapter 5.

[4] Algorithm DIANA as implemented in R package *cluster*. See also Kaufman and Rousseeuw (2005), chapter 6.

[5] These are BG, RO, IS, EE, LT, LV, PL, SK, IE and HU.

Figure 2: **Cluster Analysis Results for Divisive Clustering**



The direct computation of the mean unemployment rate for the selected set of economies has the potential drawback that the average unemployment rate for each economy may reflect an out-of-equilibrium position due to cyclical developments. To control for that possibility, we computed estimates of the output gap on quarterly data for these economies using a simple HP filtering procedure and isolated the episodes when output was sufficiently close to its potential (defined as an absolute value of the estimated output gap less than or equal to 0.2%[6]). We then averaged the corresponding unemployment rates over the identified subsamples. In this procedure all countries and observations were treated symmetrically and, in particular, no weighting scheme of any kind was applied. The resulting mean unemployment rate is a measure providing a picture of the typical historical state of unemployment when the output was close to equilibrium.

It can also be argued that the calibration of equilibrium unemployment should not be entirely backward looking but should also reflect what may be considered achievable in the future. This consideration motivates the second part of our overall approach. To this end, we introduced a measure of what the "best" outcome in terms of unemployment might be. In order to keep this approach empirically grounded, we computed the lowest historical unemployment rates for the full sample of EU countries over the

[6] While the choice of the specific threshold value is essentially arbitrary, we experimented with other values in the range 0.1–0.5% to confirm the robustness of the results.

period 1996–2015 and averaged them. While the observed minima may have been attained under specific circumstances and, in particular, at a time when the respective economy was overheating, a computation of this kind shows the historical limits reached by the respective labour markets.

To construct an estimate of equilibrium unemployment that combines a historically representative quantity with the potential "best" outcome, we obtained the final calibration value by taking the mean of the average unemployment rate computed for small output gaps and the average of the historical unemployment minima. The equilibrium unemployment rate was thus calibrated at 7.4%.

Estimation Strategy and Data

For the estimation of the model we use a more general methodology than the one used by Benes et al. Once we set up the priors of the individual parameters and locate the posterior mode, we use an adaptive random-walk Metropolis-Hastings algorithm to obtain the whole distributions of the parameters.[7] For most parameters we use the lognormal distribution and only for a few shocks in the unemployment block the inverse gamma distribution.

We estimate the model over the sample of 1999Q1–2015Q4. The series of GDP used in the model are the ones provided by the National Statistical Institute (as of March 2016) and seasonally adjusted using Bayesian methods by the Bulgarian National Bank. The measure of inflation used in the model is the HICP excluding energy goods, unprocessed food, tobacco and administrative prices. For the measure of import price, we use competitors' prices on the import side relevant for Bulgaria's main trading partners and for the international oil price we take the Brent crude oil price. The real effective exchange rate is taken from the ECB database[8]. Unemployment is taken from the Labour Force Survey and the measure used is for persons over the age of 15, seasonally adjusted by the Eurostat. Alternative versions of the model were tested with core inflation excluding energy, food, tobacco and administrative prices as well as international food prices instead of the competitors' prices on the import side. However, these alternative specifications did not yield satisfying results.

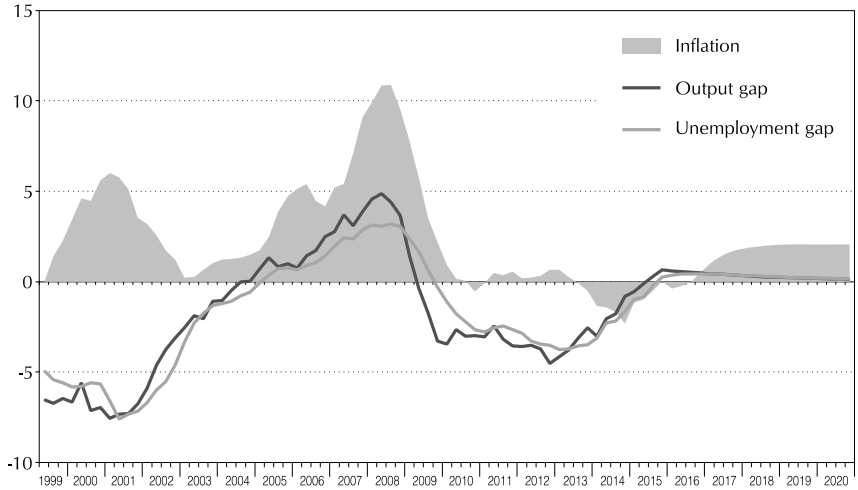
[7] Estimation was done with the Matlab toolbox IRIS.

[8] The series used in the paper is the GDP-deflators deflated effective exchange rate (EER) of the Bulgarian lev against the currencies of 19 countries.

Results

The estimates show that the trade-off between economic activity and inflation is not as clear-cut as theory suggests which is mainly due to one-off and country-specific factors. There is a large negative output gap in the period 1999Q1–2003Q1, which corresponds to the high unemployment in this period. During the same period there is also high inflation, which on the one hand can be related to the transition to market economy the country went through in terms of opening up borders to trade and privatization and on the other the period of hyperinflation the country emerged from in the end of the 90s with the introduction of the currency board. In the period 2003Q2–2008Q4 the output gap is positive and increasing and peaks in 2008Q1 (4.5%), which coincides with the peak in inflation. Throughout that period unemployment is also steadily decreasing. With the global financial crisis and the negative shock in demand, output gap turns negative in 2009Q1 and reaches a local trough in 2009Q4 (-4.5%). Coinciding with the European sovereign debt crisis there is another local trough in 2012Q4 (-4.6%). The negative output gap observed after 2009 is estimated to close in 2014Q4. Afterwards the output gap is estimated and projected to fluctuate close to zero as inflation goes up to the hard anchor of 2% annually. In general, inflation tends to exhibit stronger persistence and lag behind the output gap historically.

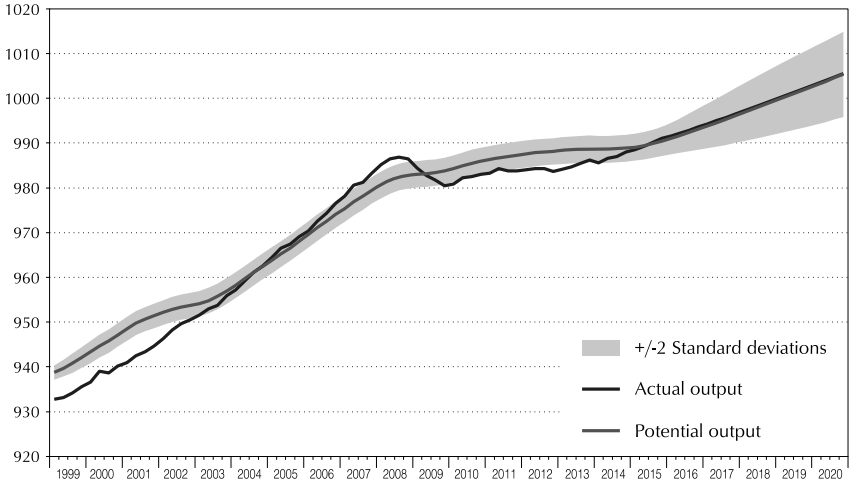
Figure 3: **Output Gap, Inflation and Unemploymen Gap**
(in percent)



Source: Authors' calculations.

Assessing the dynamics of potential output, we see that there is a significant deceleration in potential output growth in the period 2000Q4–2003Q1, which has the main contribution for the closing of the output gap in 2003. This significant deceleration can be associated with the significant increase in equilibrium unemployment. Until 2009 and the global financial crisis potential output growth is steady of around 5%. Afterwards another deceleration takes place, but much smoother than the one in the beginning of the sample. Again, this deceleration contributes to the faster closing of the negative output gap in the end of 2014. Assessing the relative variance of changes in the output gap to changes in actual output, we see that changes in the output gap dominate short-term changes in GDP (see Appendix 2). This confirms the view, which is taken as conventional for industrialized countries that output gap is more important for short-term actual output movements than the potential output.

Figure 4: **Actual and Potential Output**
(in logarithm)

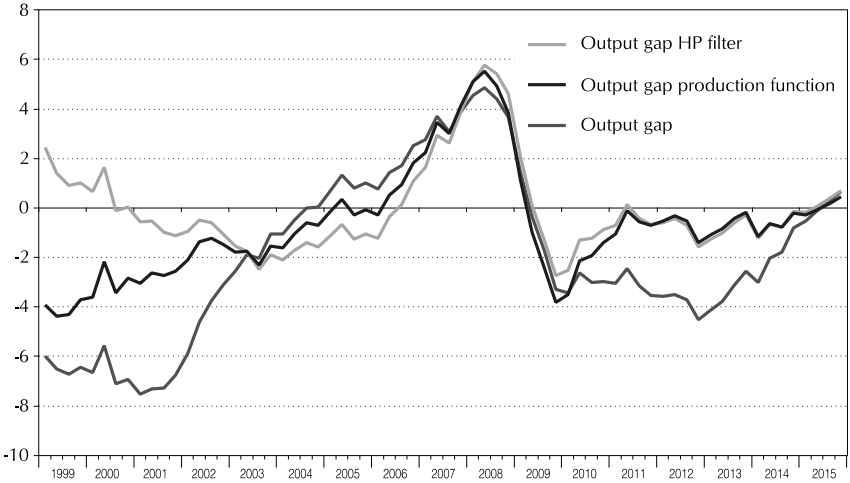


Source: Authors' calculations.

Comparing the estimated output gap with other measures obtained from alternative methods like a simple Hodrick-Prescott filter and a production function approach, we find some significant differences. The two alternative measures of output gap fluctuates much closer to zero than the one estimated by the multivariate filter model, where most of the movement is in negative territory. In the period 1999Q1–2003Q1, the output gap

estimated by the production function approach is much less negative and positive in one quarter, while the one obtained from the HP filter is positive most of the time. In the period 2003Q2–2008Q4 the three estimates are quite close to each other, even though the HP filter and the production function approach show the output gap turning positive much later than the multivariate filter estimate around 2006. During the crisis, the two alternative measures estimate a closing of the negative output gap in 2011Q1, when actual output growth accelerates, which is much sooner than the results from the multivariate filter. The results from the multivariate approach are more consistent with the undergoing structural change of the economy in 1999–2003 and macroeconomic developments in the aftermath of the global financial crisis. This comparison shows that the use of additional information about inflation and labour market developments improves on univariate methods as the Hodrick-Prescott (HP) filter.

Figure 5: **Various Measures of Output Gap**
(in percent)



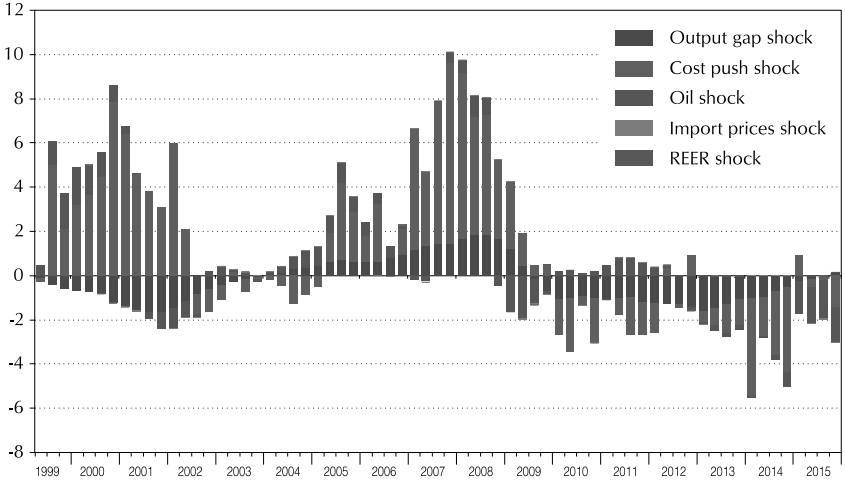
Source: Authors' calculations.

The estimated Phillips curve displayed a low sensitivity of core inflation to changes in the output gap, but a relatively higher sensitivity to speed-limit effects (see Appendix 1). Lagged inflation has a high coefficient in the Phillips curve equation, while expectations have a relatively low impact on inflation. Oil price shocks affect inflation mostly contemporaneously, while other import prices shocks have an effect with a lag. The effect of real exchange rate shocks on inflation is rather small. The coefficients remain

stable when the model is estimated over different sub-samples. The core parameters of the model, like the steady-state growth rate, were estimated within the model and the results were close to the initial calibrations. More graphs of the estimated potential output and NAIRU are given in Appendix 2.

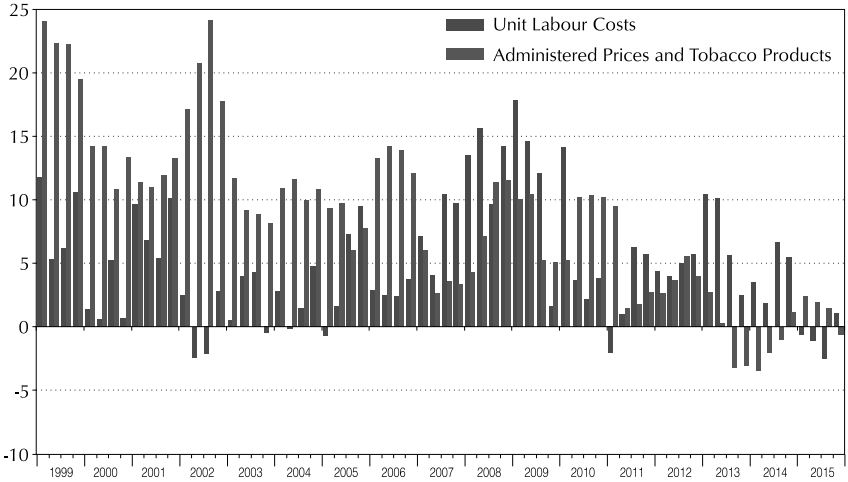
The contribution of shocks to inflation dynamics reveals that in the periods of very high inflation, the factors incorporated in the Phillips curve do not explain most of the variation in inflation. Therefore, factors not explicitly included in the model should account for that. We see the output gap shock has been systematically important for inflation dynamics throughout the sample, matching the sign of the output gap. However, the size of the contribution corresponds to the low sensitivity of inflation to economic slack. The cost push shock, which is the shock in the Phillips curve equation, has the main contribution in the three periods of very high inflation: 1999Q2–2002Q2, 2003Q2–2004Q3 and 2005Q1–2009Q2. As it was mentioned above the earliest spell of high inflation in the sample can probably be associated with the process of transition to a market economy and opening up to trade. However, the most important factors that we believe stand behind the spikes in inflation and especially the one during the pre-2009 period are administrative price changes, strong growth in unit labour costs and possibly the appreciation of housing prices (see Figure 7). It is evident that for a prolonged period between 1999Q1 and 2003Q3 administered prices grow by a double digit growth. Some of these administered prices are the prices of utilities like electricity and gas, while others are related to public services and healthcare. On the other hand, unit labour costs growth accelerates roughly in the period 2007Q1–2009Q3. During the pre-crisis period of strong demand growth, producers were able to pass on increased production costs to final consumers.

Figure 6: **Contribution of Shocks to Inflation Dynamics**
(in percentage points)



Source: Authors' calculations.

Figure 7: **Administered Prices and Unit Labour Costs**
(in annual percent changes)



Source: NSI.

On the other hand, in the period from 2009 and after it seems that the factors included in the model do manage to explain most of the inflation dynamics. In the current period of deflation demand has no contribution to the inflation dynamics as the estimated output gap closes and fluctuates around zero. On the other hand the cost push shock has a significant negative contribution, but that can be attributed to the cuts in the prices of electricity since the middle of 2013. The other major contribution to the decline in inflation is the dynamics of the international oil price since 2014Q4. In general, it seems that the Phillips curve does explain inflation better after 2009, whereas earlier a few other important factors are at play, which relate to the large portion of administered prices in the HICP basket and the processes of nominal and real convergence of the economy.

The unemployment gap seems to move close to the output gap with a lag. Its dynamics is determined by the movement of equilibrium unemployment (NAIRU), which more or less follows actual unemployment with a lag. Bigger increases in equilibrium unemployment occur roughly in the period 2001–2003, as actual unemployment spikes to almost 19% and again after 2009. Decomposing the dynamics of the equilibrium unemployment to the contribution of the shocks in the model reveals that output gap is by far the most important factor. The estimates are robust to changes in the structural unemployment parameter, which directly affects the NAIRU. Since the main focus of our model framework is the interplay between output and inflation, the results are somewhat limited in regards to unemployment. Therefore, it is hard to draw much more detailed conclusions in regards to the equilibrium unemployment and the unemployment gap. This is a reason to consider an extension of the model specification, by including indicators specific for the labour market, in order to address this issue of an undeveloped, but important part of the model.

Conclusions

The paper uses a macroeconomic model estimated with Bayesian methods and using Kalman filter to estimate a Phillips curve type relationship between output and inflation as well as produce estimates for output and unemployment gaps. The model is also used to assess the main drivers of inflation in the past decade and a half.

Modelling inflation and potential output within the theoretical framework of the New Keynesian Phillips curve yields useful results for Bulgaria. Prior views on the main drivers of inflation were confirmed in this study.

Estimates show that core inflation is highly persistent and dependent on supply-side shocks like oil prices and other import prices. Even though inflation is not very sensitive to economic slack, changes in the output gap do have a non-negligible contribution to the inflation dynamics in the sample. Before 2009, in periods of high inflation, much of the dynamics of inflation are unexplained by factors explicitly included in the Phillips curve, being driven by cost push shocks of unspecified nature. These cost push shocks are challenging to identify precisely but we conjecture that they combine the effects of spill-overs from administrative price changes to core inflation, the dynamics of labour costs and dynamics of housing prices and their effects on consumer price inflation.

However, in our view the Phillips curve does explain inflation dynamics after 2009 well, which points out that the effects of other factors like labour costs and administrative prices have been weaker lately. Also the estimates of the output and unemployment gaps seem better aligned with macroeconomic developments than the ones obtained from alternative approaches, as they incorporate information from several time series. This could be useful for medium and long-term growth forecasting, in addition to uses related to the analysis of the cyclical position of the economy. Results from such forecasting exercises can also be used in order to formulate expectations about indicators relevant to economic policy like the cyclically-adjusted budget deficit and the gross government debt. Overall, we consider the results plausible and the framework employed here merits further development both for research and for applied uses. Some further extension of the model specification and work done on parameterization is deemed necessary in order to make the model fully suitable for regular applied use and forecasting.

References

Apel, M. and P. Jansson (1999), "A theory-consistent system approach for estimating potential output and the NAIRU", Economics Letters No. 64, 271–275.

Balakrishnan, R. and J. López-Salido (2002), "Understanding UK Inflation: the Role of Openness", Bank of England Working paper No. 164.

Batini, N., J. Brian and N. Stephen (2005), "An Open-Economy New Keynesian Phillips Curve for the U.K.", Journal of Monetary Economics 52 (September), pp. 1061–1071.

Benes J., K. Clinton, R. Garcia-Saltos, M. Johnson, D. Laxton, P. Manchev and T. Matheson (2010), "Estimating Potential Output with a Multivariate Filter", IMF Working Paper, WP/10/285.

Benes J. and P. N'Diaye (2004), "A Multivariate Filter for Measuring Potential Output and the NAIRU: Application to the Czech Republic", IMF Working Pape, WP/04/45.

Bokan, N. and R. Ravnik (2012), "Estimating Potential Output in the Republic of Croatia Using a Multivariate Filter", Croatian National Bank Working Papers W-35.

Bulíř A., M. Čihák and K. Šmídková (2008), "Writing Clearly: ECB's Monetary Policy Communication", IMF Working Paper, WP/08/252.

Calvo, G. (1983), "Staggered Price in a Utility Maximizing Framework", Journal of Monetary Economics, Vol 12: 383–98.

Fabiani, S. and R. Mestre (2001), "A System Approach for Measuring Euro Area NAIRU", ECB Working Paper Series No. 65.

Friedman, M. (1968), "The Role of Monetary Policy", American Economic Review No. 58, pp. 1–17.

Fuhrer, J., Y. Kodrzycki, J. Little and G. Olivei (2009), "Understanding Inflation and the Implications for Monetary Policy: A Phillips Curve Retrospective", MIT Press.

Gali, J. and M. Gertler (1999), "Inflation dynamics: A structural econometric analysis", Journal of Monetary Economics Vol 44, pp. 195–222.

Gali, J. and J. López-Salido (2001), "A New Phillips Curve for Spain", BIS Working Paper No. 3.

Gan, G., C. Ma and J. Wu (2007), "Data Clustering: Theory, Algorithms, and Applications", ASA and SIAM: Pennsylvania.

Gordon, R. (1998), "Foundations of the Goldilocks Economy: Supply Shocks and the Time-Varying NAIRU," Brookings Papers on Economic Activity Vol 29(2), pp. 297–346.

Gordon, R. (2009), "The History of the Phillips Curve: Consensus and Bifurcation", Northwestern University, NBER, and CEPR.

International Monetary Fund (2015), "Regional Economic Issues: Central, Eastern, and Southeastern Europe", May 2015.

Iossifov, P. and J. Podpiera (2014), "Are Non-Euro Area EU Countries Importing Low Inflation from the Euro Area?", IMF Working paper WP/14/191.

James, G., D. Witten, T. Hastie and R. Tibshirani (2013), "An Introduction to Statistical Learning with Applications in R", Springer: New York.

Kaufman, L. and P.J. Rousseeuw (2005), "Finding Groups in Data: An Introduction to Cluster Analysis", Wiley-Interscience: New Jersey.

Ljung, L. (1999), "System Identification: Theory for the User", Prentice-Hall: New Jersey.

Phelps, E. (1978), "Commodity-Supply Shock and Full-Employment Monetary Policy," Journal of Money, Credit, and Banking Vol. 10 (May), pp. 206–221.

Phillips, A. (1958), "The Relation between Unemployment and the Rate of Change of Money Wage Rates in the United Kingdom, 1861–1957," Economica 25: November, no. 100, pp. 283–299.

Rotemberg, J. (1982), "Sticky Prices in the United States", Journal of Political Economy 90(6): pp. 1187–1211.

Rotemberg, J. (1983), "Aggregate Consequences of Fixed Costs of Price Adjustment", American Economic Review 73(3), pp. 433–436.

Samuelson, P. and R. Solow (1960), "Analytical Aspects of Anti-Inflation Policy", American Economic Review Papers and Proceedings 50: May, no. 2, 177–94.

Sramkova, L. (2010), "Output Gap and NAIRU Estimates within State-Space Framework: An Application to Slovakia", Financial Policy Institute to the Ministry of Finance of the Slovak Republic, Economic Analysis Paper No 16.

APPENDIX 1

Model Specification and Parameters

Model Parameters

Calibrated Parameters	Value
$G_{ss}^{\overline{Y}}$	3.1
$\overline{U}^{ss}$	7.4
π_{ss}	2.0

Bayesian Parameter Estimation

Parameter	Prior					Posterior	
	Lower Bound	Upper Bound	Log Distribution	Mean	Standard deviation	Mode	Standard deviation
α	0.1	1	normal	0.89	0.10	0.87	0.10
β	0	1	normal	0.10	0.02	0.09	0.01
Ω	0	1	normal	0.20	0.42	0.13	0.30
η_1	0	1	normal	0.80	0.40	0.63	0.00
η_2	0.01	0.15	normal	0.10	0.20	0.12	0.00
ρ	0	1	normal	0.60	0.10	0.93	0.05
ϕ_1	0	1	normal	0.57	0.06	0.66	0.06
ϕ_2	0	1	normal	0.38	0.06	0.32	0.05
τ	0	1	normal	0.50	0.09	0.28	0.08
ω	0	100	normal	3.70	1.58	4.28	1.08
λ	0.01	100	normal	5.00	0.32	4.94	0.32
γ_0^l	0	1	normal	0.30	0.32	0.20	0.48
γ_1^l	0	1	normal	0.10	0.22	0.15	0.34
γ_2^l	0	1	normal	0.10	0.32	0.19	0.02
γ_0^m	0	1	normal	0.10	0.20	0.00	0.91
γ_1^m	0	1	normal	0.16	0.20	0.00	0.38
γ_2^m	0	1	normal	0.10	0.20	0.01	0.02
γ_0^s	-1	0	normal	-0.10	0.32	0.00	1.02
γ_1^s	-1	0	normal	-0.10	0.32	-0.01	0.31
γ_2^s	-1	0	normal	-0.10	0.32	0.00	1.01

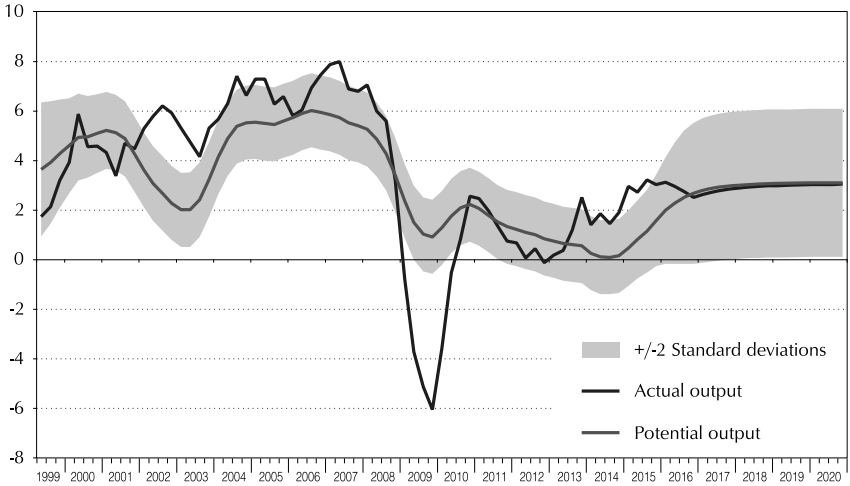
σ_{ε^y}	0.01	3	normal	0.99	0.01	1.39	0.01
$\sigma_{\varepsilon^{G\overline{Y}}}$	0.01	3	normal	0.66	0.13	1.16	0.11
σ_{ε^u}	0	3	normal	0.67	0.06	0.47	0.08
$\sigma_{\varepsilon^{\overline{U}}}$	0.01	3	invgamma	0.30	0.30	0.15	0.04
$\sigma_{\varepsilon^{G\overline{U}}}$	0	3	invgamma	0.30	0.30	0.17	0.05
σ_{ε^π}	0.01	3	normal	0.74	0.03	1.17	0.04
$\sigma_{\varepsilon^{\overline{Y}}}$	0.01	3	normal	0.21	0.08	0.26	0.08
σ_{ε^l}	0.01	3	normal	0.80	0.36	1.00	0.10
σ_{ε^m}	0.01	3	normal	0.99	0.36	1.00	0.11
σ_{ε^s}	0.01	3	normal	1.00	0.36	0.12	0.01

APPENDIX 2

Estimates of Unobservable Variables and Additional Model Output

Actual and Potential Output Growth

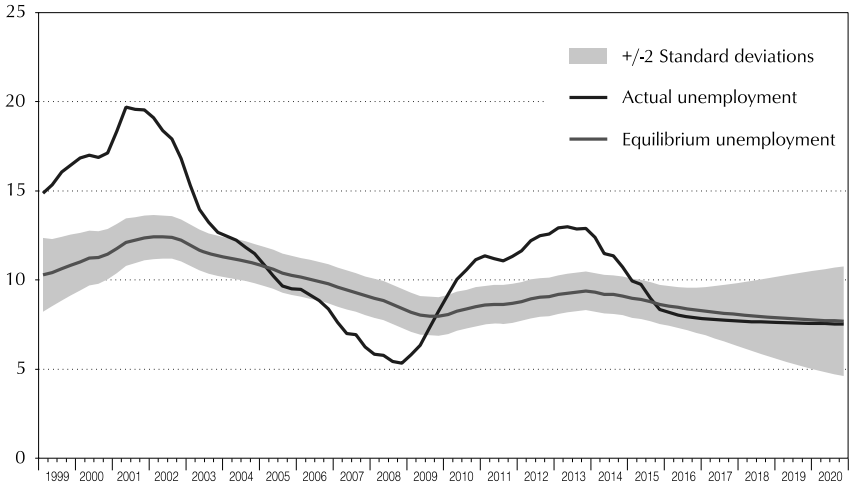
(in annual percent changes)



Source: Authors' calculations.

Actual and Equilibrium Unemployment

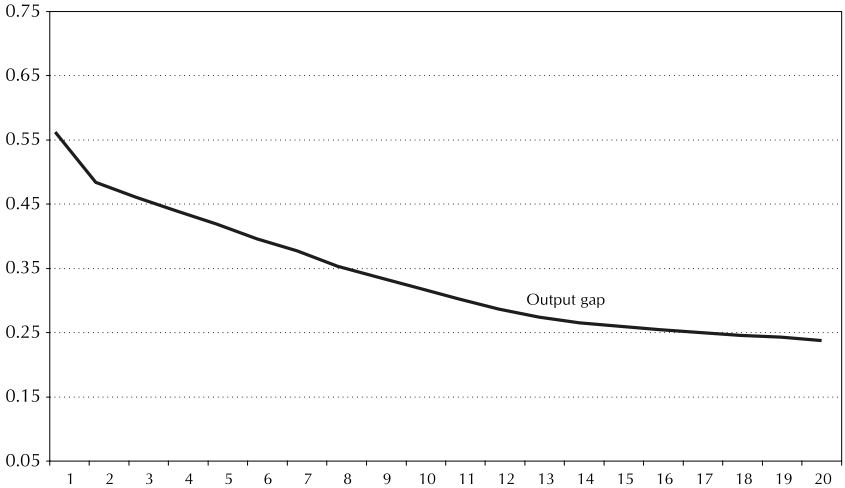
(in percent)



Source: Authors' calculations.

Variance of the Changes in Output Gap Relative to Variance of Changes in Actual Output Over One to Twenty Quarters

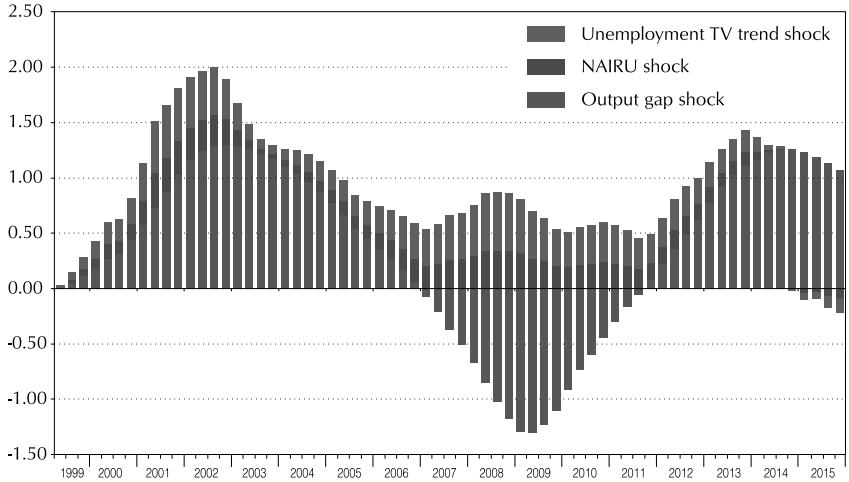
(ratio)



Source: Authors' calculations.

Contribution of Shocks to Equilibrium Unemployment Dynamics

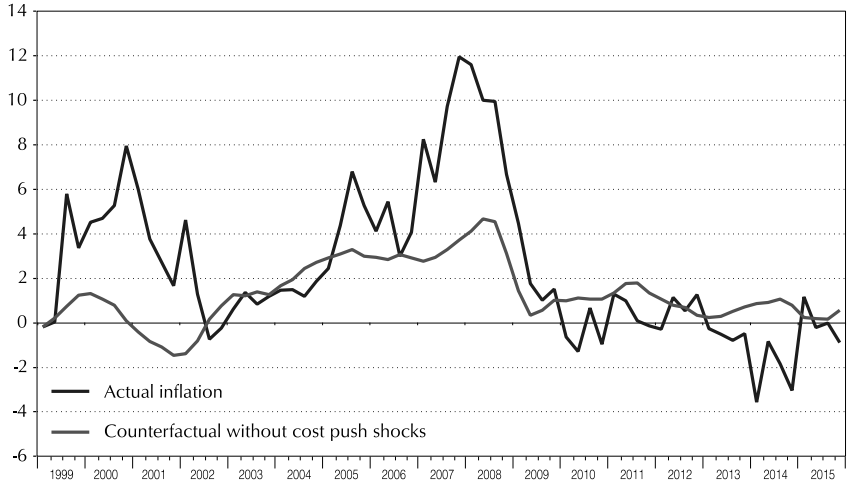
(in percentage points)



Source: Authors' calculations.

Actual and Counterfactual Inflation without the Cost Push Shocks

(in annual percent changes)



Source: Authors' calculations.

DP/105/2017

The Role of Bulgaria in Global Value Chains

Neli Ivanova, Evgeni Ivanov

Abstract. Considering the increasingly fragmented production across different countries, the high degree of trade openness of Bulgaria and the significant role of exports for the country's overall real GDP growth, the aim of this paper is to analyse the importance of participation in Global Value Chains (GVCs) for Bulgaria relative to a set of selected economies. The analysis is based on three main indicators, constructed from international input-output databases: participation in GVCs index, length of GVCs index and distance to final demand index. Our results suggest that Bulgaria is deeply integrated in GVCs mainly through manufacturing activities such as petrol refining, the production of basic metals and the production of machinery, electrical and transport equipment. Bulgaria participates predominantly in highly fragmented GVCs and specialises in processing and assembly functions, which is why the country's participation in GVCs is characterised by relatively low domestic value added content and an intense usage of foreign inputs. We also find that the products with which Bulgaria participates in GVCs are predominantly inputs rather than final products, which places them further away from end-users. The article also concludes that, although, services play a less significant role for Bulgaria's exports compared to regional peer countries, some fast-growing manufacturing sectors have considerably increased their usage of services inputs over time, in line with these sectors' increased participation in GVCs. Our results suggest that Bulgaria has managed to benefit from its participation in GVCs. Despite the significant share of foreign value added content of Bulgarian exports, the share of domestic value added trade to GDP is quite high due to the overall importance of exports relative to the size of the Bulgarian economy.

Neli Ivanova, Economic Research and Forecasting Directorate, Bulgarian National Bank, ivanova.nl@bnbank.org.

Evgeni Ivanov, Economic Research and Forecasting Directorate, Bulgarian National Bank, ivanov.e@bnbank.org.

Acknowledgment:

We would like to thank the members of the editorial board of the Bulgarian National Bank and particularly Mariella Nenova, Ass. Prof., Ph. D., and Andrey Vassilev, Ph.D., for their valuable contribution, comments and suggestions.

Резюме. Съобразявайки се с все по-фрагментираното производство в глобален аспект, с високата степен на отвореност на българската търговия и със значителната роля на износа на страната за общия растеж на БВП в реално изражение, целта на това изследване е да анализира доколко за България е важно да участва в световните вериги за създаване на стойност (Global Value Chains, GVCs) в сравнение с група подбрани икономики. Анализът се базира на три основни индикатора, конструирани от международни таблици от вида „ресурс – използване“: индекс на участие в GVCs, индекс на дължина на GVCs и индекс за отстояние от крайното търсене. Нашите резултати показват, че България е дълбоко интегрирана в GVCs главно чрез сектори на преработващата си промишленост, като нефтопреработка, производството на цветни метали машиностроене и производство на електрическо и транспортно оборудване. България участва предимно в силно фрагментирани GVCs и се специализира в дейности по обработка и сглобяване, което показва защо участието на страната в GVCs се характеризира със сравнително малко съдържание на местна добавена стойност и с интензивно използване на чуждестранни ресурси. България участва в GVCs по-скоро с междинни стоки, отколкото с крайни продукти, което отдалечава страната от крайните потребители. Изследването прави също извод, че макар услугите да играят по-маловажна роля в българския износ, отколкото в сравними държави от региона, някои бързо развиващи се сектори на преработващата промишленост значително са увеличили използването на услуги като входящ ресурс за производството си в съответствие със задълбочилото се участие на тези сектори в GVCs. Нашите резултати показват, че България успява да се възползва от своето участие в GVCs. Въпреки значителния дял на чуждестранната добавена стойност в българския износ, обемът на търговията, съдържаща местна добавена стойност, спрямо БВП е висок, тъй като съотношението на общия износ спрямо размера на българската икономика е високо.

Contents

1. Introduction .....

2. Literature Review .....

3. Data and Methodology .....
   3.1. Data .....
   3.2. Methodology .....

4. Factors to Consider When Interpreting the Results of GVCs Indicators .....

5. Structure of Bulgaria's Exports .....

6. GVCs-specific Indicators .....
   6.1. Participation Index .....
   6.2. Length of Global Value Chains Index .....
   6.3. Distance to Final Demand Index .....

7. The Role of Services .....

8. Impact of Participation in GVCs on Bulgaria .....

9. Conclusion .....

Bibliography .....

1. Introduction

The production of goods has become increasingly fragmented across countries. Due to falling trade costs, greater global openness, cooperation on trade policy and the information and communications technologies (ICT) revolution, production process has become increasingly unbundled across countries (OECD, 2013). Goods and services that were once produced in a single country have become part of global value chains (GVCs). A global value chain involves all the activities that firms engage in, at home or abroad, to bring a product to the market, from conception to final use (OECD, 2013). This means that GVCs encompass the entire process required to convert raw materials, labour, capital and knowledge into intermediate products and final goods. Production is divided into many small stages of specialisation along the chain that can be carried out where the necessary inputs are most efficiently sourced, thus achieving an increase in productivity and competitiveness. In this setup, countries tend to specialise in specific business functions rather than specific industries, such as the assembly operations for China or business services for India (OECD, 2013). Typically, a value chain includes activities like design, production, marketing, distribution and support to the final consumer, which can be performed within the same firm or divided among different firms (De Backer and Miroudot, 2014). In GVCs firms control and coordinate activities in networks of buyers and suppliers, and multinational enterprises (MNEs) play a central role (OECD, 2013). Not all GVCs are equally complex or widespread. The level of fragmentation of production is linked to the technical characteristics of products and the costs incurred when production is split among different locations (OECD, 2013).

Given the way GVCs are defined, every country in the world is to some extent part of GVCs, since it uses at least some imported inputs in its exports, or at least some of its exports end up as inputs in third countries' exports. In theory, however, a case of a country that does not participate in GVCs would be where the given country sources all inputs for its exports domestically and these exports are used only for final consumption by the rest of the world.

A country's high participation in GVCs does not simply mean that that the country has a large number of multinational enterprises (MNEs) operating on its market. Despite the crucial role of MNEs, GVCs can theoretically operate without them since the supply chains also include independent suppliers, such as domestic and foreign small and medium-sized enterprises. However, in practice this is rarely the case since investment and

trade are closely related – UNCTAD estimates that around 80% of global trade in terms of gross exports is linked to the international production networks of MNEs, either as intra-firm trade, or through arm's length transactions (UNCTAD, 2013). This is why usually countries with higher presence of FDI relative to the size of their economies tend to have higher levels of participation in GVCs and greater relative shares in global value added trade compared to their shares in global exports (UNCTAD, 2013).

Considering the growing importance of international production fragmentation, the high degree of trade openness of the Bulgarian economy and the significant role of exports for the overall real GDP growth in the country, the aim of this paper is to examine Bulgaria's participation in GVCs and to identify the production sectors with highest participation rates. The paper also aims at identifying the specific characteristics of the GVCs in which Bulgaria participates, such as the degree of production fragmentation, as well as the types of the products with which Bulgaria participates in these GVCs (whether Bulgaria participates in GVCs through the production of inputs or final products). The analysis tries to answer the question whether Bulgaria has managed to extract sufficient welfare gains from its participation in GVCs. In order to put our results in perspective and to be able to draw conclusions from our analysis we compare the results for Bulgaria to a pre-selected reference group which covers several types of economies, representative of the different specialisations in GVCs.

Our initial expectations for the results are that Bulgaria would be well-integrated in GVCs, given that it is a small, open economy, which is also a member of the European Union (EU). The structure of gross exports and imports of the country and the importance of primary commodities (such as petrol and metals), as well as the country's low endowment of natural resources, suggest that Bulgaria would focus more on processing and assembly functions in relatively more fragmented GVCs, and its products would likely be positioned further away from end-use consumers, compared to more developed economies.

The contribution of this paper is that it applies a structured empirical analysis that uses indicators based on international input-output tables in order to assess the importance of GVCs participation for Bulgaria, relative to a set of pre-selected economies. To our knowledge this paper is the first attempt in existing literature to apply such type of detailed analysis for the case of Bulgaria.

The rest of the paper is structured as follows: Section 2 gives a short overview of existing studies that are relevant to our topic. Section 3 then pre-

sents the data used for our analysis, as well as the methodological notes on the construction of the used indicators. Section 4 outlines some factors that can affect the interpretation of these GVCs indicators. Section 5 provides an overview of the structure of Bulgarian exports. Section 6 focuses on presenting and discussing the results from the GVCs-specific indicators for Bulgaria. Then we have another section that discusses in more details the role of services for Bulgarian exports, followed by a section that tries to give an intuition of how participation in GVCs has affected the Bulgarian economy. The paper ends with concluding remarks.

2. Literature Review

The fragmentation of production across countries is not a new phenomenon. What is new is its increasing scale and scope. The concept of GVCs can be traced back to the end of the 1970s with some work on the „commodity chain" (De Backer and Miroudot, 2014), the focus of which was to trace all the sets of inputs and transformations that lead to an „ultimate consumable" and to describe a linked set of processes that add up to this final product (Hopkins and Wallerstein, 1977). The concept of „global commodity chain" was later introduced in the work of Gary Gereffi (1994), describing as an example the apparel commodity chain, from the raw materials (such as cotton, wool or synthetic fibres) to the final products (garments). In the 2000s, there was a shift in terminology from „global commodity chain" to „global value chain", the latter coming from the analysis of trade and industrial organisation as a value added chain in the international business literature (Porter, 1985).

Since globalisation has entered a phase of radical reductions in international communication and co-ordination costs and firms can offshore many tasks that used to be considered as non-traded, the examination of countries' position within international production networks has become even more crucial (Baldwin, 2006). Studies on GVCs trade have been done at a firm level (Alfaro et. al (2016), Dedrick et al. (2009)), industry level (De Backer et al. (2014), Sturgeon et al (2008)), country level (OECD (2013), ITC (2014)) as well as regional level (Timmer et al. (2013), IDE (2011), ILO (2016)). Work on the role of countries in GVCs provides a framework for the analysis of the competitiveness position of countries, their trade patterns, as well as opportunities for market expansion.

GVCs trade can be measured in several ways such as firm surveys, customs statistics that record trade flows under special schemes of tariff reduction or exemption and the Standard International Trade Classifica-

tion (SITC) classifying goods as being intermediate or final (Park et al., 2013). However, due to the limitations associated with these methods, analysis based on input-output tables has become the preferred option. Using input-output tables Koopman et al. (2011) estimate that accounting for the use of exports as inputs into exports by another country is crucial for understanding the functioning of GVCs. They provide a framework for decomposing a country's gross exports into value added components by source. In the study all value added in a country's exports is distributed to its original sources, and individual sources and destinations of value added are expressed at either a country-wide or industry aggregate level. Hummels et al. (2001) use input-output tables for 14 countries and show that growth in vertical specialisation accounted for about one-third of the growth in overall exports in the period 1970–1990. They provide evidence that it is developing countries that source inputs from developed countries which are then transformed into exports for developed economies. Another study by Banga (2013) applies a value added analysis using harmonised world input-output tables and estimates countries' extent of participation in GVCs in terms of their share in total value added created by GVCs. The study shows that around 61 per cent of total value added created in GVCs is due to OECD countries.

Analysis of a country's participation in GVCs by sector is key for establishing a country's place in GVCs. Daudin et al (2011) show evidence that standard trade statistics do not give an accurate representation of the relative dependence of different sectors on international demand. Their analysis of 66 regions and 55 sectors shows that some sectors have more value added trade than export trade which implies that their products have been traded as inputs in other goods. Furthermore, Koopman et al. (2010) suggest that a country's revealed comparative advantage in a certain sector may change when using value added in exports as opposed to official trade data which suggests that accounting for GVCs could modify substantially our understanding of trade patterns and revealed comparative advantage. In terms of the services sector, GVCs studies prove that the importance of services for trade has been more than undermined by standard trade statistics. Timmer et al. (2013) show that with the increasing participation in GVCs, the importance of services has also increased which is associated with the fact that the production of final goods in many sectors is largely dependent on service activities.

The analysis of GVCs trade brings up the question of potential positive effects from participation in GVCs for countries. Though the answer to that question is largely dependent on the specific role of each country

in GVCs, high participation in GVCs is generally seen as beneficial for countries, since studies find a positive correlation between participation in GVCs and GDP per capita growth rates (UNCTAD, 2013). Furthermore, studies such as Rahman and Zhao (2013) and IMF (2013) find evidence that the direction of causality runs from increased participation in GVCs to expansion of the production possibility frontier of the economy and a consequent spur of overall employment and growth. Economies with the fastest growing GVCs participation have some 2 percentage points above the average GDP per capita growth rates (UNCTAD, 2013). In addition, participation in GVCs can also be used by developing countries to build up productive capacity, improving their prospects for longer-term industrial upgrading. Participation in GVCs is considered a factor facilitating the integration and sustainability of country's exports and a main channel for technology and knowledge transfer.

Having established the importance of understating the role of countries in GVCs, we focus on providing a framework for the analysis of the case of Bulgaria.

3. Data and Methodology

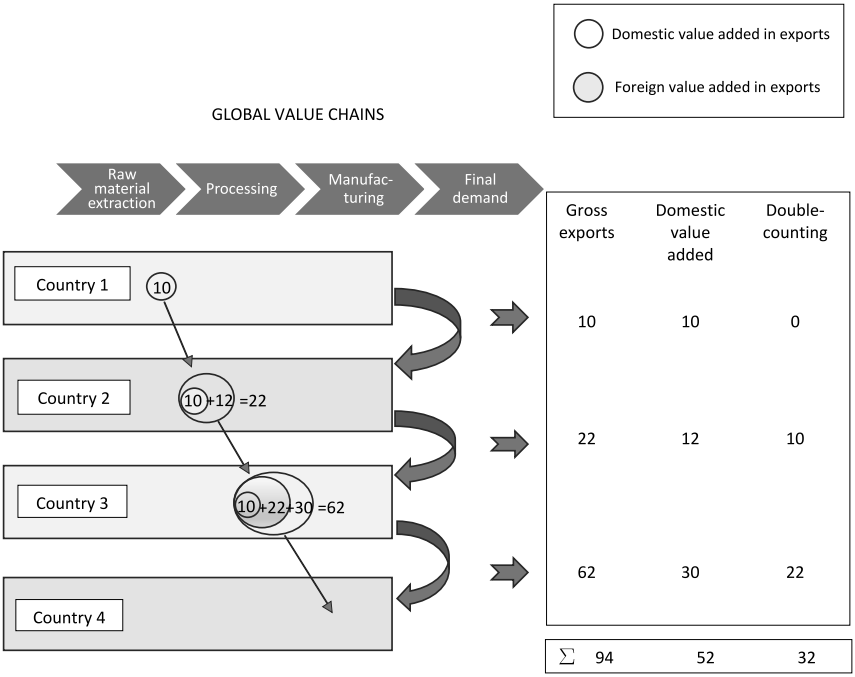
3.1. Data

With world trade and production being increasingly structured around GVCs, the sourcing and usage of inputs and components from multiple suppliers based in different countries challenges standard trade statistics (Antras et al., 2012). This is due to the fact that standard trade statistics are recorded in gross terms and as inputs cross borders multiple times, the value of intermediate inputs traded along the value chain is recorded several times (De Backer and Miroudot, 2014). In order to solve this problem, value added trade statistics can be used as it eliminates this double counting by showing where value is created along the production chain.

To illustrate the difference between standard trade statistics and value added statistics consider the following example (see Figure 1): Country 1 extracts iron ore and exports it for EUR 10 to Country 2, which processes it into steel and then exports it to Country 3. County 2 adds value of EUR 12 to the iron ore by processing it and so exports EUR 22 to Country 3. Conventional trade statistics show gross exports of EUR 22, while only EUR 12 of value added has been generated in Country 2's production. To illustrate further, consider that Country 3 uses the imported steel for production of safe deposit boxes, adding value of EUR 30 to the steel

and exports the safe deposit boxes to Country 4 for final use. At this stage standard statistics would register gross exports of EUR 62 even though only EUR 30 of value added was created in Country 3. In other words, conventional trade statistics can overstate the importance of exports for GDP and make it difficult to visualise the production chain (De Backer and Miroudot, 2014)

Figure 1: **Difference between Gross Statistics and Value Added Statistics**



Source: based on UNCTAD (2013).

This is the reason behind the initiative of the OECD together with WTO to build a database of trade flow in value added terms that is based on a global model of international production and trade networks. This global inter-country input-output (ICIO) model is created using national and harmonised input-output tables (IOTs), national supply and use tables (SUTs), bilateral trade in goods by industry and end-use category (BTDIxE).[1] The model is used for the creation of the Trade in Value Added database

[1] http://stats.oecd.org/Index.aspx?DataSetCode=TIVA2015_C1

(TiVA) which includes 61 economies and 34 industries in its October 2015 edition. The TiVA database can be used to extract useful information about the characteristics of GVCs, their development over time and the nature and impact of economic globalisation (OECD, 2015). Despite its usefulness, trade in value added statistics is still used as complementary to conventional trade statistics due to the significant delays in their publication and the low frequency of the data.

This analysis of the role of Bulgaria in GVCs is based on two separate OECD databases and is presented in a multi-country multi-industry framework. In particular, for most indicators related to GVCs that we have analysed we have used the October 2015 edition of the value- added database of OECD (TiVA). In the TiVA database data are available for the years 1995, 2000, 2005, 2008, 2009 and 2011. For two of the discussed indicators in the paper („length of GVCs" and „distance to final demand") we have used the May 2013 OECD Global Value Chains Indicators database and in order to establish possible trends we have examined the indicators for all available years (1995, 2000, 2005, 2008 and 2009).

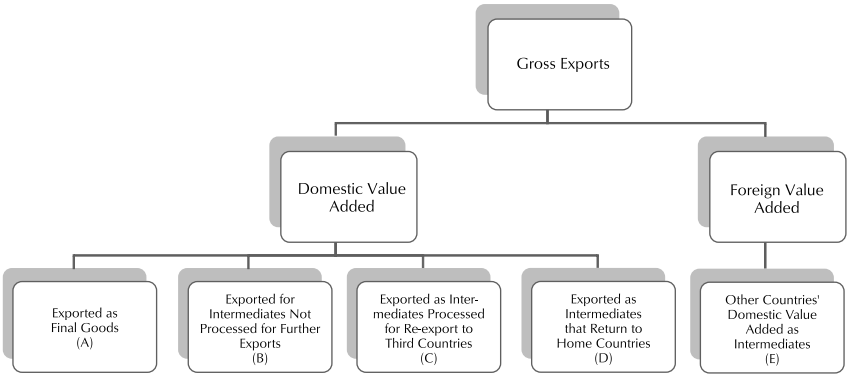
As our analysis is comparative we have chosen a reference group that covers the main types of economies in terms of different trade pattern in GVCs. The reference group includes economies with similar economic and historical development to Bulgaria (new Member States of the EU (NMS)[2]), main trading partners of Bulgaria and developed economies (the EU), natural resource exporters (Russia), large developed economies (the US), large developing economies (China) and technologically advanced economies with rather limited natural resources (Japan). In our analysis we emphasise on the comparison between Bulgaria, NMS and the EU as we believe this would allow us to draw the most relevant conclusions due to the economic similarities.

3.2. Methodology

In order to better understand the methodology used in this paper it is worth explaining in more details the way in which domestic and foreign value added are intertwined in gross exports. To achieve this goal we will graphically illustrate the decomposition of gross exports into its main five value added components.

[2] This group comprises of the Czech Republic, Estonia, Croatia, Latvia, Lithuania, Hungary, Poland, Romania, Slovenia and Slovakia.

Figure 2: **Decomposition of Gross Exports into Value Added Components**



Source: IMF (2013).

The first two components (A and B) comprise the exports that are independent from the country's participation in supply chains, whereas the rest of the components represent exports that are related to the country's participation in supply chains. The participation of an economy in GVCs can be divided into two components: upstream (components (C)-(D)), which include domestic value added that is processed for further exports, and downstream (component (E)). A large share of the foreign value added in total exports can be interpreted as a sign that a country is a downstream participant in the supply chain, usually specialising in processing and assembly functions (IMF, 2013).

In order to conduct a more in-depth analysis of the role of Bulgaria in GVCs we have used several GVC-specific indicators, namely "participation index", "length of GVCs" and "distance from final demand".

Participation Index.[3] The extent to which a country is engaged in global value chains is indicated by the "GVC participation index". Countries are involved in GVCs both as users of foreign inputs and as suppliers of intermediaries[4] in other countries' exports and both of these elements are reflected in the index (OECD, 2013). Moreover, the GVCs participation index as developed by Koopman et al. (2010) is measured as a percentage of foreign inputs in gross exports (backward participation) and domesti-

[3] Calculations are made based on the TiVA database updated as of October 2015. Latest available data entry in the database is for 2011.

[4] We define intermediate products as goods and services consumed as inputs by a process of production, excluding fixed assets.

cally produced inputs used in third countries' exports (forward participation) (De Backer and Miroudot, 2014). In formal terms:

$$\text{Participation index} = \frac{IV_{ik}}{E_{ik}} + \frac{FV_{ik}}{E_{ik}}, \text{ for country } i \text{ and industry } k$$

where ***IV*** reflects the domestic value added embodied in third countries' gross exports (forward participation), ***FV*** indicates the foreign value added embodied in gross exports (backward participation) and ***E*** stands for the country ***i***'s gross exports.

In its essence, the indicator measures both the import content (foreign value added content) of exports and the share of exports made of domestic value added used in third countries to produce other exports. The higher each one of these components is, the higher the participation of a country in GVCs. A larger value of the index shows greater integration in GVCs which, as indicated earlier, is considered beneficial for an economy.

Length of Global Value Chains.[5] A complete understanding of how a country integrates into the production networks requires more than just looking at the participation index as a high ***IV*** share could correspond to the use of expensive raw materials in a simple chain, whereas a high ***FV*** share could be added at one go at the final stage of the production process. While the participation index underpins the importance of vertical specialisation and the degree to which a country is involved in a vertically fragmented production process, the "length of GVCs" is a complementary indicator that shows the average number of production stages required to realise a product or provide a service in a given final industry (OECD, 2015). The "length of GVCs" index allows for a more in-depth analysis of trade patterns as well as the gains from trade that a country exploits (Grossman and Rossi-Hansberg, 2008).

The average number of stages involved in the production of good ***i*** (N_i) depends on the number of stages required for each intermediate good (Fally, 2012). The index can be calculated at the industry level or aggregated at country level. Furthermore, it can be decomposed to reflect domestic production stages and foreign production stages by distinguishing between domestic inputs and foreign inputs.

The index of number of production stages was introduced by Fally (2012). It is calculated using input-output tables and is defined as a weighted-average number of plants involved sequentially in the production of good ***i***.

[5] Data for the index is extracted from the OECD Global Value Chains Indicators database updated as of May 2013. Latest available data is for 2009.

$$N_i = 1 + \Sigma_j \mu_{ij} N_j$$

where μ_{ij} is the share of inputs from industry j used to produce a good in industry i.

In our work we have used the index "length of GVC" based on the concept developed by Fally (2012) and calculated in the multi-country, multi-industry framework of the OECD Global Value Chains database. It is expressed through the following matrix:

$$N = u.(I\text{-}A)^{-1}$$

where N is a row vector with the indexes for all countries i and industries k, u is a row unit vector, and $(I\text{-}A)^{-1}$ is the well-known Leonitief inverse matrix which represents the total production value that is directly and indirectly required to produce for final demand (De Backer and Miroudot, 2014).

The index takes the value of 1 if there is a single production stage in the final industry. The value of the index increases as intermediate inputs from the same industry or other industries are used in the production of the final good or service (OECD, 2015).

Distance to Final Demand.[6] Once the depth and length of GVCs is assessed, the question of where an economy is located in the value chain remains. In order to establish the position of a country in the value chain we refer to an indicator called "distance to final demand". The index defines whether a country is situated "upstream" or "downstream" in the GVC, depending on its specialisation. An "upstream" country is one which produces the raw materials in the beginning of the production process, whereas a "downstream" country is associated with the production of products closer to final consumers.

The "distance to final demand" is measured in terms of production stages and starting from one industry in a certain country it measures the number of production stages left before the goods or services reach final consumption. It is introduced by Fally (2012) and Antras et al. (2012). For each product Fally (2012) defines the distance to final demand (D_i) as:

$$D_i = 1 + \Sigma_j \varphi_{ij} D_j$$

[6] See ref. 5

where φ_{ij} denotes the share of the production of good ***i*** that is used as an intermediate good in industry ***j***.

Similarly, in a multi-country and multi-industry framework the OECD Global Value Chains database calculates "distance to final demand" as:

$$D = u.(I - G)^{-1}$$

where ***D*** is a row vector with the indexes for all countries ***i*** and industries ***k***, ***u*** is a row unit vector, and $(I - G)^{-1}$ is best known as the output inverse or Gosh inverse matrix (De Backer and Miroudot, 2014).

The higher the value of the index, the more "upstream" the country or industry is – the more it is involved in the production of inputs at the beginning of the value chain. In case the entire production of a good is used for final consumption, the value of the index is one (Fally, 2012).

4. Factors to Consider When Interpreting the Results of GVCs Indicators

Economies are not all equally engaged in GVCs, just as they are not equally engaged in international trade. The role of individual countries in international production networks depends on different factors, such as the size of their economy, their level of development and the composition of their exports. As a result, countries with very different characteristics may be very similar in the ranking of GVCs participation. In this section we present some considerations, related to the interpretation of the indicators for participation in GVCs.

- **Level of fragmentation**. We should note that the level of fragmentation of production depends on the technical characteristics of products. Not all products can be produced in multiple stages. For example, vertical specialisation is less typical for services sectors, since face-to-face contact between the provider and the consumer is usually required in this area. Traditionally, a selected number of manufacturing industries has been at the forefront of value chain segmentation. The electronics and automotive industries, where products can be broken down into discrete components that can be separately produced, easily transported, and assembled in low-cost locations, have led the way in GVCs expansion and consequently rank highest by share of foreign value added in trade. A number of industries that incorporate and process outputs from extractive industries and traded

commodities (*e.g.* petroleum products, plastics, basic chemicals) follow closely behind (UNCTAD, 2013).

- **Share of natural resources in exports**. Countries with significant shares of natural resources in their exports, such as Russia, Brazil, Indonesia and Saudi Arabia, tend to have higher relative domestic value added trade shares, since these commodities are at the "beginning" of GVCs and require little foreign inputs (UNCTAD, 2013). Due to this outsized downstream usage of their export products in third countries' exports, the measures for participation of natural resource exporters in GVCs appear elevated compared to other countries with similar characteristics.
- **Size of the economy**. Generally, small economies tend to score better than large economies in terms of participation in GVCs. This is because large economies like the US and Japan tend to display a greater degree of self-sufficiency in production for exports (large economies rely less on foreign inputs). Moreover, they also have lower downstream participation levels because of the focus on exporting the so-called final demand goods and services, *i.e.* those not used as intermediates in third countries' exports (UNCTAD, 2013). Small economies tend to have higher import contents of exports and also export more intermediate rather than final products, which is indicative of a high degree of participation in GVCs. Such examples are the countries from East and South-East Asia, where GVCs participation is among the highest in the world, reflecting the region's specialisation in export-oriented highly segmented manufacturing and processing activities. Due to the nature of the products they export, these countries also need to import more to generate exports, which in turn reduces the domestic value added embodied in their exports. However, despite the low domestic value added content of these countries' exports, the overall contribution of value added trade to their GDP is usually high because of the overall importance of trade relative to the size of their economies.
- **Development stage of the economy**. According to Fally (2012), developed and developing countries tend to specialise in different stages along the value chain. In particular, richer countries like the US and Japan have a comparative advantage in goods that involve fewer production stages and which are closer to final demand.
- **Importance of Services**. Production in GVCs increases the need for the co-ordination and efficient linking of production stages across

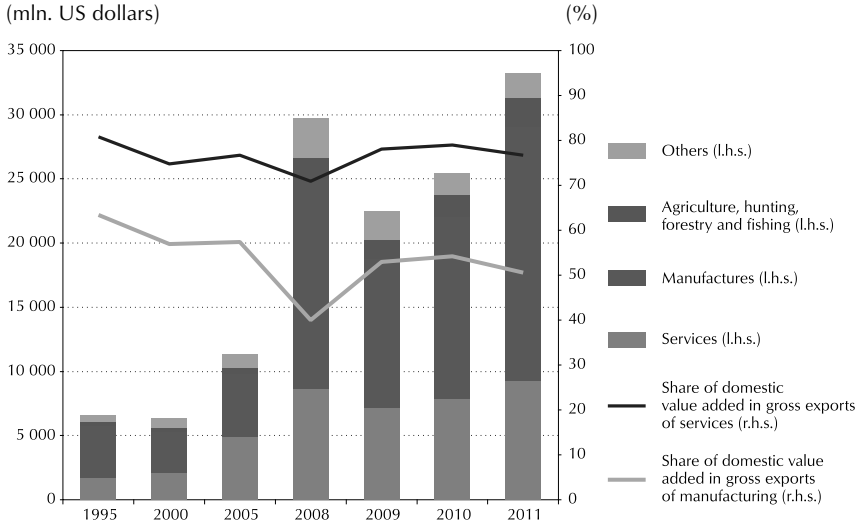
countries, with services playing a particularly prominent role. Services tend to have higher importance for overall exports in developed countries like the US, Germany and France compared to developing countries (OECD, 2013). Countries that rely more on exports of natural resources, such as Russia, Brazil and Canada, and manufacturing-intensive economies, such as China and Korea, have relatively low share of services value added in total exports. Despite this difference, the contribution of services has been increasing over time for both groups of countries. It should also be noted that services tend to have higher content of domestic value added because of the low usage of foreign inputs.

Based on these insights we could expect that as a small open economy that is part of the EU, Bulgaria would be well integrated in GVCs. Given the significance of primary commodities (such as petrol and metals) for the gross exports of Bulgaria, as well as the country's low endowment of natural resources, it follows that foreign inputs are likely to have a large contribution to the overall participation of the country in GVCs. Such participation characteristics are likely to place the products with which Bulgaria participates in GVCs relatively far away from final consumers. In the following two sections we analyse whether these expectations hold by discussing the structure of Bulgaria's exports and by examining the indicators discussed in Section 3.

5. Structure of Bulgaria's Exports

Before we focus on the indicators related to Bulgaria's participation in GVCs it is worth discussing the structure of Bulgaria's exports, which in recent years has established itself as one of the main drivers of economic growth in the country. An industrial breakdown of gross exports allows us to identify the sectors that are at the core of this export-led growth and after examining Bulgaria's participation in GVCs to be able to conclude whether these sectors are the ones in which Bulgaria participates in GVCs the most. As can be seen in Figure 3 the increase in gross exports in the period between 2000 and 2011 has been dominated mainly by the expanding exports of the manufacturing sector. Exports of services have also increased in the period under consideration but the growth was not as significant as the one in the manufacturing sector. It is interesting to note that exports of services are much less volatile than the manufacturing ones and during the financial crisis we can see that overall gross exports fell largely due to the lower manufacturing exports.

Figure 3: **Structure of Bulgarian Gross Exports**[7]



Source: OECD TiVA database, own calculations.

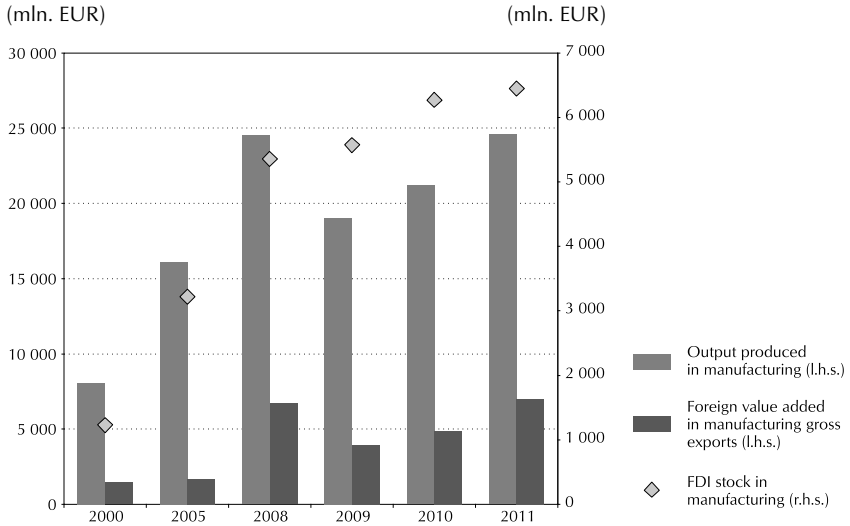
The large increase in the gross exports of the manufacturing sector between 2000 and 2008 can be partly explained with production expansion, following significant FDI inflows in this period (between 2000 and 2008 the stock of FDI in the manufacturing sector increased more than four times). These investments were mainly aimed at upgrading and expanding existing production units and as a result by 2008 total production in the manufacturing sector had increased more than three times compared to its 2000 level.

It should be noted, however, that the foreign value added in gross exports of the manufacturing sector increased more than the increase in the output of the sector (see Figure 4). This can partly be explained with the fact that major enterprises in the sector were bought by large multinational companies in the early 2000s and were integrated in the international supply chains of those companies. Data from EUROSTAT shows that in the

[7] In this paper data for **Total Manufactures** as specified in the OECD TiVA database includes the following sub-categories: food products, beverages and tobacco, textiles; textile products, leather and footwear; wood, paper, paper products, printing and publishing; chemicals and non-metallic mineral products; basic metals and fabricated metal products; machinery and equipment nec; electrical and optical equipment; transport equipment; manufacturing nec: recycling. Data for **Total Services** incorporates as follows: wholesale and retail trade; hotels and restaurants; transport and storage, post and telecommunication; final intermediation; real estate, renting and business activities.

Bulgarian manufacturing sector the share of production value generated by foreign-owned[8] firms in total production value has increased from 37.1% in 2003 to 54.6% in 2011. This has led Bulgarian affiliates to increase their usage of foreign inputs, produced elsewhere in the supply chain, which in turn led to an increase in the foreign value added component of these companies' exports.

Figure 4: **Manufacturing Sector in Bulgaria**

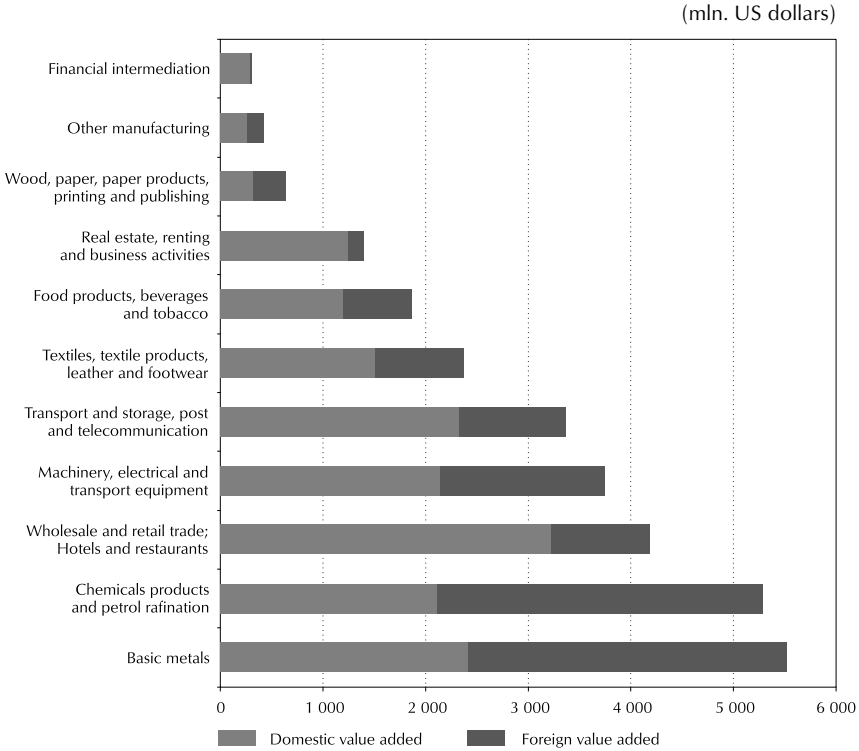


Source: NSI, BNB, OECD TiVA database, own calculation.

Despite that manufacturing appears to be the engine of Bulgarian exports growth between 2000 and 2011 (see Figure 3) we should have in mind that the benefits that the Bulgarian economy gets from exporting (in the form of employment, skills and income growth) could be larger in the services sector due to its higher level of domestic value added in exports. A closer look at the industry breakdown confirms that, indeed, manufacturing exports have a significant share of foreign value added and in 2011 two of the top three sectors in terms of exported domestic value added are services sectors (see Figure 5). This ambiguous role of the services sectors is why we will examine their contribution to overall trade performance in more detail in Section 7.

[8] According to the EUROSTAT definition, foreign-owned firms are the ones where the controlling institutional unit is resident in a different country from the one where the institutional unit over which it has control is resident. http://ec.europa.eu/eurostat/cache/metadata/en/fats esms.htm

Figure 5: **Value Added in Bulgarian Exports (2011)**



Source: OECD TiVA database, own calculations.

Now that we have established the general industrial structure of Bulgarian exports we can examine Bulgaria's role in GVCs and see whether the industries that led the overall exports growth are also the ones that are the most active participants in GVCs.

6. GVCs-specific Indicators

6.1. Participation Index

The participation index provides a direct way to establish the degree to which a country is involved in GVCs. As indicated in Section 3, by using this index, we are able to show the extent to which Bulgaria is taking part in vertically fragmented production chains as an importer of foreign inputs (backward participation) and as a supplier of intermediate products used in other counties' exports (forward participation).

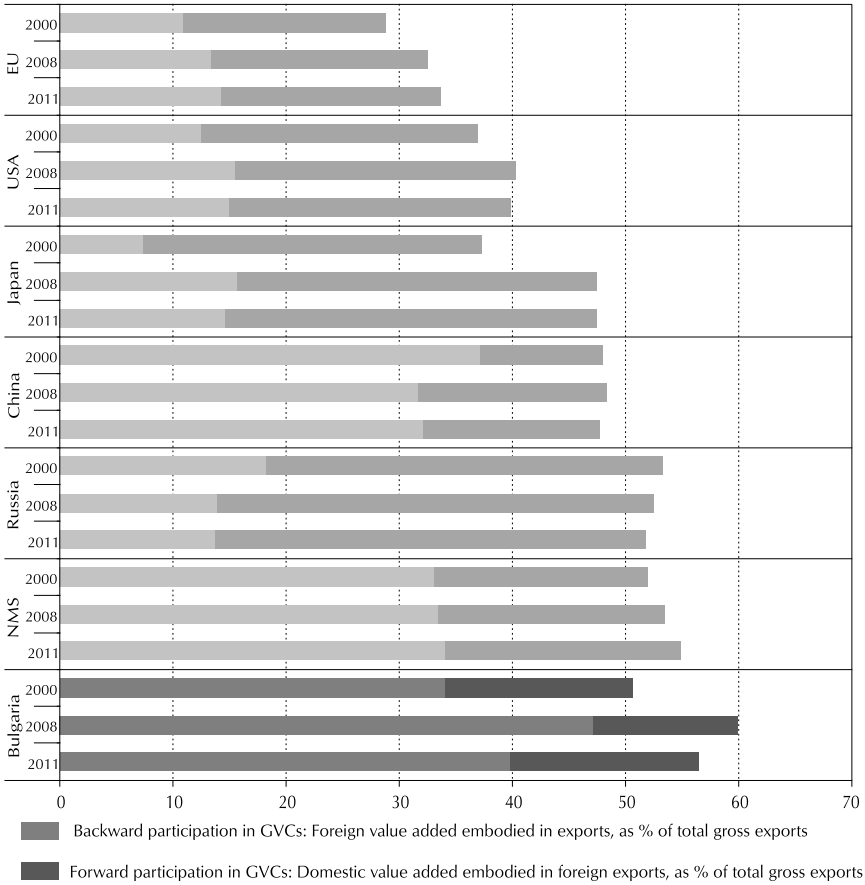
The participation index for Bulgaria is higher than the participation indices for the rest of the countries in the reference group (see Figure 6), which is mainly due to the high degree of backward participation of the country. This finding is consistent with our initial expectations. As a small open economy, Bulgaria sources more foreign inputs than large economies such as the US and the EU. Moreover, the large dependency on foreign inputs of the Bulgarian participation is even more evident when comparing it with a natural resource exporter such as Russia. This predominant backward participation of Bulgaria could be beneficial for the country since studies find that increases in foreign value added are associated with consequent increases in domestic value added (IMF, 2013) and that foreign value added in exports is a major channel of productivity increases through knowledge and technological transfers (Van der Marel, 2015). The result for Bulgaria's GVCs participation index is most similar to the one for the NMS[9], both in terms of participation intensity and in terms of composition (with backward participation accounting for the largest part of the index). This is not surprising, given the similar level of economic and institutional developments in these countries, related to their membership in the EU.

Looking at the period 2000–2008 we observe that, with the exception of Russia, the countries in the reference group have recorded an increase in the participation in GVCs and this suggests an increase in the vertical specialisation of production (Di Mauro et al., 2013). This increase in the participation index is particularly evident for Bulgaria and could be related to the EU pre-accession period and the potential deepening of the integration of Bulgaria in EU GVCs.

Even though the global financial crisis in 2009 has contributed to a slowdown in the GVCs integration, Figure 6 illustrates that by 2011 the average participation in GVCs shows signs of recovery. For the case of Bulgaria the GVCs participation index in 2011 is lower than the level registered in 2008. It should be noted, however, that the country's forward participation has increased, which is a sign that Bulgaria is increasingly participating in the production of inputs used in third countries' exports. Even though Bulgaria hasn't fully recovered its pre-crisis GVCs participation rate by 2011, it still exhibits the highest value of the participation index compared to the rest of the countries in the reference group, which is indicative for the high integration of Bulgaria in GVCs.

[9] From here on in this analysis NMS should be understood as the averaged result across the NMS countries.

Figure 6: **GVCs Participation across Countries**

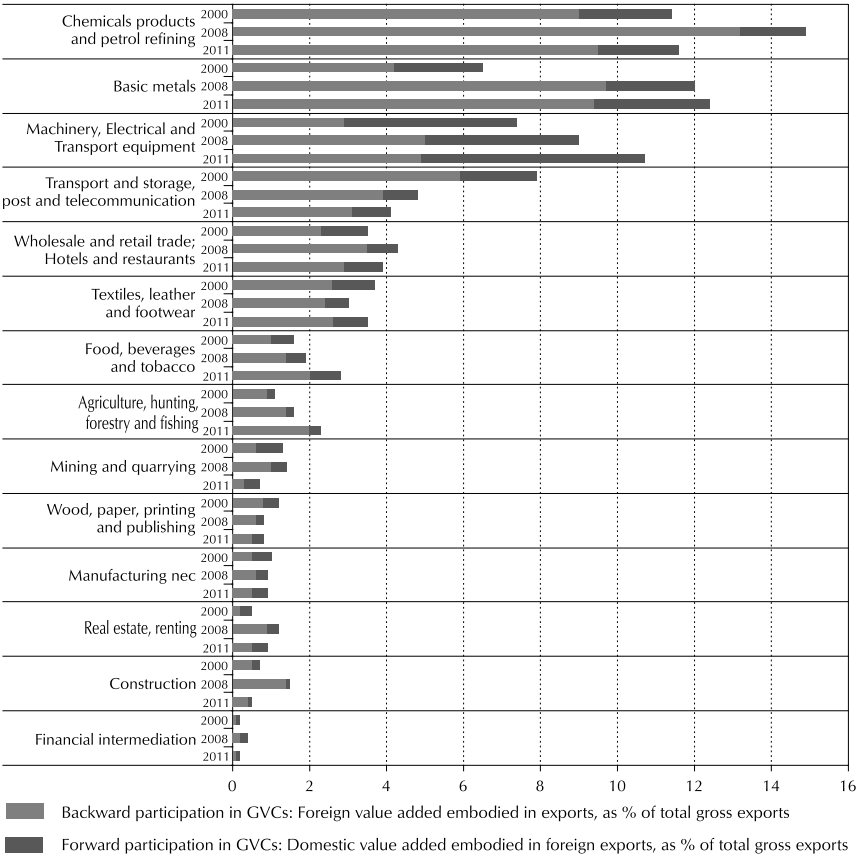


Source: OECD TiVA database, own calculations.

An industrial breakdown of the participation index (see Figure 7) shows that Bulgaria is most active in GVCs in the processing sectors of the manufacturing industry and mainly as an importer of foreign intermediate products (backward participation). By 2011 chemical products and petrol refining, basic metals and machinery, electrical and transport equipment had established themselves as the sectors with the highest participation index for Bulgaria. Overall, the country's role in GVCs in these sectors is mainly through the foreign value added embodied in exports (backward participation) but its forward participation has registered an increase compared to

its 2008 level. Moreover, Bulgaria's participation in the machinery industry has accelerated notably and this acceleration was defined to a large extent by Bulgaria's forward participation. This is evidence that Bulgaria is moving towards integration in supply chains in technology intensive sectors with higher value added production.

Figure 7: **Bulgaria's Participation in GVCs across Industries**



Source: OECD TiVA database, own calculations.

We should note that our results for the participation index and mainly the backward participation component of the index might be affected by price effects. The effect of prices on the amount of foreign value added in a country's exports is discussed by Cappariello and Felettigh (2015) and

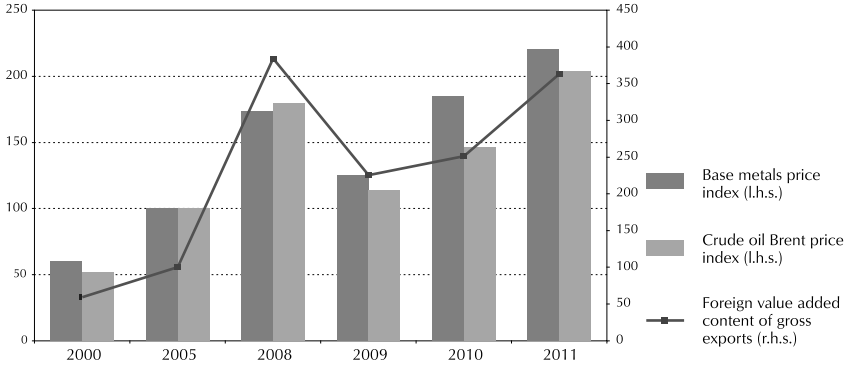
Nagengast and Stehrer (2015). Both papers conclude that the effect of prices is not the main driver of changes in the foreign value added content of gross exports but, nevertheless, should not be overlooked.

Since the value added is calculated from nominal data, an increase in prices of foreign inputs is likely to lead to an increase in foreign value added content of exports and, thus, to an upward shift in the relative importance of the backward component of the GVCs participation index to the forward one (given that we hold domestic value added more or less constant). This is quite evident in countries that specialise in processing since companies in that sector generally add certain value to the inputs they process. This domestic value added (which would represent the company's mark-up) is more or less stable in the short run since companies are constrained by their production technologies. On the other hand, input costs may vary significantly (like the international price of petrol and base metals) and since gross exports and domestic and foreign value added components are calculated using nominal data, it is likely that the foreign value added component of gross exports is more volatile.

This is relevant for the case of Bulgaria, where most of the participation in the GVCs and foreign value added is concentrated in the processing industry (for example, petrol products, metals and chemical products). This in turn explains why we see such significant movements in the foreign value added content of gross exports in line with the international business cycle, while the domestic value added, embodied in foreign exports, is much more stable (see Figure 6). According to Figure 8 in periods of cyclical upturns, when global economic activity is accelerating and prices of primary commodities (and thus prices of processing inputs) are rising, the foreign value added of gross exports tends to increase. On the other hand, in periods of global economic downturns, when global demand weakens and international prices of primary commodities are falling, the foreign content of gross exports also tends to fall.

Figure 8: **Relationship between International Prices of Commodities and Foreign Value Content of Bulgarian Gross Exports**

(index 2005=100, US dollars)

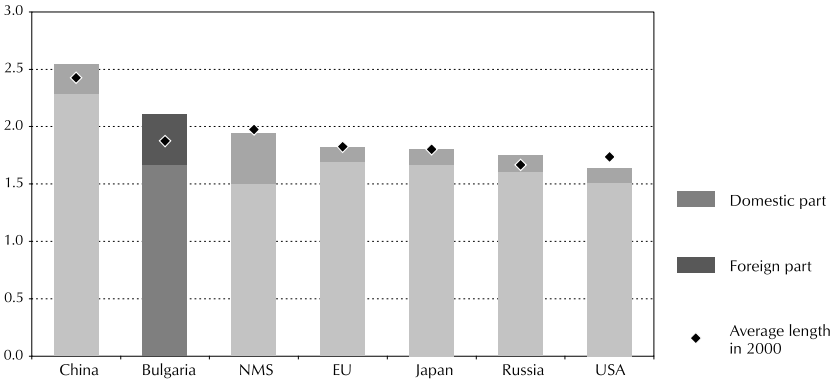


Source: ECB, OECD TiVA database, own calculations

6.2. Length of Global Value Chains Index

Thus far we have discussed the most widely used GVCs indicator, namely the participation index. However, in order to get a full picture of the role of Bulgaria in GVCs, this analysis looks at a complementary index which gives more information on the particular characteristics of the GVCs in which the country participates. The indicator for the average length of GVCs could be interpreted as showing the degree of production fragmentation in GVCs (see Figure 9).

Figure 9: Average Length of GVCs across All Industries (2009)

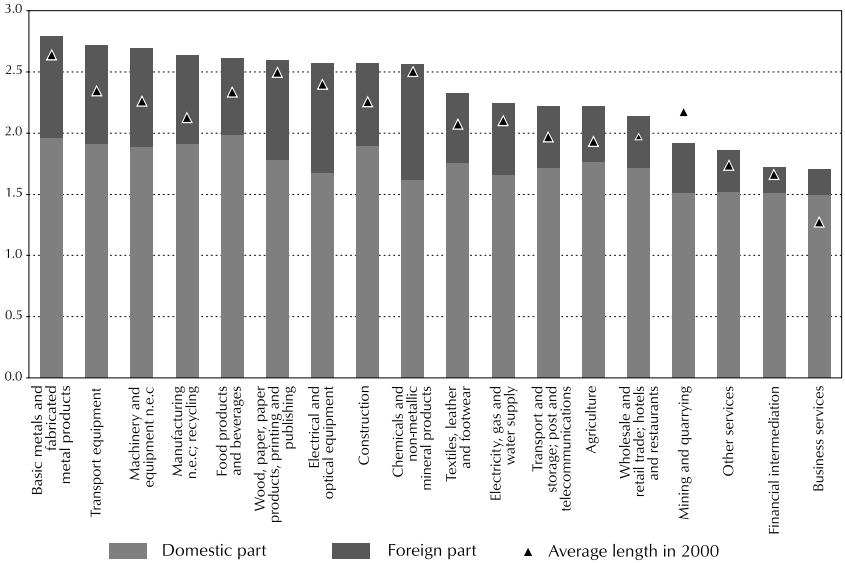


Source: OECD GVCs indicators database, own calculations.

As expected in our initial hypothesis, Bulgaria has one of the highest average lengths of GVCs across the considered countries, which is in line with the fact that Bulgaria has the highest rate of participation in the group and that its participation is dominated by sectors with highly fragmented production processes.

As noted in Section 3, production stages can be divided into domestic and foreign and through that the average length index can be divided into a domestic part and a foreign part. The 'length of GVCs' indicator shows that for all countries in our reference group, including Bulgaria, the domestic part is much higher than the foreign part. However, we should indicate that Bulgaria and the NMS have the highest foreign share of the average length of GVCs which could be attributed to the lower level of self-sufficiency they have compared to bigger and more developed economies like the US and Japan. Furthermore, this relatively higher foreign part is in accordance with the significant share of backward participation that both Bulgaria and the NMS register.

Figure 10: **Length of GVCs by Industry (2009)**



Source: OECD GVCs indicators database, own calculations.

The index shows that for the case of Bulgaria, unlike the US, Japan and the EU, there has been an increase in the average length of GVCs in the period 2000–2009 (see Figure 10). The notable increase of the index for

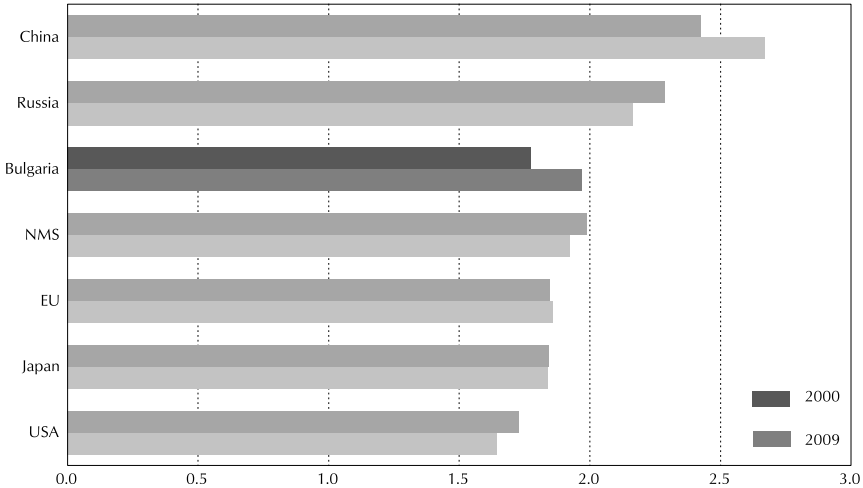
Bulgaria compared to the rest of the reference group could be due to the increasing role of the manufacturing sector for Bulgaria's exports. Fally (2012) finds that fragmentation of production in the US has been decreasing over time mainly because of the increasing role played by services relative to the manufacturing sector in the US economy. In the case of Bulgaria it could be the opposite – the increasing role played by the manufacturing sector compared to that of the services sector could have led to the increase in the average length of GVCs that the country participates in. We can assess this hypothesis by investigating the industrial breakdown of the average length of GVCs for Bulgaria and its dynamics.

At industry level, the average length of GVCs index shows highest values for Bulgaria for industries in which the participation in GVCs index is also high. In the period 2000–2009 sectors with high fragmentation of production such as machinery and equipment, transport equipment and basic metals have registered significant increases in the average length of their GVCs. This confirms our hypothesis that the overall increase in the average length of GVCs in which Bulgaria participates is a result of an increase in the importance and length of GVCs related to the manufacturing sector. Conversely, the length of the chains for Bulgaria in the services industries such as financial intermediation is quite low and has registered a decrease in the period 2000–2009 due to the generally low fragmentation of production in these sectors (see Section 4).

6.3. Distance to Final Demand Index

In this section we examine the position of Bulgaria in GVCs through the distance to final demand index. Using this index we establish whether Bulgaria's participation is at the beginning of the GVCs or closer to the final consumers. The value of the index is to a large degree defined by the production specialisation of the country. According to Figure 11, Bulgaria, similarly to the NMS, has a high distance to final demand index. The index establishes Bulgaria in the reference group as more involved in the initial stages of the production process (an "upstream country") and further away from final consumers, compared to large developed economies like the USA, EU and Japan. The only countries that register higher distance to final demand are China and Russia. In the case of China this can be explained with the country's role as a world leader in terms of assembly and processing functions. Russia, on the other hand, is a country rich in natural resources, which are typically used at the beginning of the production process. This explains why Russia's exports are far from final demand.

Figure 11: **Distance to Final Demand for Bulgaria**



Source: OECD GVCs indicators database, own calculations.

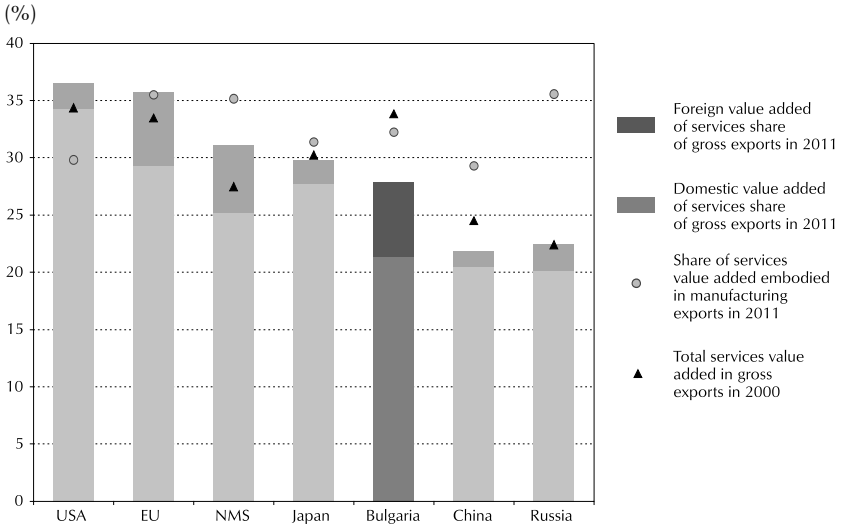
Bulgaria's distance to final demand index has increased between 2000 and 2009 (see Figure 11). This can be interpreted as a sign that the country is deepening its specialisation in the production of inputs. Studies find that on average, most countries move upstream (their distance to final demand increases) which is explained with the overall increase in the length of GVCs and the outsourcing phenomenon. When the production of some inputs is outsourced, their value added is moved backward to the industries supplying intermediate inputs and the distance to final demand increases (De Baker and Miroudot, 2014). In the case of Bulgaria the increase in the distance to final demand between 2000 and 2009 is likely to reflect the increased outsourcing as part of the country's integration in production processes with higher fragmentation. The fact that Bulgaria has moved upstream is consistent with the observed increase in the length of the GVCs that the country participates in.

7. The Role of Services

Production in GVCs increases the need for the co-ordination and efficient linking of production stages across countries, with services playing a particularly prominent role. Production in GVCs relies heavily on logistics, information and communications technologies and therefore on efficient network infrastructures and complementary services (OECD, 2013). In

fact, a significant part of the international production networks of MNEs are geared towards providing services inputs to their country affiliates (UNCTAD, 2013). The importance of services becomes evident once data on trade in value added is used. According to UNCTAD (2013), the share of services in gross exports worldwide is only around 20%, while around 40% of the value added inputs to exports is contributed by service-sector activities since most manufacturing exports require significant services inputs for their production. Therefore, in order to complete our analysis of the role of Bulgaria in GVCs we look further at the role of services.

Figure 12: **Services Value Added (2011)**



Source: OECD TiVA database, own calculations.

According to Figure 12 the share of services value added in gross exports of Bulgaria is lower than the one for developed economies like the USA, EU and Japan. It is also lower than the average for its EU peers from Central and Eastern Europe and, thus, suggests that Bulgaria's services sector is still underdeveloped compared to the other NMS. Unlike the NMS and the EU as a whole the share of services value added in total Bulgarian gross exports has fallen compared to 2000, which can be explained by the increasing importance of manufacturing sector exports (relative to that of the services sector exports) for the overall growth of gross exports (illustrated in Figure 3). We could expect that as Bulgaria catches up with more developed economies the importance of services value added for

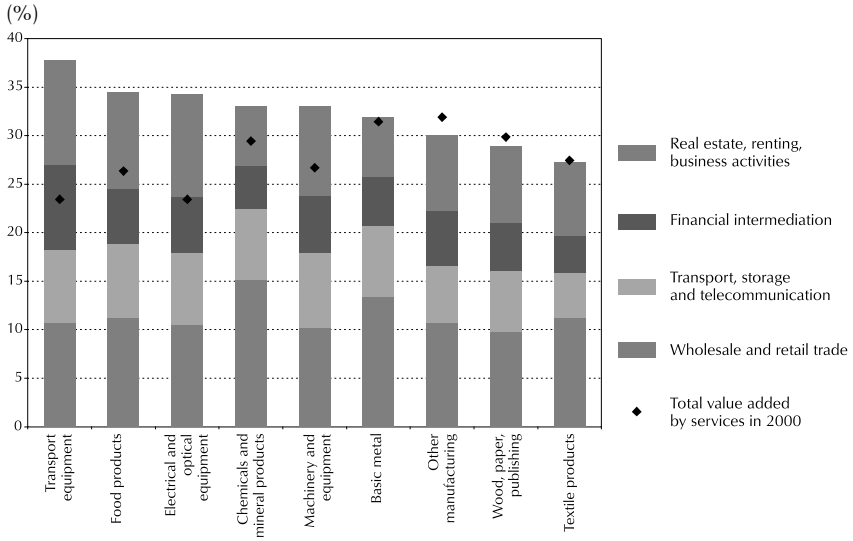
exports will increase, becoming a vital source of productivity growth and competitiveness.

Despite the lower importance of the services sector for the country's overall gross exports, Bulgaria shares some similar characteristics with NMS and the EU in terms of the relatively high share of foreign services value added in gross exports and the high share of services value added embodied in manufacturing exports. Both characteristics are related to the high fragmentation of production across countries in the EU. Numerous studies like IMF (2013) show that since the 1990s large labour cost differentials together with geographical proximity and cultural similarities have led many firms from developed euro area countries (especially Germany) to shift part of their manufacturing production (like motor vehicles and other transport equipment) to Central and East European countries, while retaining control over functions with high domestic value added such as production management, design, marketing and R&D. This can explain the high share of foreign services value added in gross exports and the high share of overall services value added in manufacturing exports in Central and East European countries.

As mentioned earlier, manufacturing includes more and more value added from service-related activities. Firms use logistics, communication services and business services in order to facilitate the efficient functioning of the production process in GVCs. At the same time firms use services like design, development, marketing, warranties and after-sales care to differentiate, customise and upgrade products, allowing them to capture more value. According to OECD (2013), value created by services as intermediate inputs represents over 30% of the total value added in manufactured goods. The results for Bulgaria are similar – almost 33% of Bulgaria's manufacturing exports represent services value added (see Figure 12). This is lower compared to the percentage in the rest of the NMS and the average level for the EU, which may reflect the different product structure of Bulgarian manufacturing exports.

Distribution and transport services contribute more than half of the services value added in Bulgarian manufacturing exports as they provide the necessary links in supply chains. Business and financial services which are aimed at improving the efficiency of goods production inside the GVCs generally contribute to a lesser extent in the case of Bulgaria (see Figure 13). It appears that manufacturing sectors that have considerably increased their usage of services inputs are part of the more complex and higher value added machinery production industry.

Figure 13: **Services Value Added Embodied in Manufacturing Exports by Industry, Bulgaria (2011)**



Source: OECD TiVA database, own calculations.

Although there are signs of increasing usage of services in fast-growing higher value added sectors such as machinery production, the role of services for Bulgaria's exports remains less significant than the one for more developed economies. The underdeveloped services sector is one of the reasons why Bulgaria's participation in GVCs is mainly through the usage of foreign value added in its exports.

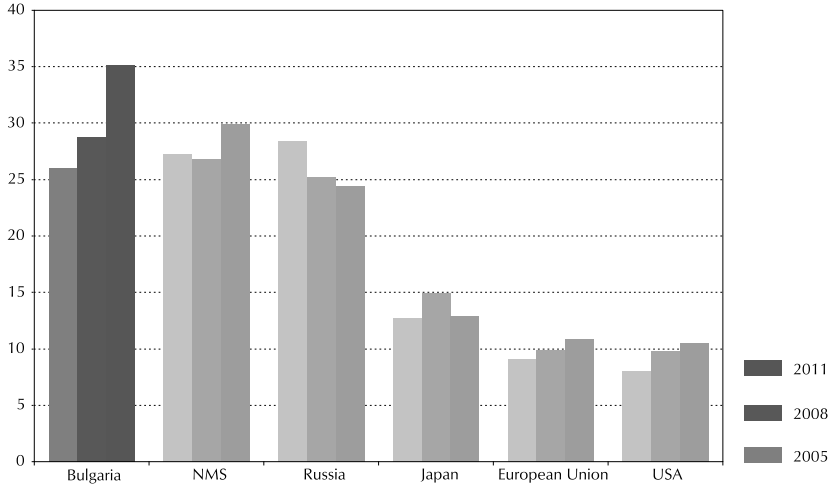
8. Impact of Participation in GVCs on Bulgaria

Our results so far suggest that Bulgaria's exports are predominantly intermediary products, which are far away from final consumers and have low domestic value added content. This poses the question of whether Bulgaria is able to benefit from its, as we saw, extensive participation in GVCs.

Despite the fact that domestic value added content of exports is relatively low for Bulgaria the share of domestic value added in exports to GDP is higher than that of the other countries from the reference group because of the large volume of exports relative to the size of the economy (see Figure 14).

Figure 14: **Domestic Value Added Embodied in the Country's Gross Exports as a Share of GDP**

(%)

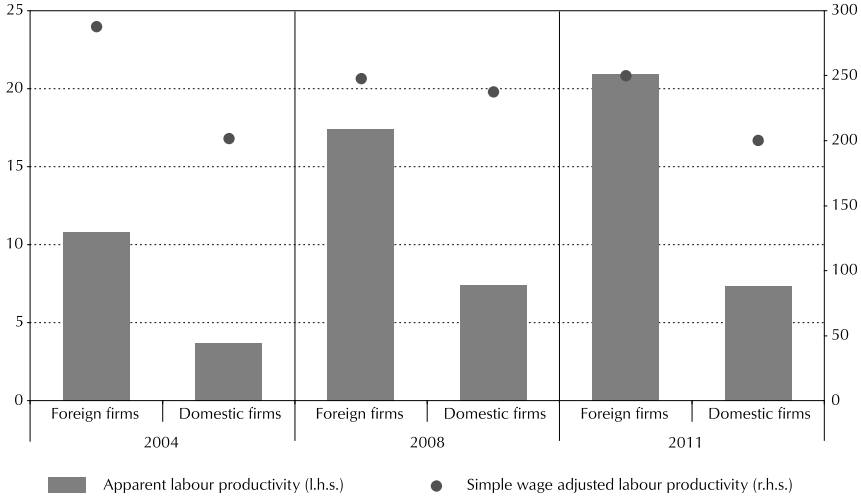


Source: OECD TiVA database, IMF WEO April 2016, own calculations.

Figure 14 shows that in the period 2005–2011 the domestic value added embodied in gross exports as a share of GDP increased most significantly in Bulgaria compared to the other reference countries and regions. This also corresponds to a period of increasing participation of Bulgaria in GVCs. This would imply that the increased exports that resulted from Bulgaria's deeper integration in GVCs has stimulated the generation of domestic value added, which is likely to have boosted economic activity by increasing the demand for labour and investment.

Section 5 illustrated that foreign direct investment has played a crucial role for the deeper integration of Bulgaria in GVCs and that foreign-owned firms account for a large part of the production value (around 39% of total production value in 2011), especially in the sectors that register highest GVCs participation rates for Bulgaria (*e.g.* the manufacturing sector, where foreign-owned firms account for around 55% of total production value in 2011). Evidence suggests that foreign-owned firms in Bulgaria exhibit higher productivity rates compared to domestically-controlled firms and thus raise the overall productivity of the economy (see Figure 15).

Figure 15: **Productivity of Foreign- and Domestically-owned Firms in Bulgaria**



Note: *Apparent Labour Productivity is calculated as gross value added per person employed; Simple wage adjusted labour productivity is calculated as gross value added by personnel costs, expressed in percent (%).*

Source: EUROSTAT, own calculations.

What is more, the presence of these more productive foreign-owned firms on the local market is likely to have positive spill over effects on the rest of the economy through knowledge and technology transfers, as well as through increased competition.

9. Conclusion

The production of goods has become increasingly fragmented across countries due to the expansion of GVCs, which defines the need for policy makers to be well-informed about the position of their countries in such chains. In this paper we analyse the role of Bulgaria in GVCs by using value added data from the OECD TiVA (October 2015) and OECD Global Value Chains (May 2013) databases. The use of these databases offers a number of advantages in examining participation in GVCs compared to standard gross trade statistics. However, it should be noted that the low frequency of the data in the databases as well as the significant lag with which the data is published represent a limitation to the analysis.

For the purpose of comparison we present the results for Bulgaria alongside those for a chosen reference group of countries within which we

pay particular attention to the countries from Central and Eastern Europe that joined the EU after 2004. The analysis on the participation in GVCs is centred around three main indicators, namely "Participation in GVCs", "Length of GVCs" and "Distance to final demand". The paper also discusses the role of services for Bulgaria's export performance, given their increasing importance in the operations of GVCs.

Our findings show that Bulgaria has increased its involvement in GVCs in the years prior and after the country's accession in the EU and as of 2011 registers the highest GVCs participation rate among the countries from the chosen reference group. Bulgaria's involvement in GVCs is mainly in terms of the manufacturing sector, where the increased participation in supply chains was facilitated by the large inflows of FDI, especially in the years before the global financial crisis. Our results point to the conclusion that Bulgaria participates in GVCs mainly as an "upstream" country, which specialises in processing functions. As such Bulgaria requires the use of considerable amounts of foreign inputs which in turn leads to a high share of foreign value added in total Bulgarian exports. Despite that the share of domestic value added in Bulgaria's exports appears to be low, the welfare gains from trade are larger for Bulgaria than for the rest of the countries in the reference group due to the large volume of trade relative to the size of the Bulgarian economy. Our results also suggest that given Bulgaria's specialisation in processing functions, the country participates in highly fragmented production chains and exports mostly intermediary products, which are situated far away from final consumers. The increased participation of Bulgaria in highly fragmented production chains of the manufacturing sector between 2000 and 2009 has led to an increase in both the average length of GVCs in which Bulgaria participates and the distance to final demand of the country's exports. Our analysis shows that the services sector in Bulgaria remains less developed compared to advanced economies and the NMS, both in terms of the direct exports of services and the extent to which services are used to facilitate the efficient production of other sectors in GVCs. However, as Bulgaria catches up with more developed economies the importance of services value added for exports will increase, becoming a source of productivity growth and competitiveness. Such a tendency can already be observed in some fast-growing higher value added manufacturing sectors, where the importance of services' inputs has been increasing.

Bibliography

Alfaro, L., P. Antras, D. Chor and P. Conconi (2016), *"Internalising Global Value Chains: A Firm-Level Analysis"* NBER Working Paper No. 21582.

Antras, P., D. Chor, T. Fally and R. Hillberry (2012), *"Measuring the Upstreamness of Production and Trade Flows."* National Bureau of Economic Research Working Paper 17819.

Baldwin, R. (2006), *"Globalisation: The Great Unbundling(s)"*, Economic Council of Finland.

Banga, R. (2013), *"Measuring Value in Global Value Chains"*, UNCTAD, Background paper no. rvc-8.

Cappariello, R. and A. Felettigh (2015), *"How Does Foreign Demand Activate Domestic Value Added? A Comparison among The Major Euro-Area Economies"*, Banca D'Italia Working Papers, No. 1001.

Daudin, G., C. Rifflart and D. Schweisguth (2011), *"Who Produces for Whom in the World Economy?"* Canadian Journal of Economics, 44(4): 1403–1437.

De Backer, K., S. Miroudot (2014), *"Mapping Global Value Chains"*, ECB Working Paper Series No 1677/May 2014.

Dedrick, J., K. Kraemer and G. Linden (2009), *"Who Profits from Innovation in Global Value Chains?: A Study of the iPod and Notebook PCs"*, Industrial and Corporate Change, p. 1–36.

Di Mauro, F., H. Plamper and R. Stehrer (2013), *"Global Value Chains: A Case for Europe to Cheer up"*, Competency Policy Brief, ECB p. 3.

Fally, T. (2012), *"Production Staging: Measurement and Facts"*, University of Colorado-Boulder, August 2012.

Gereffi, G. (1994), *"The Organization of Buyer-Driven Global Commodity Chains: How US Retailers Shape Overseas Production Networks"*, in G. Gereffi and M. Korzeniewicz (eds), Commodity Chains and Global Capitalism, Westport, CT: Praeger, p. 95–122.

Grossman, G.M. and E. Rossi-Hansberg (2008), *"Trading Tasks: A Simple Theory of Ojfshoring"*. The American Economic Review, 98(5): 1978–1997.

Hopkins, T. and I. Wallerstein (1977), *"Patterns of Development of the Modern World-System"*, Review, Vol. 1, No. 2, p. 157–170.

Hummels, D., J. Ishii, and K-M. Yi (2001), *"The Nature and Growth of Vertical Specialization in World Trade"*, Journal of International Economics, 54(1): 75–96.

IMF (2013), *"Multi-Country Report"*, IMF Country Report No. 13/263.

Institute of Developing Economies (2011), *"The Dynamics of Local Learning in Global Value Chains: Experiences from East Asia"*. Japan External Trade Organisation. Palgrave MacMillan.

International Labour Organization (2016), *"Promoting Decent Work in Global Supply Chains in Latin America and the Caribbean. Key issues, Good Practices, Lessons Learned and Policy Insights "*, ILO Americas Technical Report 2016/1.

International Trade Centre (2014), *"Global Value Chains in Services: A Case Study on Costa Rica"*, Doc. No. DMD-14-257.E.

Koopman, R., W. Powers, Z. Wang and S.-J. Wei (2011), *"The Value added Structure of Gross Exports: Measuring Revealed Comparative Advantage by Domestic Content in Exports"*, Paper presented at the 2012 AEA Annual Meeting, Chicago, 7 January.

Koopman, R., W. Powers, Z. Wang, S. Wei (2010), *"Give Credit Where Credit is Due: Tracing Value Added in Global Production Chains"*, National Bureau of Economic Research, Working Paper 16426.

Nagengast, A. and R. Stehrer (2015), *"The great collapse in value added trade"*, European Central Bank, Working Paper Series, No. 1833/July 2015.

OECD (2013), *"Interconnected Economies: Benefiting from Global Value Chains"*, Report, p. 8.

OECD (2015), *"Trade in Value Added (TiVA) Indicators"*, Guide to Country Notes.

Park, A., G. Nayyar and P. Low (2013), *"Supply Chain Perspectives and Issues"*. WTO Publications

Porter, M. (1985), *"Competitive Advantage: Creating and Sustaining Superior Performance"*, New York: The Free Press.

Rahman, J. and T. Zhao (2013), *"Export Performance in Europe: What Do We Know from Supply Links?"*, IMF Working Paper, WP/13/62.

Sturgeon, T., J. Van Biesebroeck and G. Gereffi (2008), *"Value Chains, Networks and Clusters: Reframing the Global Automotive Industry"*, Journal of Economic Geography, J Econ Geogr (2008) 8 (3): 297–321.

Timmer, M., B. Los, R. Stehrer and G. De Vries (2013), *"Fragmentation, incomes and jobs: an analysis of European competitiveness"*, Economic Policy October 2013, p. 613–66.

UNCTAD (2013), *"Global Value Chains and Development"*, United Nations Publication UNCTAD/DIAE/2013/1.

Van der Marel, E. (2015), *"Positioning on the Global Value Chain Map: Where do u want to be?"* ECIPE Occasional Paper No. 01/2015.

DP/106/2017

Labour Market Model with On-the-job Search, an Agent-based Approach

Mariya Mitkova

Abstract. We use agent-based methodology to simulate a labour market with on-the-job search. Workers continue to search while employed in order to find a job in a better paying firm and climb the wage ladder. In particular, we study the effect of the intensity of job search of both employed and unemployed workers on the unemployment rate and the average unemployment duration. The main result is that whenever employed workers continue to apply for new jobs intensively, i.e. by sending out many applications per period, it becomes very difficult for young and inexperienced workers to enter the labour market because they are never preferred over a more experienced worker. Introducing adaptive behaviour under which workers can adjust their search strategy according to whether they are successful in finding a job or not and whether they are already in a firm which offers high salaries or not, remedies to some extent the congestion effect created by the search of employed workers.

Резюме. Използвана е агентно ориентирана методология за симулиране на пазар на труда с търсене на работа без напускане на съществуващата. Работниците продължават да търсят работа, докато са заети, за да намерят работа във фирма с по-високо заплащане и да заемат по-високоплатена длъжност. По-конкретно, изучава се ефектът от интензифициране на търсенето на работа от страна както на заетите, така и на безработните върху равнището на безработица и върху нейната средна продължителност. Основният резултат е, че когато заетите продължат интензивно да търсят работа, например чрез многократно кандидатстване през съответния период, става твърде трудно за младите и неопитни работници да излязат на трудовия пазар, тъй като те никога не биха били предпочетени пред по-опитните работници. Въвеждането на адаптивен модел на поведение, при който работниците могат да приспособят своята стратегия на търсене на работа според това дали успешно или неуспешно търсят работа и дали вече са или не са във фирма, която предлага високо заплащане, неутрализира в някаква степен ефекта на пренасищане, породен от търсенето на работа от страна на заети лица.

This research was supported by the scholarship program of the Bulgarian National Bank.

Mariya Mitkova, Dept. of Business Administration and Economics, Bielefeld University.
E-mail address: mmitkova@wiwi.uni-bielefeld.de

Acknowledgment:
I would like to thank Andrey Vassilev, Ph.D., for his supervision and greatly helpful comments and contribution. I would also like to thank the two anonymous referees for their insightful suggestions.

Contents

1. Introduction .....

2. Related Literature .....

3. Model Description .....
   3.1. Initial Conditions .....
   3.2. Step Dynamics .....
   3.3. Pseudocode .....

4. Search Intensity .....
   4.1. Benchmark Case .....
   4.2. No On-the-job Search, 1 Application Per Period .....
   4.3. On-the-job Search With Multiple Applications .....
   4.4. Single Runs with Different Distributions over Workers' Search Intensity .....

5. Adaptive Behaviour .....

6. Concluding Remarks and Possible Extensions of the Model .....

References .....

A. Appendix .....

1. Introduction

In the following paper, a labour market model with on-the-job search is considered. Empirical studies suggest that a large fraction of the separations between workers and firms occur because of job-to-job transitions. The estimates vary from country to country (see Jolivet et al. (2006)) but generally point to a high rate of job-to-job movements. Therefore, we incorporate an on-the-job search mechanism in a simple agent-based simulation of the labour market and study how workers' search intensity, both on-the-job and when unemployed, affects the labour market outcomes in terms of unemployment rate and duration until finding a job. For the purpose we test multiple theoretical scenarios which do not rely on empirical data. We investigate what the labour market outcomes will be assuming different distributions over workers' search strategies. More precisely, we are able to show how macro-level frictions emerge, preventing the market from clearing even though on the micro-level workers try hard to find a job by sending out many applications per period. This project is a first step towards developing a more comprehensive model of the labour market.

The agent-based approach is becoming an increasingly popular tool for modelling and analysis across different disciplines. Broadly speaking, agent-based models (ABMs) consist of autonomous agents who interact according to some set of rules. The modeller assigns these rules and as a result we can observe a macro level outcome arising from a micro level behavioural mechanism. Making use of the advancement in computer capabilities, the agent-based approach can be helpful in modelling various complex economic interactions. The "agents" in this context could be consumers, workers, firms but also institutions, governed by some rules of behaviour (Tesfatsion, 2006). One of the advantages of agent-based modelling over the conventional economic modelling tools is that it allows to relax some of the strong assumptions often made in theoretical economic models. For instance, we can go beyond the "representative, rational, utility-maximizing agent" framework and consider agents which are heterogeneous in one or many of their characteristics.

In this model we introduce firm heterogeneity by letting firms draw different bargaining strategies and worker heterogeneity in terms of search strategies, age and experience. The main results suggest that high search intensity is not necessarily beneficial for the workers. On the contrary, it increases the degree of friction on the labour market due to coordination failure on the side of the job-searching workers and competition for workers on the side of the firms. Similarly, Albrecht et al. (2003) find, in an

analytical framework, that introducing multiple applications by unemployed workers exacerbates the coordination failure during the matching process.

Using agent-based methodology we are able to show that allowing multiple applications together with on-the-job search can considerably increase the unemployment rate and the duration until finding a job. To my knowledge, this question has not been explicitly considered in other agent-based labour market models. We find that whenever workers search very intensively for a first job or for a "better" job than their current one, they create a big queue for the available vacancies and decrease the probability of actually forming a match. More precisely, when employed workers continue to send out applications very actively they block the opportunities for young and inexperienced workers to be chosen by the firms. Allowing for multiple applications further exacerbates the market frictions because employed workers are also more picky. They accept only offers that are better than their current job, which means that they often reject offers. Once a worker rejects the job offer, the vacancy remains open until the next round of applications. The lack of coordination between workers in the application process also increases the unemployment duration of young workers. Therefore, depending on workers' search intensities, some vacancies remain open for many periods and the market cannot clear.

2. Related Literature

The implementation of on-the-job search in frictional labour market settings has a strong appeal because empirical evidence has documented that a large portion of the separations between workers and firms in the developed economies occur due to job-to-job transitions. Mattila (1974) was the first study to point out, using U.S. data, that job-to-job transitions account for 50 to 60% of job quits. Pissarides (1994) estimates that 20% of the new hired workers in the U.S. come directly from another job. In a more recent study, Fallick and Fleischman (2004) find that two-fifths of new jobs in the U.S. in the period 1994–2003 are due to job-to-job transitions.

Results for Europe point to smaller but still considerable job-to-job transition rates. Using panel data on 10 European countries and the U.S., Jolivet et al. (2006) find that one can divide the sampled countries into three categories: such with high-, middle- and low- job-to-job turnover rates. High job-to-job transition rates are estimated for the U.K. and Denmark. Germany, the Netherlands, Ireland and the U.S. fall in the middle group, while Belgium, France, Italy, Portugal exhibit lower job-to-job turnover rates. Pissarides and Wadsworth (1994) find, using data from Britain's Labour Force

Survey (1983–1984), that job-to-job transitions constituted three quarters of all job separations in the considered period.

Moreover, the introduction of on-the-job search to the standard models with search equilibrium, is shown to bring desirable features. Mortensen (1994), in one of the early contributions to the topic, finds that search during employment helps explain the Beveridge curve, the procyclical quits as well the countercyclicity of the flows in and out of unemployment. Pissarides (1994) finds that on-the-job search induces higher cyclical volatility of the vacancies and more persistent unemployment than the standard search equilibrium model. In both models the wage is a fixed split of the surplus of the match.

Numerous studies, on the other hand, focus on why workers continue searching when they are already employed. The predominant incentive assumed in the literature is that workers search on the job in order to climb a wage ladder. Mortensen (1990) (as cited in Christensen et al. (2005)) finds that the distribution of wages that on-the-job seekers face stochastically dominates the one over unemployed applicants because employed workers increase their wage by moving from job-to-job. Burdett et al. (2011) extend the model with on-the-job search by assuming that workers accumulate experience. They find that in the first 10 years of their career workers typically change several jobs and increase their wage rapidly. Moreover, the approach leads to a "desirable" wage distribution, *i.e.* humped-shaped with "fat" right tail, which is also the wage distribution shape found in many empirical studies (see, for example, Neal and Rosen (2000), Bontemps et al. (2000)).

Christensen et al. (2005) implement endogenous search effort and find that on-the-job search declines as wages increase. Workers who receive low wages have a higher incentive to search while employed because they can expect higher returns. This theoretical result is also verified by an empirical study which uses data from the Danish Integrated Database for Labour Market Research. Job tenure also reduces the probability that a worker looks for a new employer. On the other hand, younger, skilled and part-time workers are more likely to search while employed (Pissarides and Wadsworth (1994)).

Another stream of literature emphasises educational and skill job mismatches as incentives for on-the-job search. Allen and Van der Velden (2001), for example, find using a comparative international study of the labour market outcomes of people with tertiary education in Japan and eleven European countries, that skill mismatches decrease the job satisfac-

tion and increase the incentive for on-the-job search. Educational or skill job mismatches, however, go beyond the scope of the model perused in this paper.

The labour market exhibits large heterogeneity among the participants, both employees and employers. Numerous agent-based models, some of which very large scale, have already been used to model this diversity with focus on reproducing some stylized facts in labour economics and considering what-if scenarios in order to evaluate the effects of different policies. Neugart et al. (2012) divide the labour market agent-based models into two broad categories: the so-called partial models and labour market models which are part of a larger macroeconomic framework and which investigate the interaction between several markets. One example of the second class is the EURACE model (Deissenberg et al., 2008) which is aimed at modelling the European economy (implements both real and financial markets) and reproducing empirical regularities that emerge at the macro level. The model is used to investigate the outcome of different policy scenarios.

Some examples of agent-based models, which are confined to interactions in the labour market are Guerrero and Axtell (2013), Lewkovicz et al. (2009), Neugart (2004), Gemkow and Neugart (2011) and Richiardi (2004,2006)[1]. The labour market model presented in Ballot (2002), called ARTEMIS, is calibrated to the French labour market in the period the 1970s oil crisis and reproduces the gross flows dynamics and some labour mobility patterns observed it that period. An intermediary which assists firms and workers to meet is implemented, modelled as a temporary help firm, which offers only short term contracts. On-the-job search is also present with the assumption that the unemployed search more intensively and those who are employed under a temporary contract strive to find a permanent position. Firms' internal labour markets are also considered, as the firms' decisions on promotions, lay-offs, job creation and destruction, dismissals and hiring standards. Workers, on the other hand, make decisions about entry and exit into the labour market, applications, registration for a temporary help firm, on-the-job search and abandoning search in case of discouragement. Both firms and workers adapt their behaviour based on past experience.

Besides reproducing the empirically documented flow dynamics, the results of the simulation suggest that the intermediaries have a profound

[1] For a review of agent-based labour market models see Neugart et al. (2012).

effect on the aggregate matching in the market by reducing unemployment and enabling workers to gather experience.

Lewkovicz et al. (2009) build on the ARTEMIS model to study age discrimination on the labour market against senior workers (defined as older than 50 years) and explore how a new type of work contract, the so called Unified Contract, which is an alternative to the definite- and indefinite-term contracts, might influence the unemployment rate in this group. Very briefly summarized, the Unified Contract is indefinite, allows for accumulation of seniority and unconditional firing and does not bear legal justification costs upon firing. The main result of the simulations is that under Unified Contract the unemployment rate for young and middle-aged workers decreases, while it increases for senior workers. Moreover, the last effect is high enough so that the overall unemployment rate also slightly increases.

Neugart (2004), on the other hand, simulates the matching process between firms and workers and studies the emerging matching function. The results suggest that the matching function exhibits decreasing returns to scale which contradicts the conventionally assumed constant returns to scale matching function. Moreover, labour market policies are found to have an effect on the matching function which implies that assuming an exogenous matching technology might be inappropriate.

Gemkow and Neugart (2011) implement an endogenous social network and investigate its effects on hiring and inequality. The authors find that referral hiring decreases when the market is more volatile, *i.e.* when there is higher variability in the labour demand of firms. Sustaining the "friendships" in the network is costly and during volatile time it is more likely that the contacts of a given agent will be also searching for a job. Hence, the result shows that during such times referral hirings decrease and subsequently the distribution of individual incidences of unemployment becomes more equal.

Richiardi (2006) models a labour market with on-the-job search and entrepreneurship, which successfully reproduces some stylized facts such as the Beveridge, Wage and Okun curves. In the model workers differ with respect to their productivity and inclination to entrepreneurship. Firms, on the other hand, have differential growth potential and make decisions regarding the number of vacancies and the wages they offer. There is also a possibility of firm exit if the firm generates negative profits or if all employees leave. Dissatisfied workers may look for a new job or may decide to become employers themselves. Interestingly, Richiardi (2006) finds that a value of the search intensity of 1 application every 3 periods, where one

period is one week, is appropriate. Moreover a higher search intensity is associated with lower unemployment rate under increasing vacancy cost. This result, at first glance contradictory to ours, is, however, not directly comparable since in this model we do not assume a positive vacancy cost.

With the model presented here, we aim to study the interaction of the agents confined to the labour market, which gives us the opportunity to isolate the effect of on-the-job search and the search intensity on the market outcomes. We wish to incorporate on-the-job search in a labour market with heterogeneous firms and *ex-ante* homogeneous workers. The idea is derived from the theoretical models described above, but the agent-based environment allows us to further investigate the mechanism.

3. Model Description

There are two types of agents in this model: firms and workers and we simulate a simple matching process between them. The workers want to find a job and the firms wish to fill their open vacancies. We assume that the firms are infinitely-lived and equally sized. The heterogeneity among them arises from their bargaining strategies, *i.e.* the share of the match surplus they are willing to offer to their workers. The worker and the firm are bargaining over the split of the surplus of the match similarly to Pissarides (1994) and Mortensen (1994). Workers are heterogeneous with respect to their age, experience and search strategy. In the context of the model, search strategy is defined as the number of applications each worker sends per period, where one period is defined as one quarter of a year.

The matching mechanism employed in this model is of urn-ball type. The urn-ball matching is quite intuitive in terms of understanding how coordination failure might arise[2]. As it can be inferred by the name, we can think of the firms as urns in which workers place their applications, the balls. Coordination failure comes from the fact that workers send out their applications randomly which may lead to some firms receiving too many applications per vacancy while others receive none. Hence, the lack of information about the actions of other agents creates frictions in the labour market which prevents it from clearing.

Moreover, allowing for workers to send more than one application per period does not necessarily alleviate the problem and coordination failure might persist (see Albrecht et al. (2003)). Under the case of many applica-

[2] For a comprehensive review of the types of matching functions refer to Petrongolo and Pissarides (2001).

tions per worker per period, it is more likely that all vacancies will receive an application. Nevertheless, this does not ensure that a match will be formed since now firms are competing for the best workers, which in our set-up are the most experienced ones. Hence, if a very good candidate has sent multiple applications s/he might receive several offers and reject most of them, leaving the vacancies open until the next period.

For each vacancy, the firm chooses the applicant who will generate the highest match surplus **S**, *i.e.* the one with the highest experience ($x_{i,t}$). The match surplus generated by an employed worker i at time t is an increasing, concave function of the worker's experience:

$$\mathbf{S} = \frac{x_{i,t}^{1-\sigma}}{1-\sigma}, \ \ \sigma > 0 \tag{1}$$

Throughout the simulations we set $\sigma = 0.5$. Moreover, we assume that worker's experience is not job-specific and is perfectly transferable between firms. That is, the worker will generate the same surplus whenever matched with either one of the firms. The wages of workers with the same experience, however, differ between firms depending on the employer's bargaining strategy β. Hence, the wage of worker i employed in firm j at time t can be written as:

$$w_{ijt} = \phi(x_{i,t}, \beta_j) = \beta_j \left(\frac{x_{i,t}^{1-\sigma}}{1-\sigma} \right) \tag{2}$$

It follows that wage dispersion within and across firms is generated. This mechanism provides an obvious incentive for the workers to search for an employer with whom they can earn a larger share of the surplus, given that workers with the same experience produce the same match surplus in all firms. Similarly to Burdett et al. (2011), we expect that most of the young workers will start their careers in low-paying firms and will change several employers in pursuit of higher remuneration.

The wages are normalized between 0 and 1 and a "minimum wage" of 0.01 is introduced. This is done in order to avoid a wage of 0 for workers with no experience. However, given the set-up of the model, the introduction of a minimum wage does not alter the market outcomes, since the only criteria for hiring workers is the experience.

The model is created in RePast and the used programming language is Java. The next two sections describe in more detail the initial conditions of the model as well as what happens in each period.

3.1. Initial Conditions

In the beginning of the simulation the agents are created. Since firms are equally sized, it must hold that the number of workers that could be employed in one firm is: $\frac{NumberofWorkers}{NumberofFirms}$.

Additionally, firms draw their bargaining strategy from a Gaussian distribution with mean 0.5 and standard deviation 0.1. Given that the bargaining strategy is defined as a fraction of the match surplus, it can take values in the range [0,1]. If a firm draws a value outside of that range, it disregards it and makes a new draw. Further, the worker's age ($a_{i,0}$) is drawn from an uniform distribution over:

$$A = \{18 + 0.25n \mid n = 0,..., 179\}$$

and his/her experience ($x_{i,0}$) is randomly drawn from the set:

$$\{0.25\text{n} \mid \text{n} = 0, ..., 4(a_{i,0} - 18)\}.$$

In the initialization of the model, all workers are unemployed, which implies that the first few periods of the simulation should be disregarded from the analysis. Note also that the model could be calibrated to shorter periods than a quarter of a year by adjusting the age updating process, which should decrease the unemployment duration. Throughout all simulations, however, a period of a quarter of a year is assumed.

3.2. Step Dynamics

In the beginning of each period firms post vacancies if they have not already filled their positions. Workers, both employed and unemployed, follow their search strategy and send a number of applications for the open vacancies accordingly. Workers can observe the mean of the bargaining strategies across all firms but do not know *ex-ante* what is the bargaining strategy of the particular firm for which they apply.

Next, each firm gathers the list of applicants and sorts it according to workers' experience. Then, the firm chooses the most experienced applicant and offers the position to him/her. If the firm has more then one vacancy, it chooses the second most experienced applicant and offers the next vacancy to him/her and so on. Upon making an offer, the firm reveals its bargaining strategy to the worker.

At this point some workers might have received job offers and they proceed to decide whether to accept them or not. If a worker has sent out more than one application, s/he might have also received more than one offer. An unemployed worker then sorts the job offers according to the bargaining strategy of the firms. S/he chooses the employer with the high-

est β, *i.e.* the one who is willing to offer the biggest share of the match surplus to the employee and accepts the offer. If the unemployed worker has received only one offer, s/he takes it. In this model unemployed workers have high incentive to accept even "bad" offers because of the on-the-job search mechanism. Workers with higher experience are valued more by the firms and therefore applicants have no incentive to remain unemployed in order to wait for better offers.

Employed workers, on the other hand, also sort their job offers according to the firm's bargaining strategy. They next compare the best offer they have received with their current employers' β and change jobs if $\beta_{current}<\beta_{offer}$. Otherwise, they remain in their old position.

At the end of each period workers' age ($a_{i,t}$) is updated:

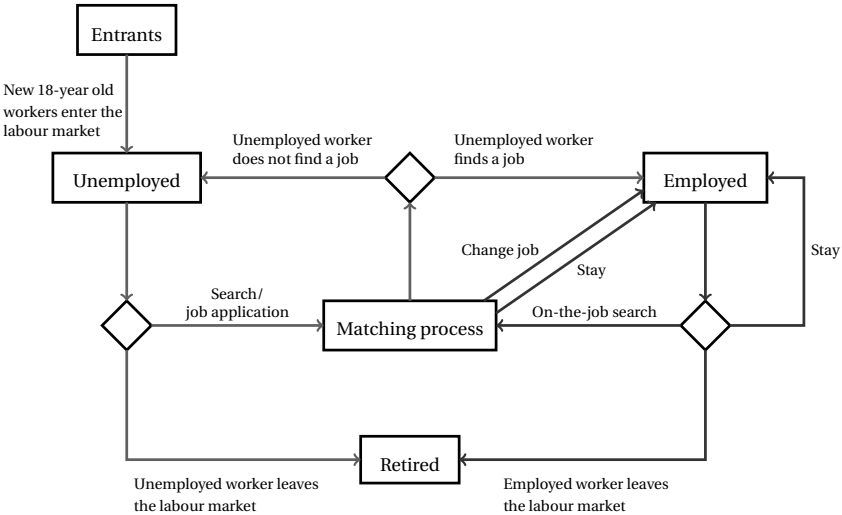
$$a_{i,t+1} = a_{i,t} + \frac{1}{4} \tag{3}$$

Whenever a worker reaches the age 63, s/he retires and a new 18-year-old unemployed worker with no experience is "born". Hence, the number of workers on the market is kept constant. Also the experience of the workers ($x_{i,t}$) is updated at the end of the period according to:

$$x_{i,t+1} = \begin{cases} x_{i,t} + \frac{1}{4}, \text{if worker } i \text{ was employed in period } t \\ \\ x_{i,t}, \text{ otherwise.} \end{cases} \tag{4}$$

Finally, data is collected and the period ends. Figure 1 depicts the described process and the next subsection presents a pseudocode of the model. In figure 1 the red arrows show the possible transitions of unemployed agents in the given period, while the blue arrows follow the possible paths of employed agents.

Figure 1: **Labour market model with on-the-job search**



3.3. Pseudocode

The following pseudocode presents in detail what happens in each period of the simulations.

1 period = 1 quarter
Initial conditions:
Create **n** infinitely-lived firms and **m** workers
Workers are allocated age ($a_{i,0}$), drawn from the set: $A = \{18 + 0.25n \mid n = 0,..., 179\}$ and experience level ($x_{i,0}$) drawn from the set $\{0.25n \mid n = 0,..., 4(a_{i,0} - 18)\}$
Firms are allocated a bargaining strategy β
Workers are assigned to the pool of unemployment
Dynamics:
for each period t
  for each firm j
    if number of workers ($numWor_j$) < $\frac{m}{n}$
      post $\frac{m}{n} - numWor_j$ vacancies
    else
      do nothing
  end for each firm j
  for each worker i
    if worker is unemployed
      observes all firms with $numWor_j < \frac{m}{n}$
      applies randomly to g of them according to the search strategy

```
    if worker is employed in firm with bargaining strategy β < β_mean
        observes all firms with numWor_j < m/n
        applies to k firms according to the search strategy
    if worker is employed in firm with bargaining strategy β > β_mean
        observes all firms with numWor_j < m/n
        applies to l firms according to the search strategy
end for each worker i
Hiring
for each vacancy of firm j
    if workers apply
        make an offer to the one with highest experience x_{i,t}
        if two or more applicants have the same experience, make an offer to one
        of them at random
    if no workers apply
        vacancy remains open
end for each vacancy of firm j
Job transition
for each worker i
    if unemployed worker receives job offers
        chooses the firm with highest β
    else
        remains unemployed
    if employed worker receives job offers
        changes firms if β_offer > β_current
    else
        remains with current employer
end for each worker i
Market update
for each worker i
    update age
    if age = 63
        worker is replaced by a new unemployed worker with age=18 and
experience=0
    else
        worker remains on the labour market
    if worker is employed
        update experience
    else
        keep old experience
end for each worker i
```

Data collection at end of each period

end for each period t

4. Search Intensity

In this set-up, we address the question: how the search intensity affects the market outcome. Since we consider infinitely-lived firms which never lay off workers, the only options for a separation between a firm and a worker arise from the job-to-job transition of workers and from retirement. Hence, once a worker finds a job, s/he remains employed until retirement. Within the framework of the model we explore the link between search intensity and the average duration until finding the first job and between search intensity and the unemployment rate.

4.1. Benchmark Case

In the benchmark case we assume, rather unrealistically, that each worker sends one application per period regardless of his/her employment status. Figure 2 displays a histogram of the wages (panel a) at the end of the 400th period (equivalent to 100 years) and the hazard rates during the 400 periods (panel b). The hazard rates provide an empirical estimate of the probability of moving into employment (job-finding rate) and the probability of changing jobs (job-to-job transition rate) at a given period t. The job-finding rate is calculated as the fraction of unemployed workers who manage to find a job until the end of the period. The job-to-job transition rate, on the other hand, is estimated as the fraction of employed workers who change jobs within the period.

Figure 3 shows the unemployment rate, average wage (panel a) and the average unemployment duration (panel b) generated by the simulation. We can observe that the model creates a high unemployment rate (the value stabilizes close to 30%) and a high average unemployment duration. The underlying reason for this result could be inferred from the hazard rates diagram (fig 2b). We see that the job-to-job transition rate is on average higher than the job-finding rate. Once workers become employed and start accumulating experience, they have an advantage over unemployed workers when being selected for positions. Since, in this scenario, workers search with the same intensity regardless of their employment status, it becomes easier for workers who are already employed to transition between firms then for the unemployed workers to find a job. This can be deduced from the algorithm of the matching process according to which firms offer one position to a single worker. If many workers applied the firm chooses the one with the highest experience who might then refuse the offer. Hence, no match is created for the period even though less experienced or unemployed workers might have been willing to accept the offer.

Figure 2: **Simulation results, single runs, one application *per* period: t = 400, 100 firms, 1000 workers**

(a) Histogram of the wages **(b)** Hazard rates

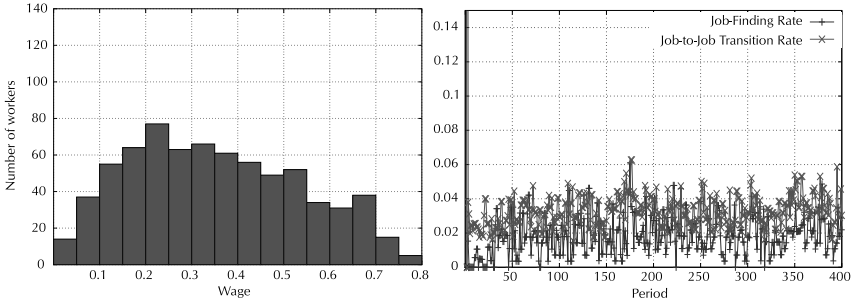


Figure 3: **Simulation results, single runs, one application *per* period: t = 400, 100 firms, 1000 workers**

(a) Unemployment and average wage **(b)** Average unemployment duration

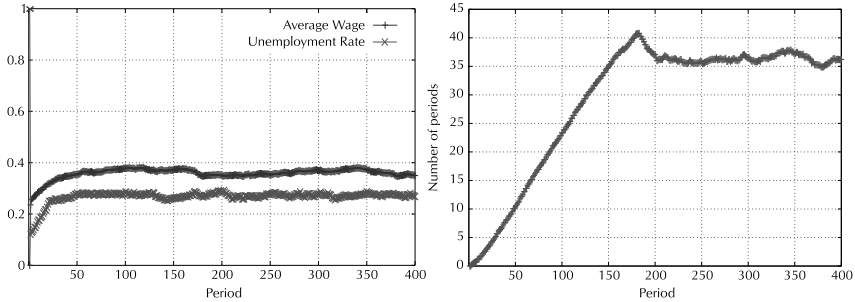
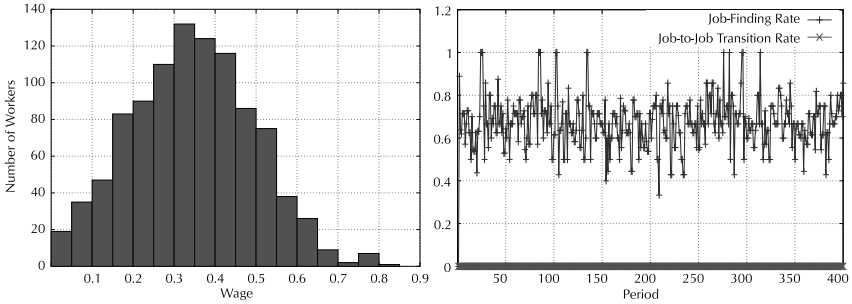


Figure 4: **Simulation results, single run, no on-the-job search: t = 400, 100 firms, 1000 workers**

a) Histogram of the wages **(b)** Hazard rates



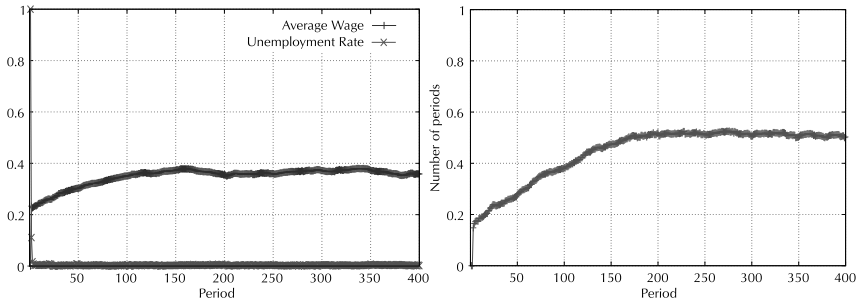
4.2. No On-the-job Search, 1 Application Per Period

To illustrate this further, in the next experiment we forbid the search while employed. Without on-the-job search, the model generates almost full employment. Since once the worker finds a job, s/he remains employed until retirement, the lack of job-to-job transitions, job destruction and competition coming from more experienced workers, create prerequisite for achieving almost full employment. On average, workers manage to find work within their first period of unemployment which implies very low degree of search frictions. The results are displayed in figures 4 and 5.

We can observe a high job-finding rate which fluctuates around 0.65. A value of 0.65 for the job-finding rate means that in the given period 65% of the unemployed agents have found a job. Comparing the hazard rates under the no job-to-job transition case with the result in the previous section where all workers send one application per period, irrespective of their employment status (see fig. 2b), we can see how much the chances of unemployed workers worsen when on-the-job search is allowed. The value of the job-finding rate in that case fluctuates between 0 and approximately 0.06, which indicates that in the "best" period of the simulation only about 6% of the unemployed agents moved into employment.

Figure 5: **Simulation results, single run, no on-the-job search: t = 400, 100 firms, 1000 workers**

(a) Unemployment and average wage **(b)** Average unemployment duration



If we look at the wage histograms generated from the two scenarios we can observe that the histogram from the simulation with on-the-job search is flatter, with relatively large concentration of workers with highest and lowest wages (fig. 2a), whereas when we forbid workers to change jobs, the number of workers seems more concentrated around the average wage value (fig. 4a). The underlying reason behind this result comes from the fact that in the no on-the-job search case almost all workers manage to find a

job fast and begin accumulating experience. Since the wage in this model is an increasing function of the experience, most workers manage to generate enough experience so that to reach a salary close to the average, even if they are in a firm which has a low Nash bargaining coefficient.

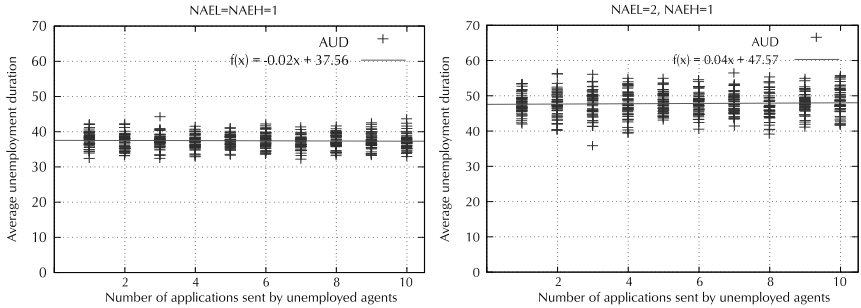
On the other hand, the case with on-the-job search allows experienced workers to move up the wage ladder relatively fast and find employment in the high-paying firms, *i.e.* the most experienced workers who will have high salaries in any firm, manage to also move to the firms with the highest Nash bargaining coefficients. This might explain the fatter right tail of the wage histogram with on-the-job search (fig. 2a). The reason why there is also a relatively high concentration of workers with the lowest salaries in this case could possibly be due to the fact that once workers begin accumulating experience, they want to move to better firms. Hence the firms with lowest Nash bargaining coefficients are constantly looking for new workers who inevitably come from the pool of unemployment, where workers have experience of 0. Since workers have no incentive to stay in the low-paying firms, such firms suffer higher worker turnover, and possibly need to constantly hire workers with no or very little experience who earn the lowest wages. This might explain why there is a relative concentration of workers in the lowest part of the histogram, whereas in the case with no on-the-job search there is a pronounced peak in the number of workers with wages closer to the mean.

This experiment shows how important on-the-job search is in this model and how it can create a very high degree of frictions and subsequently high unemployment rate and high average unemployment duration. In the next sections we explore further how different search intensities alter the results. For the purpose we run multiple simulation altering the search strategies of workers, whereby we will distinguish between unemployed workers, workers employed in “worse than average” firms and workers employed in “better than average firms”.

Figure 6: **Average unemployment duration *vs.* search intensity of unemployed agents**

(a) One application *per* employed worker *per* period

(b) One application per period *per* worker employed in a "good" firm, two applications *per* period *per* worker employed in a "bad" firm



4.3. On-the-job Search With Multiple Applications

For the simulations, we alter workers' search intensity, *i.e.* the number of applications they send per period. Unemployed workers send between 1 and 10 work applications. The assumption is that unemployed workers always apply for jobs. On the other hand, workers employed in both better and worse than average firms send between 0 and 10 applications per period. Starting with the first value of the parameters, they are incremented by one for each run. All possible combinations are replicated 50 times. There are 100 workers and 10 firms and firms can employ up to 10 workers. The values of the variables at the 250th period of each run are reported. 250th period is chosen based on the observation from the single runs that all variables have stabilized by that time.

Worse than average firms are the ones with bargaining strategy β lower than the mean of bargaining strategies across all firms while better than average firms are the ones with bargaining strategies above the mean. Workers can observe the mean and also know the value of the β of their current employer. Intuitively, we can think that workers who know that they are already in a "good" firm will reduce the number of applications they send per period or will stop searching altogether. This possibility is explored in section 6. For the moment we assign a search strategy for each worker and vary it exogenously, depending on whether or not the worker has a job and the type of firm s/he is employed in.

Figure 6 shows the result of the simulation runs for average unemployment duration of unemployed workers for different search intensity, holding the search intensity of employed workers constant. Table 1 summarizes the used abbreviations.

Table 1: **Variables**

Abbreviation	Description
aud	average unemployment duration
nau	number of applications sent by unemployed workers per period
nael	number of applications per period sent by workers employed in "worse than average firms"
naeh	number of applications per period sent by workers employed in "better than average firms"
jtj	job-to-job transition rate
jf	job-finding rate
avewage	average wage
nael_sq	nael squared
naeh_sq	naeh squared

From figure 6a we can see that, whenever the search intensity of employed workers is held constant, at one application per period, increasing the number of applications of unemployed workers has almost no effect on the average unemployment duration. The slope of the fitted line is negative but very close to zero. Repeating the same experiment, but this time increasing the number of applications of workers employed in worse than average firms to two per period, reveals similar pattern. The search intensity of the unemployed agents does not seem to have an impact on the average unemployment duration and in this case the slope of the fitted line is even positive (and again very close to 0).

Next, the output of the simulations is used in OLS regression models in order to study the correlation between workers' search intensity and average unemployment duration and unemployment rate, respectively. In table 2 the dependent variable in all 4 regression models is average unemployment duration, while in table 3 the dependent variable is the unemployment rate. Plotting the data revealed a curvilinear relationship between both the average unemployment duration and the unemployment rate and the number of applications sent by employed workers (nael and naeh), see Appendix A. Therefore, also a squared terms of the two variables is included.

Table 2: **Regression results, 50 replications of each case, dependent variable: average unemployment duration**

VARIABLES	(1) Model 1	(2) Model 2	(3) Model 3	(4) Model 4
nau	0.0343*** (0.0125)	0.0353*** (0.0117)	0.0383*** (0.0117)	0.0370*** (0.0106)
nael	1.889*** (0.0146)	1.755*** (0.0134)	1.718*** (0.0134)	5.220*** (0.0523)
naeh	3.157*** (0.0137)	3.046*** (0.0132)	3.005*** (0.0133)	5.102*** (0.0447)
jtj		15.84*** (0.836)	17.26*** (0.836)	4.872*** (0.731)
jf		-57.54*** (1.147)	-57.20*** (1.130)	-46.08*** (0.960)
avewage			11.67*** (0.633)	7.416*** (0.597)
nael_sq				-0.346*** (0.00442)
naeh_sq				-0.207*** (0.00396)
Constant	42.89*** (0.160)	44.20*** (0.156)	40.01*** (0.282)	33.21*** (0.293)
Observations	60,500	60,500	60,500	60,500
R-squared	0.635	0.676	0.678	0.735

Robust standard errors in parentheses

***p<0.01, ** p<0.05, * p<0.1

The search intensity of employed workers has a significant, positive impact on the unemployment rate. On the other hand, the search intensity of unemployed workers has a very little (coefficients close to 0, see table 3) and insignificant effect on the unemployment rate. This results indicates that higher search activity of employed workers increases the unemployment rate, while increasing the search intensity of the unemployed can do very little to counteract this effect. Moreover, the results in table 2 suggest that unemployed workers cannot reduce the time until finding their first job by increasing the number of applications sent per period. On the contrary, the variable "number of applications sent by unemployed workers" (nau) has a positive and significant effect on the average unemployment duration (table 2). Even though the estimated coefficients are statistically significant, they are close to 0 in all four specifications of the regression. This indicates that although statistically significant, this values are not economically significant. The high significance level could also be driven by the large number of observations generated by the simulations.

Table 3: **Regression results, 50 replications of each case, dependent variable: unemployment rate**

VARIABLES	(1) Model 1	(2) Model 2	(3) Model 3	(4) Model 4
nau	7.27e-05 (0.000127)	8.17e-05 (0.000120)	0.000154 (0.000118)	0.000139 (0.000104)
nael	0.0236*** (0.000150)	0.0222*** (0.000140)	0.0213*** (0.000137)	0.0616*** (0.000522)
naeh	0.0423*** (0.000139)	0.0414*** (0.000136)	0.0404*** (0.000136)	0.0640*** (0.000443)
jtj		0.269*** (0.00842)	0.303*** (0.00826)	0.161*** (0.00676)
jf		-0.540*** (0.0127)	-0.532*** (0.0122)	-0.405*** (0.0103)
avewage			0.278*** (0.00649)	0.229*** (0.00593)
nael_sq				-0.00397*** (4.35e-05)
naeh_sq				-0.00233*** (3.91e-05)
Constant	0.365*** (0.00162)	0.375*** (0.00161)	0.275*** (0.00293)	0.198*** (0.00292)
observations	60,500	60,500	60,500	60,500
R-squared	0.743	0.770	0.776	0.826

Robust standard errors in parentheses

***p<0.01, **p<0.05, * p<0.1

Nevertheless, one possible reason why there is a positive correlation between the average unemployment duration and the search intensity of unemployed workers might also stem from a congestion effect that unemployed workers create by being "too active" in their search. If a worker with no experience is lucky enough to be selected by more than one firm at a given period s/he is preventing another unemployed worker from finding a job. However, the search intensity of employed workers has a far higher impact on the average unemployment duration which can be seen by comparing the coefficients of the independent variables in table 2 (nau *vs*. nael and naeh). The result suggests that on-the-job search considerably worsens the market outcome for unemployed workers and that increasing the search intensity of young and inexperienced workers cannot counterbalance this effect.

4.4. Single Runs with Different Distributions over Workers' Search Intensity

A further experiment that could be conducted with the model is to see how the unemployment rate and the average unemployment duration change with different choices of distributions over workers' search intensity. In the presence of empirical data on this distribution, the model could be calibrated to closer simulate the effect of workers' search intensity on the market outcomes. For now, we consider different theoretical distributions over the search strategies of the workers and compare the results. For this purpose we assume that whenever the worker is "born" s/he draws his/her search intensity strategy from some discrete distribution and sticks to this strategy until retirement. As above, the search strategy is defined as the number of applications the worker sends per period.

We assume a uniform distribution over the search strategy as the benchmark case and then also test the outcomes under Poisson and Binomial distributions. In the first scenario "newborn" workers draw randomly a search strategy from the interval {1,2}. In this case there are only two possible search strategies: low- and high search intensity. We compare this set-up with a scenario in which we add one more option for search strategy, *i.e.* new workers draw a their strategy from the interval {1,2,3}.

For the Poisson distribution also two scenarios are considered. In the first case, young workers draw their search strategy from a Poisson distribution with mean $\lambda = 1$, while in the second case we consider a mean of $\lambda = 2$. It must be noted, however, that in both cases a search strategy of 0 applications per period is possible, hence the unemployment rate and average unemployment duration might be higher due to workers whose strategy is to not participate in the labour market.

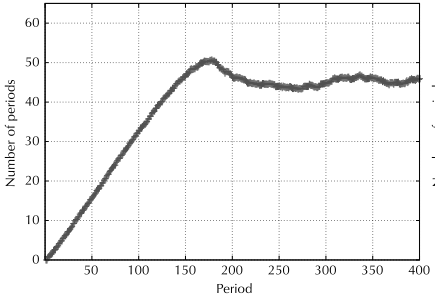
Last but not least, we consider a binomial distribution with parameters $n = 10$, $p = 0.1$ and $n = 10$, $p = 0.2$. This discrete probability distribution is generated as the outcome of 10 failure/success experiments, where the probabilities of success in each experiment is given by the parameter p. Once again, the probability of 0 as a search strategy is not ruled out.

The following figures displays the outcomes for single runs under the different scenarios. Figure 7 shows the average unemployment duration during the 400 periods of the single runs in the six considered cases, figure 8 depicts the unemployment rate and the average wage, figure 9 displays the hazard rates and figure 14 in the appendix shows the wage histograms. What we observe can be inferred also from the results in the previous section: perhaps a bit counter-intuitively, distributions with higher mean,

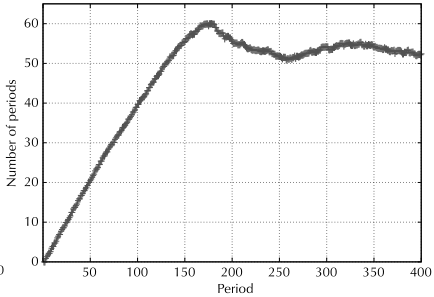
i.e. the ones that generate larger search intensity, also lead to higher unemployment and higher average unemployment duration. This result seems robust to the choice of distribution over the search intensities.

Figure 7: **Average unemployment duration, simulation results, single runs, 100 firms, 1000 workers, t = 400**

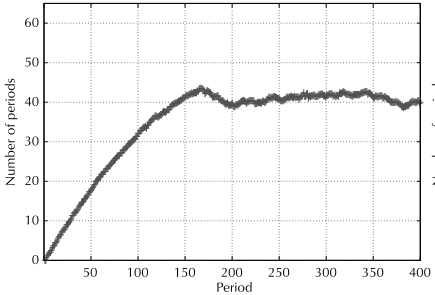
(a) Uniform distribution in the interval [1,2]



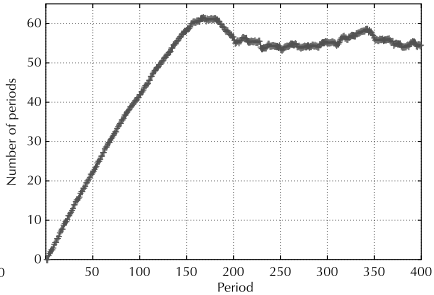
(b) Uniform distribution in the interval [1,3]



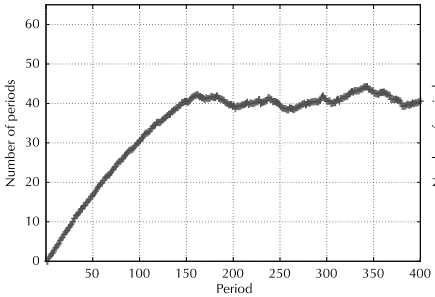
(c) Poisson distribution with λ = 1



(d) Poisson distribution with λ = 2



(e) Binomial distribution with n = 10, p = 0.1



(f) Binomial distribution with n = 10, p = 0.2

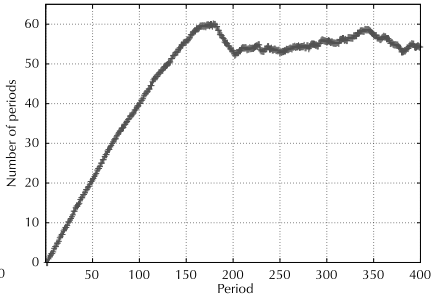
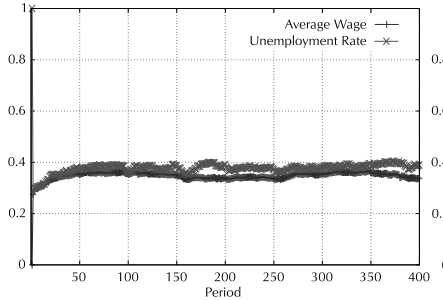
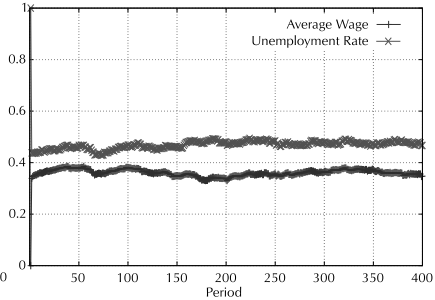


Figure 8: **Unemployment and average wage, simulation results, single runs, 100 firms, 1000 workers, t = 400**

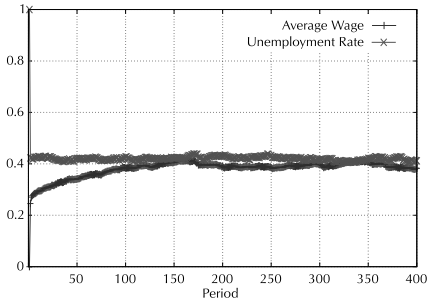
(a) Uniform distribution in the interval [1,2]



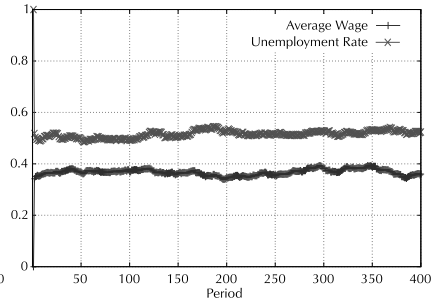
(b) Uniform distribution in the interval [1,3]



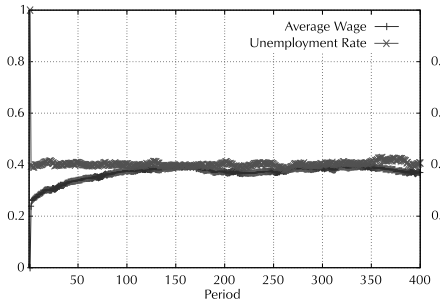
(c) Poisson distribution with $\lambda = 1$



(d) Poisson distribution with $\lambda = 2$



(e) Binomial distribution with $n = 10$, $p = 0.1$



(f) Binomial distribution with $n = 10$, $p = 0.2$

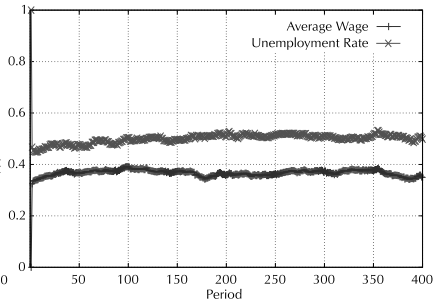
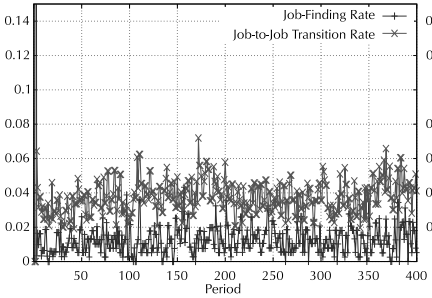
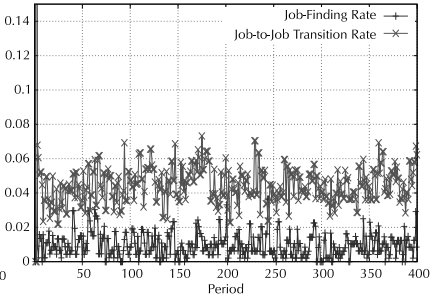


Figure 9: **Job-finding rate and job-to-job transition rate, simulation results, single runs, 100 firms, 1000 workers, t = 400**

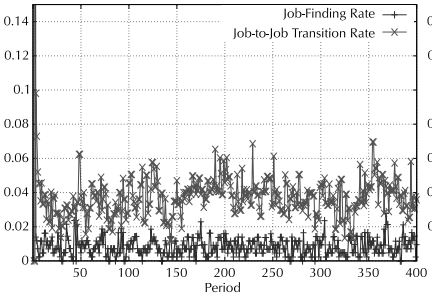
(a) Uniform distribution in the interval [1,2]



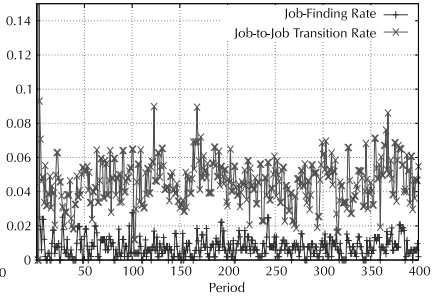
(b) Uniform distribution in the interval [1,3]



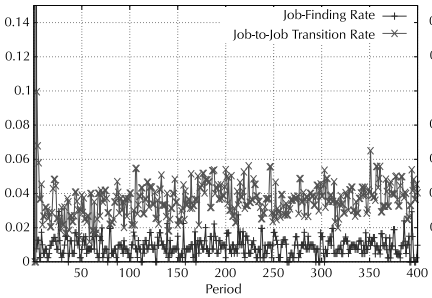
(c) Poisson distribution with λ = 1



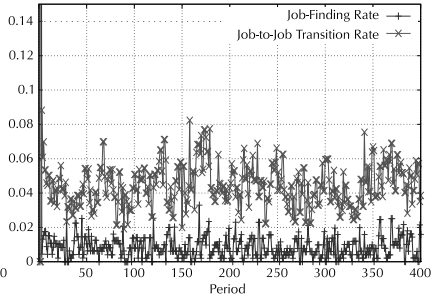
(d) Poisson distribution with λ = 2



(e) Binomial distribution with n = 10, p = 0.1



(f) Binomial distribution with n = 10, p = 0.2



5. Adaptive Behaviour

Last but not least, we assume that workers do not stick with their search strategy throughout their life on the labour market. Instead, workers who are unsuccessful in finding a first job can increase their search intensity. This is modelled in the following way: all initial workers draw their search intensity from a uniform distribution in the interval [1,3]. If a worker did not manage to find a job in the current period, s/he increases his/her search intensity by sending an additional application in the next period. An unemployed workers might increase his/her search intensity up to 20 applications per period if s/he is not successful in finding a position. For modelling the search strategy of employed workers, we rely on the empirical findings, according to which workers who earn high wages are less likely to search on-the-job (see section 2).

Since the wage in the model is an increasing function of the firm's bargaining strategy β, the probability that an employed worker sends out applications in a given period is decreasing in the bargaining strategy of the worker's current employer. More precisely, in each period an employed worker in firm j refrains from applying for new jobs with a probability βj, where βj was the Nash bargaining coefficient of firm j. The higher the βj, the lower the probability that the worker will search for a new job in the current period. On the other hand, with probability $(1 - \beta j)$ the employed worker sends out between 1 and 3 applications in the given period, where s/he draws the search intensity from an uniform distribution. In each period employed workers who continue their search make a new draw from the distribution. Hence, agents who already have a job will not necessarily search in each period and the ones that are employed in the best firms will be less likely to apply for new positions.

The results from the simulation are displayed in figures 10 and 11. We can compare them to the case in which new workers draw their search strategy randomly from the interval [1,3] and stick to it until retirement (figures 7b, 8b, 9b, 14b). While in the former case, the unemployment rate and the average unemployment duration fluctuate around 45–50% and 55 periods, respectively, under adaptive behaviour both variables drop significantly. The unemployment rate fluctuates between 25 and 30% and the average unemployment duration moves between 35 and 40 periods.

One of the underlying reasons for this result is that inexperienced workers do not compete with all employed workers in each period which gives

them a better chance to get a job offer. Moreover, unemployed workers have the ability to increase their search which also raises the probability that they will send an application to a firm that has not received applications from experienced workers. Comparing the hazard rate figures, 9b and 11a, is indeed indicative of the described phenomenon. While in the former, the job-to-job transition rate lies above the job-finding rate in almost all periods, in the latter we can observe quite a few periods in which the fraction of unemployed workers who moved into employment is higher than the fraction of employed workers who changed jobs. Nevertheless, the unemployment rate remains fairly high.

Figure 10: **Simulation results, single run: t = 400, 100 firms, 1000 workers**

(a) Histogram of the wages **(b)** Average unemployment duration

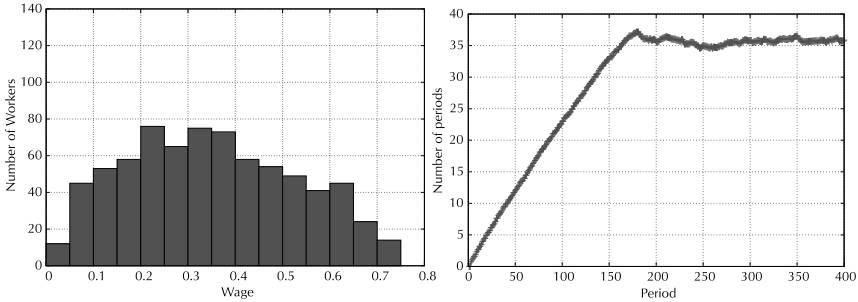
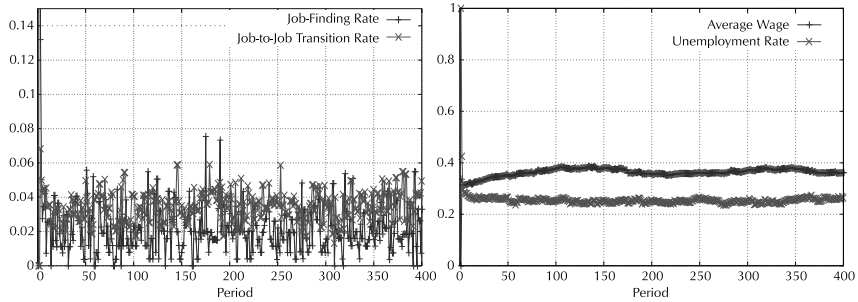


Figure 11: **Simulation results, single run: t = 400, 100 firms, 1000 workers**

(a) Hazard rates: Job-finding rate and job-to-job transition rate **(b)** Unemployment rate and average wage



6. Concluding Remarks and Possible Extensions of the Model

The model presented in this paper hints towards the possibilities that the agent-based methodology can offer as a modelling tool and is a first step towards creating a fuller labour market model with more realistic features. We investigated how the search intensity of both employed and unemployed workers affects the market outcomes in terms of unemployment and duration until finding a first job. Perhaps a bit counter-intuitively, the model suggests that very high search intensity is "bad" for the market since it creates a lot of congestion and increases the unemployment duration. Moreover, if employed workers are very active in searching for positions in better-paying firms, it takes considerably longer for the young and inexperienced workers to find a job. In this set-up of the model high search intensities aggravate the coordination failure in the market.

Under no on-the-job search scenario, the model creates almost full employment. It takes very few periods for new workers to get hired because they do not face competition from older workers. When we allow for on-the-job search, however, it becomes increasingly more difficult for the inexperienced, young workers to find a job since they are never preferred over a workers who has at least a little bit of work experience. Moreover, increasing worker's search intensity, *i.e.* allowing for multiple application per worker per period leads to very high degree of frictions caused by coordination failure on both sides of the labour market. Employed workers do not know which are the "better" firms and send applications randomly, hence blocking the entrance of young workers into the market and on the other hand, firms are competing for the best workers which leaves some vacancies unfilled due to the fact that firms make offers which are ultimately rejected.

There are several directions in which the current model can be augmented and made more realistic. First of all, in its current state the model generates unemployment rate that is too high compared to real-world data. There are several reasons for the high level of unemployment generated in the simulations which in further stages of the model can be corrected.

The first one lies in the assumption that the experience of workers is perfectly transferable between firms. Hence, the on-the-job search mechanism leaves little chance for the new workers to enter the market. If the older, more experienced workers are too active in their search, the young and inexperienced ones "get stuck" on the queue of applicants and rarely receive an offer. If there is some kind of penalty for switching jobs in terms

of loss of experience, this effect might be reduced. Also, including a search cost in terms of loss of leisure time should conceivably reduce the search intensity of employed workers under adaptive behaviour because their spare time is already scarce in comparison to that of unemployed workers.

Moreover, we can consider a two-step offering process, in which firms who have received a rejection in the current period, make an offer to the next best applicant from their application list. In this case the model might be brought closer to reality since usually firms would not wait until the next quarter to make a new offer for their vacancy. Such mechanism is expected to generate more plausible outcomes. It can be also extended to a multi-step process in which firms continue offering the position to next best applicant upon rejection.

Last but not least, we can consider the possibility to fire workers and to incorporate firm exit. Both of which would further augment the model and introduce the possibility to analyse various array of research questions.

Nevertheless, in its current version, the model provides some insight into two frictions that act simultaneously on the labour market. Not only are workers competing for the available positions but also firms are competing for the best workers which makes the formation of an actual match more difficult. On the one hand, a coordination failure arises from the urn-ball type matching in which workers send out their applications randomly. Hence, some firms receive more applications than vacancies they have while others receive less, which prevents the market from clearing. And on the other hand, allowing for multiple applications and on-the-job search introduces a further source of friction in the labour market which is shown to substantially increase the unemployment rate and the duration until finding a first job.

References

Albrecht, J. W., IP A. Gautier and S. B. Vroman (2003). Matching with multiple applications. *Economics Letters,* 78(1): 67–70.

Allen, J. and R. Van der Velden (2001). Educational mismatches versus skill mismatches: effects on wages, job satisfaction, and on-the-job search. *Oxford economic papers,* 53(3): 434–452.

Ballot, G. (2002). Modeling the labor market as an evolving institution: model artemis. *Journal of economic behavior & organization,* 49(1): 51–77.

Bontemps, C., J.-M. Robin and G. J. Van den Berg (2000). Equilibrium search with continuous productivity dispersion: Theory and nonparametric estimation. *International Economic Review,* 41(2): 305–358.

Burdett, K., C. Carrillo-Tudela and M. G. Coles (2011). Human capital accumulation and labor market equilibrium. *International Economic Review,* 52(3): 657–677.

Christensen, B. J., R. Lentz, D. T. Mortensen, G. R. Neumann and A. Werwatz (2005). On-the-job search and the wage distribution. *Journal of Labor Economics,* 23(1): 31–58.

Deissenberg, C., S. Van Der Hoog and H. Dawid (2008). Eurace: A massively parallel agent-based model of the european economy. *Applied Mathematics and Computation,* 204(2): 541–552.

Fallick, B. and C. A. Fleischman (2004). Employer-to-employer flows in the us labor market: The complete picture of gross worker flows.

Gemkow, S. and M. Neugart (2011). Referral hiring, endogenous social networks, and inequality: an agent-based analysis. *Journal of Evolutionary Economics,* 21(4): 703–719.

Guerrero, O. A. and R. L. Axtell (2013). Employment growth through labor flow networks. *PloS one,* 8(5): e60808.

Jolivet, G., F. Postel-Vinay and J.-M. Robin (2006). The empirical content of the job search model: Labor mobility and wage distributions in europe and the us. *European Economic Review,* 50(4): 877–907.

Lewkovicz, Z., D. Domingue and J.-D. Kant (2009). An agent-based simulation of the french labour market: studying age discrimination. In *The 6th Conference of the European Social Simulation Association, eds, Bruce Edmonds and Nigel Gilbert, (9 2009).*

Mattila, J. P (1974). Job quitting and frictional unemployment. *The American Economic Review,* 64(1): 235–239.

Mortensen, D. T. (1994). The cyclical behavior of job and worker flows. *Journal of Economic dynamics and control,* 18(6): 1121–1142.

Neal, D. and S. Rosen (2000). Theories of the distribution of earnings. *Handbook of income distribution,* 1: 379–427.

Neugart, M. (2004). Endogenous matching functions: An agent-based computational approach. *Advances in Complex Systems,* 7(02): 187–201.

Neugart, M., M. Richiardi et al. (2012). Agent-based models of the labor market. *LABORatorio R. Revelli working papers series,* 125.

Petrongolo, B. and C. A. Pissarides (2001). Looking into the black box: A survey of the matching function. *Journal of Economic Literature,* 39(2): 390–431.

Pissarides, C. A. (1994). Search unemployment with on-the-job search. *The Review of Economic Studies,* 61(3): 457–475.

Pissarides, C. A. and J. Wadsworth (1994). On-the-job search: some empirical evidence from britain. *European Economic Review,* 38(2): 385–401.

Richiardi, M. (2004). A search model of unemployment and firm dynamics. *Advances in Complex Systems,* 7(02): 203–221.

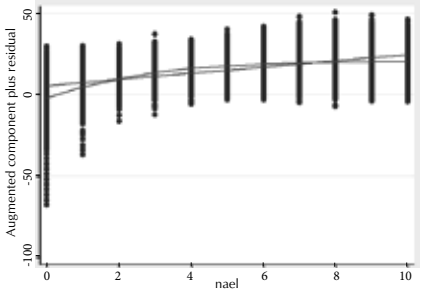
Richiardi, M. (2006). Toward a non-equilibrium unemployment theory. *Computational Economics,* 27(1): 135–160.

Tesfatsion, L. (2006). Agent-based computational economics: A constructive approach to economic theory. *Handbook of computational economics,* 2: 831–880.

A. Appendix

Figure 12:

(a) Average unemployment duration (aud) and Number of applications per period sent by workers employed in "worse than average firms" (nael)



(b) Average unemployment duration (aud) and Number of applications per period sent by workers employed in "better than average firms" (naeh)

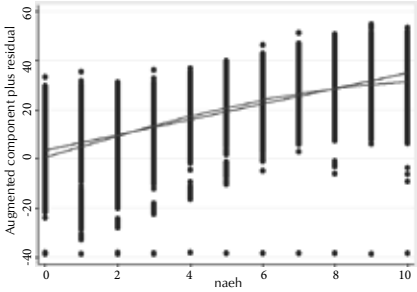
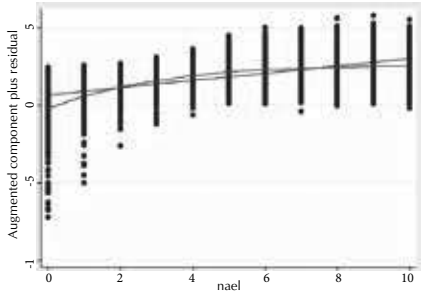


Figure 13:

(a) Unemployment rate (ur) and Number of applications per period sent by workers employed in "worse than average firms" (nael)



(b) Unemployment rate (ur) and Number of applications per period sent by workers employed in "better than average firms" (naeh)

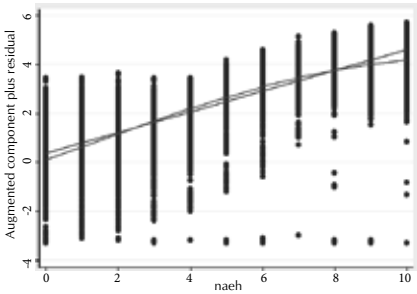
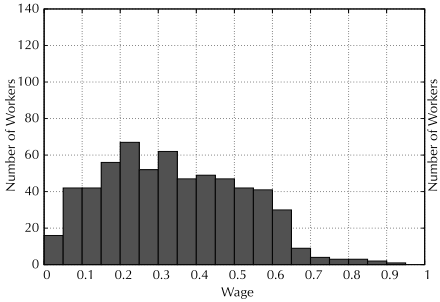
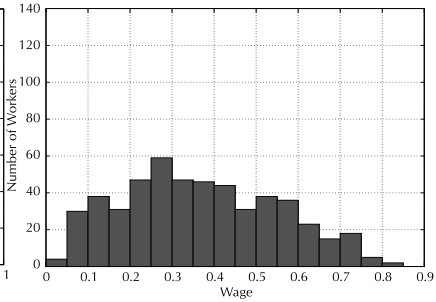


Figure 14: **Wage distribution at the 400th period, simulation results, single runs, 100 firms, 1000 workers, t = 400**

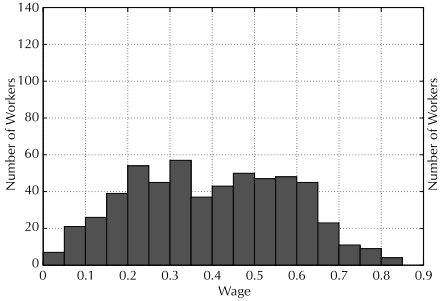
(a) Uniform distribution in the interval [1,2]



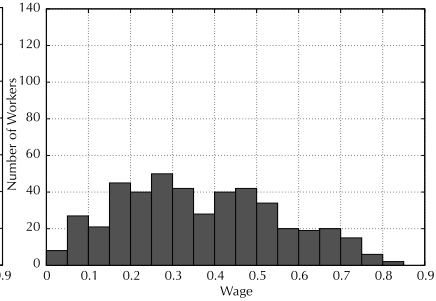
(b) Uniform distribution in the interval [1,3]



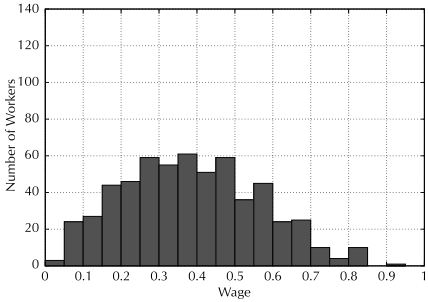
(c) Poisson distribution with λ = 1



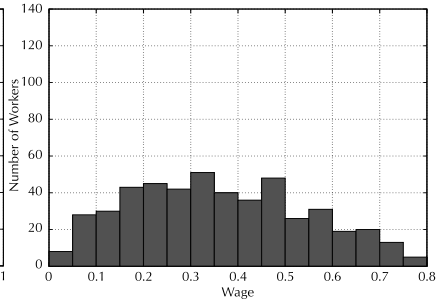
(d) Poisson distribution with λ = 2



(e) Binomial distribution with n = 10, p = 0.1



(f) Binomial distribution with n = 10, p = 0.2



DP/107/2017

Неравновесен модел на кредитния пазар в България

Михаил Михайлов

Резюме. Целта на това изследване е с помощта на неравновесен модел на кредитния пазар да оцени функции на търсене и предлагане на различните видове кредити в България, като въз основа на резултатите се идентифицират периодите на по-голямо несъответствие между търсенето и предлагането на заемни ресурси. Като цяло резултатите показват, че въпреки някои временни несъответствия е налице висока степен на корелация между търсенето и предлагането на кредити. Бързото нарастване на кредита в България през периода 2004–2008 г. е следствие от активизирането в еднаква степен на търсенето и предлагането на заемни ресурси. Аналогично, резкият спад в обема на новоотпуснатите кредити в периода на глобалната криза и в последвалия период на бавно икономическо възстановяване е обусловен от отслабване както на търсенето, така и на предлагането на кредити.

Abstract. Using disequilibrium model of the credit market, this study aims at modelling supply and demand functions of various categories of loans in Bulgaria and based on that at identifying periods of significant gap between the supply and demand of loans. Although some temporary gaps were identified, the results suggest that in general there is a high degree of correlation between the supply and demand of loans. Increases of loan demand and loan supply underpinned to the same extent the rapid credit growth in Bulgaria in the period 2004-2008. The sharp decline in the amount of newly extended loans during the global crisis and the subsequent period of slow economic recovery was similarly driven by both lower demand and supply of credit.

Михаил Михайлов, д-р по икономика, е главен експерт в БНБ,
e-mail: mihaylov.m@bnbank.org

Съдържание

Въведение .....
I. Характеристики на кредитния пазар в България .....
II. Обзор на изследвания с оценени функции
на търсене и предлагане на кредитния пазар.....
III. Използвани данни.....
IV. Неравновесни модели за кредитния пазар
в България .....
Заключение.....
Приложения.....
1. Тестове за единичен корен
и коинтеграция .....
2. Оценки на търсенето и предлагането
на различните видове кредити .....
Използвана литература .....

Въведение

Кредитната дейност на банките често е обект на анализ в икономическата литература. Изследването на свързаните с нея въпроси е от особено значение за страните, в които финансовата система е доминирана от банковия сектор, а предоставяните от банките кредити са основният източник на заемни ресурси за предприятията и домакинствата. Изучаването на факторите, влияещи върху търсенето и предлагането на банкови кредити, както и разглеждането на кредитната дейност в контекста на общото състояние на икономическата среда, са важни заради значителното отражение, което динамиката на кредита оказва върху икономическата активност и финансовата стабилност.

За времето от 2000 г. досега динамиката на кредитната дейност в България премина през няколко фази: от период на потисната кредитна активност в годините до 2004 г., през няколко години на прекомерно висок растеж на кредита в периода 2004–2008 г., до спад на кредитния растеж по време на глобалната криза и последвалия период на бавно икономическо възстановяване. Целта на това изследване е да даде количествена оценка на нивата на търсене и предлагане на кредити, като същевременно потърси отговор на въпроса, каква е степента на евентуално несъответствие между тях. За целта са оценени отделни функции на търсенето и предлагането на заемни ресурси по видове кредити, като използваните в литературата методи са адаптирани в съответствие с характеристиките на местната икономика и банкова система. Анализът е проведен в рамките на неравновесен модел на кредитния пазар, който се базира на допускането, че наблюдаваните пазарни обеми отразяват по-ниската от стойностите на търсенето и предлагането на заемни ресурси. По-конкретно, превишение на търсенето над предлагането на кредити би могло да се наблюдава при негъвкавост на лихвения процент по кредитите във възходяща посока, дължаща се на нежелание на банките да увеличат предлагането на кредити дори при повишение на лихвеното ниво заради наличието на асиметрична информация и риска по-високият лихвен процент да привлече ненадеждни кредитополучатели. Периодите, в които предлагането на кредити превишава търсенето на заемни ресурси, също заслужват внимание заради потенциалния риск високата склонност на банките да разширяват кредитната дейност в тези периоди да не отговаря напълно на условията в икономиката и кредитоспособността на икономическите агенти.

Анализът е структуриран в четири части. В първата част е направен дескриптивен анализ на структурните и цикличните характеристики на кредитния пазар в България. Във втората част е предложен обзор на изследвания, в които са оценени отделни функции на търсенето и предлагането на кредити. Третата част представя източниците на използваните за целите на изследването данни, а в четвъртата е оценен неравновесен модел за всеки един от сегментите на кредитния пазар в България. В заключението са обобщени резултатите от анализа и са очертани насоките за използване на резултатите.

I. Характеристики на кредитния пазар в България

При описанието на структурните и цикличните характеристики на кредитния пазар в България от ключово значение е изборът на период, върху който се базира анализът. Това изследване се фокусира върху времето от началото на 2000 г. до края на 2016 г., като причините за този избор са две. От една страна, периодът е достатъчно дълъг, за да позволи формулирането на заключения относно процесите на кредитния пазар в различните фази на икономическия и финансовия цикъл. От друга страна, този период се отличава с относително стабилни структурни характеристики. Независимо от някои изменения в структурата на банковата система като например консолидацията на кредитни институции и обявяването в несъстоятелност на Корпоративна търговска банка през 2014 г., като цяло през периода не се наблюдават промени, които биха поставили под въпрос валидността на резултатите и възможностите за тяхното практическо използване.

Една от особеностите на средата, в която банките в България осъществяват дейността си, е функционирането на въведения през 1997 г. паричен съвет. Той оказва влияние върху кредитния пазар в страната по два начина. Първо, отстранявайки несигурността, свързана с колебанията на курса на резервната валута, паричният съвет е фактор за стабилността на макроикономическата среда, като по този начин допринася за удължаването на инвестиционния хоризонт на икономическите агенти. От своя страна това влияе благоприятно върху търсенето и предлагането на заемни ресурси. Второ, заради фиксирания курс на лева към еврото и свободното движение на капитали паричният съвет осигурява тясна връзка на монетарните условия в България с тези в еврозоната, като възможните отклонения се свеждат само до различията в рисковите характеристики на средата.

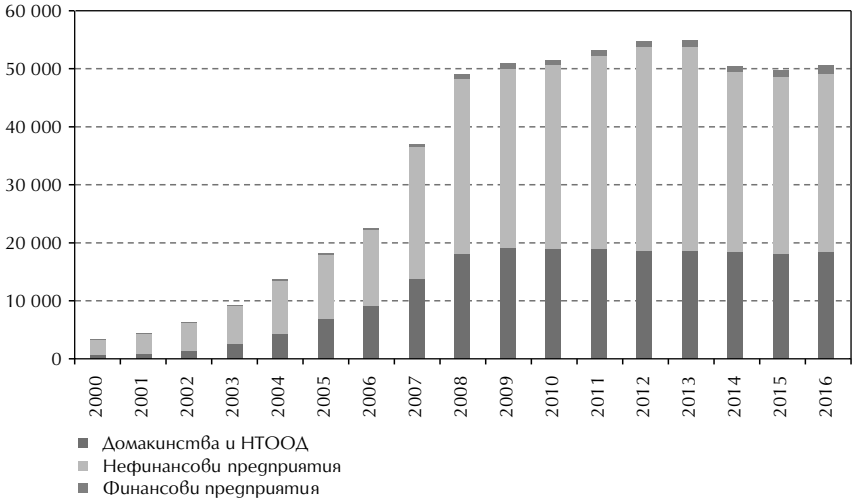
Това е фактор, който се отразява върху степента на синхронизация на цикличните колебания в България с тези в еврозоната, включително и върху тенденциите в кредитната активност. Друга характеристика на средата е, че кредитните институции са обект на консервативна макропруденциална политика, провеждана от БНБ. Тази политика се отличава с по-високи капиталови и ликвидни изисквания към банките в сравнение с прилаганите в други страни, като същевременно има антицикличен характер и варира в различните фази на кредитния и икономическия цикъл. Например, идентифицирането на зараждащи се дисбаланси на кредитния пазар в периода 2004–2008 г. наложи предприемането на стъпки за ограничаване на кредитния растеж. Обратно, след като глобалната криза засегна българската икономика, регулаторната политика бе смекчена в известна степен, като същевременно бяха взети мерки за засилване на устойчивостта на банките спрямо негативни процеси в обкръжаващата ги среда.

Кредитният пазар в България се характеризира с преобладаващ дял на кредитите за нефинансови предприятия. В началото на разглеждания период този дял е около 80%, като след 2005 г. с постепенното активизиране на кредитирането на домакинства той се понижава до около 60%. Делът на вземанията от домакинства нараства от около 20% в началото на периода до 37% към края на 2016 г., докато кредитите за финансови предприятия формират само около 2% от заемните ресурси, които местната банкова система предоставя на неправителствения сектор.

На фона на разликата в дяловете на кредитите за нефинансови предприятия и за домакинства в общия обем на кредитите за неправителствения сектор впечатление правят значителните различия в относителния дял на банковите кредити в структурата на привлечените от тях средства. Например, основният източник на заемни средства за нефинансовите предприятия са не кредитите от местната банкова система, а финансирането от чужбина, което формира повече от половината от привлечените от корпоративния сектор заемни ресурси и за това допринася силната отвореност на българската икономика. За разлика от предприятията домакинствата не разполагат с широк кръг алтернативни източници на финансиране, като почти изцяло използваните от тях заемни средства представляват кредити, получени от местните банки.

Графика 1. **Кредити за неправителствения сектор**

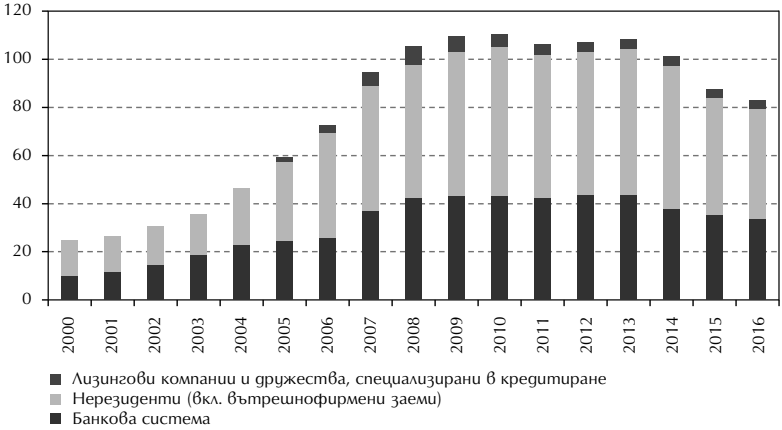
(млн. лв.)



Източник: БНБ, аналитична отчетност на други ПФИ.

Графика 2. **Задължения на нефинансовите предприятия по видове кредитори**

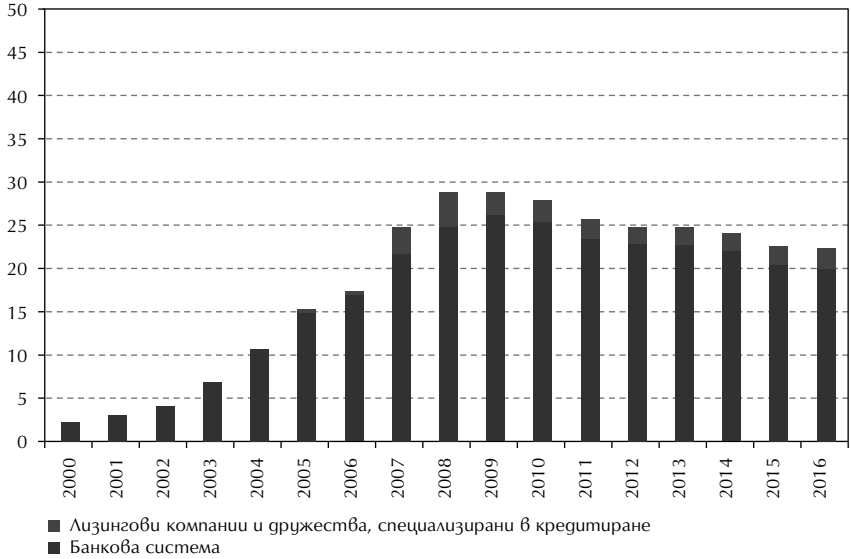
(% от БВП)



Източници: БНБ, НСИ.

Графика 3. **Задължения на домакинствата по видове кредитори**

(% от БВП)

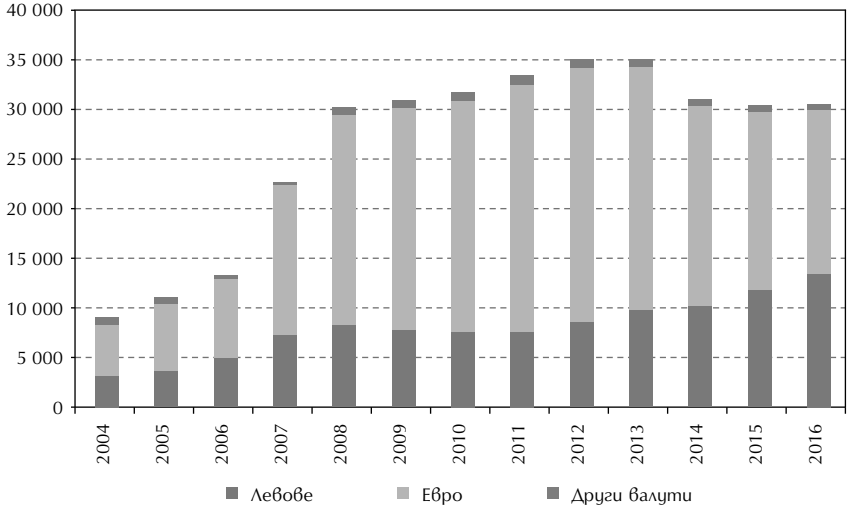


Източници: БНБ, НСИ.

Във валутната структура на кредитите за неправителствения сектор преобладават кредитите в левове и в евро, като към края на 2016 г. делът на кредитите в други валути е едва 1.5%. Тази структура на кредитния пазар е обусловена от действието на режима на паричен съвет, който премахва несигурността по отношение на валутния курс спрямо еврото, както и от еднаквото третиране на лева и резервната валута. Незначителният дял на кредитите, деноминирани в други валути, може да се разглежда и като свидетелство за нежеланието на икономическите агенти да поемат валутен риск. Основната разлика във валутната структура на кредитите за нефинансови предприятия и за домакинства се свежда до относителния дял на кредитите в евро. При нефинансовите предприятия преобладаваща част от обема на кредитите са в евро, което вероятно е свързано с активните търговски и финансови взаимоотношения на този сектор с еврозоната. Обратно, при домакинствата относителният дял на заемното финансиране в левове е по-голям.

Графика 4. **Кредити на нефинансови предприятия по валути**

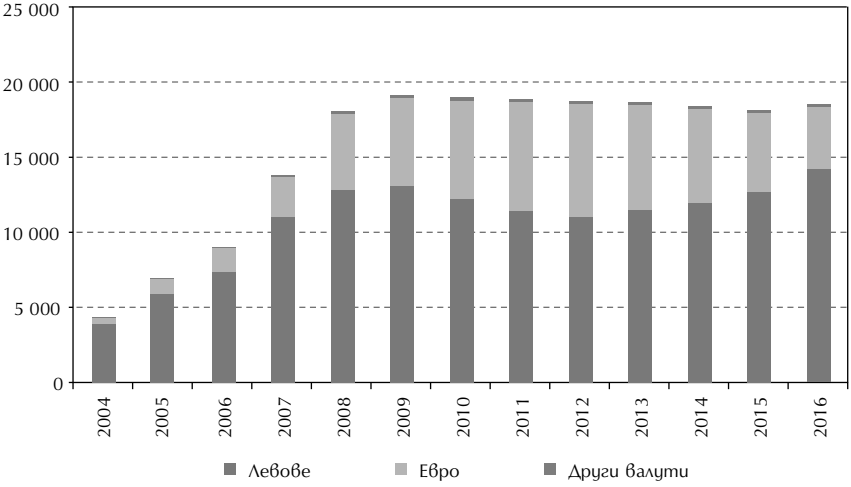
(млн. лв.)



Източник: БНБ, аналитична отчетност на други ПФИ.

Графика 5. **Кредити на домакинства и НТООД по валути**

(млн. лв.)



Източник: БНБ, аналитична отчетност на други ПФИ.

От гледна точка на анализа за цикличното развитие на кредитния пазар в България от 2000 г. досега могат да се обособят три периода на кредитна активност. Представа за това дават тенденциите в годишния растеж на новоотпуснатите кредити (графика 6), годишния растеж на салдата на кредитите за неправителствения сектор (графика 7) и годишното изменение на съотношението „кредити за неправителствения сектор/БВП“ (графика 8).

Първият период обхваща времето от началото на 2000 г. до средата на 2004 г. Той се характеризира с предпазливо поведение на икономическите агенти, намиращо изражение в ограничена задлъжнялост на неправителствения сектор, с наличие на значителни буфери в банковата система и слаба кредитна активност, произтичаща както от ограниченото предлагане на заемни ресурси, така и от потиснатото търсене на кредити от страна на предприятията и домакинствата. Въпреки че през този период се наблюдават сравнително високи темпове на кредитен растеж, те са резултат главно от ефекта на ниската база.

През периода от средата на 2004 г. до края на 2008 г. е налице рязко ускоряване на кредитирането и бързо нарастване на задлъжнялостта. Макроикономическата среда през този период се характеризира с висока икономическа активност, подкрепена от благоприятните тенденции във външната среда и силното вътрешно търсене. Допълнителна роля за това има значителният приток на капитали в българската икономика. Бързото нарастване на доходите и оптимистичните очаквания, свързани с присъединяването на България към ЕС, водят до нарастване както на търсенето, така и на предлагането на кредити, като за високата кредитна активност на банките допълнително допринасят конкуренцията сред кредитните институции, склонността към поемане на по-висок риск и значителното увеличение на депозитите на резиденти и на финансирането от банките майки. Бързият растеж на кредита в този период поражда опасения от натрупването на цикличен системен риск в банковата система, което е причина за прилагането на мерки за ограничаване на кредитната активност. Въвеждането на такива мерки, сред които и използването на допълнителни задължителни резерви, успява да допринесе за известно забавяне на кредитния растеж в рамките на 2005 г. и 2006 г., но премахването им е последвано от ново рязко активизиране на кредитирането през 2007 г., когато при новоотпуснатите кредити е отчетен 58% годишен растеж, салдата на кредитите за неправителствения сектор отбелязват годишно нарастване с 64%, а съотношението „кредити за неправителствения сектор/БВП“ се увеличава с 23 пр.п. в рамките на една година.

Третият период започва в края на 2008 г., когато глобалната икономическа криза засяга и българската икономика посредством отслабването на външното и вътрешното търсене и свиването на характерните за предходните години значителни входящи потоци по финансовата сметка на платежния баланс. Отчетеният през 2009 г. спад на БВП в реално изражение е последван от продължителен период на потисната икономическа активност. В условията на висока несигурност в икономическата среда намаляват както нагласите на икономическите агенти да поемат нов дълг, така и склонността на банките да разширяват кредитната си дейност. През 2009 г. обемът на новоотпуснатите кредити на предприятия и домакинства отбелязва спад с 39% на годишна база, а годишният растеж на кредитите за неправителствения сектор се забавя рязко до около 3% и остава под това равнище до края на разглеждания период. С изключение на 2014 г., когато заради изваждането на Корпоративна търговска банка от обхвата на паричната статистика е отчетен по-значителен спад в кредитните салда, нивото на кредитите за неправителствения сектор през третия период не претърпява съществени изменения.

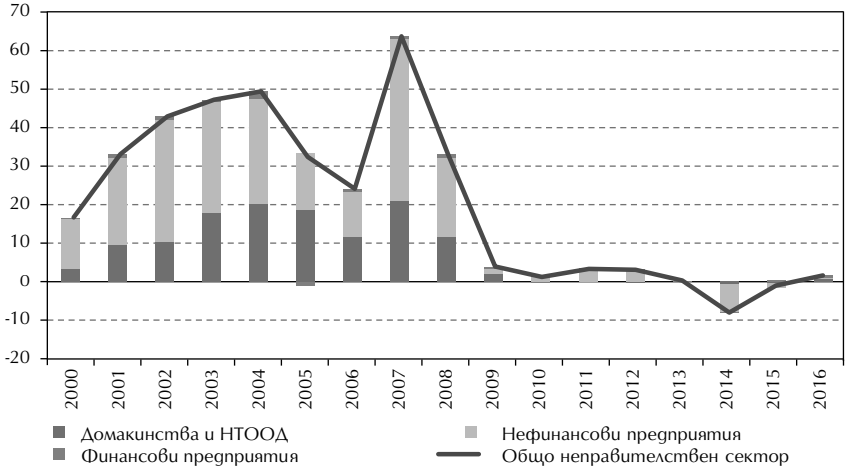
Графика 6. **Годишен растеж на новоотпуснатите кредити и принос по сектори**

(%, пр.п.)

Източник: БНБ, лихвена статистика.

Графика 7. **Годишен растеж на кредитите за неправителствения сектор и принос по сектори**

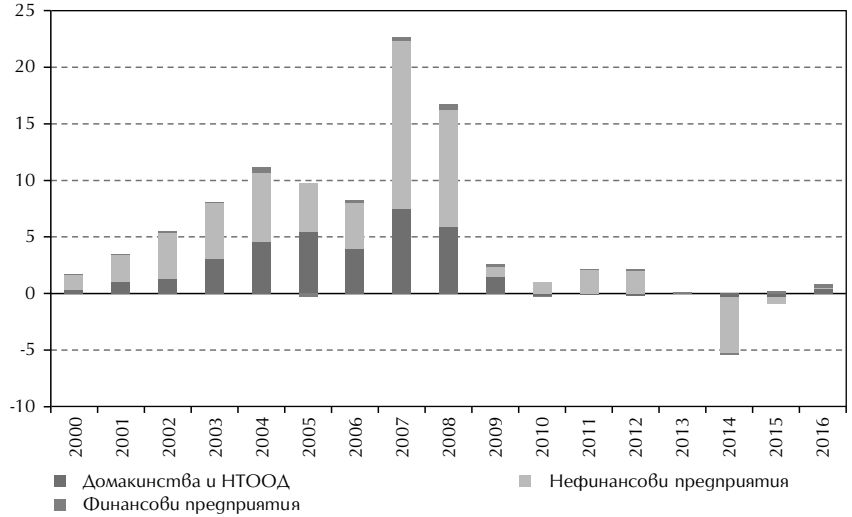
(%, пр.п.)



Източник: БНБ, аналитична отчетност на други ПФИ.

Графика 8. **Изменение на съотношението „кредити за неправителствения сектор/БВП“**

(пр.п.)



Източник: БНБ, аналитична отчетност на други ПФИ.

Представеният тук дескриптивен анализ дава само предварителна представа за тенденциите в кредитната дейност на българската банкова система през последните години. По-нататък в изследването ще бъде направена количествена оценка на влиянието, което различните фактори оказват върху търсенето и предлагането на кредити в България. За да се разберат по-ясно основните характеристики на използвания подход в същинската част от анализа обаче, първо е представен преглед на научната литература по темата за моделирането на търсенето и предлагането на заемни ресурси.

II. Обзор на изследвания с оценени функции на търсене и предлагане на кредитния пазар

Сред по-ранните изследвания с оценени в рамките на система уравнения функции на търсене и предлагане на кредити е статията на Melitz и Pardue (1973), които оценяват модела си за кредитния пазар в САЩ въз основа на динамични редове за времето от 1951 г. до 1969 г. В голям брой публикации оценяването на отделни функции на търсенето и предлагането на кредити се прави в контекста на евентуално неравновесие на пазара на заемни ресурси. Методологията на тези изследвания се базира на статиите на *Fair* и *Jaffee* (1972), *Fair* и *Kelejian* (1974) и *Maddala* и *Nelson* (1974), които представят алтернативни методи за оценяване на функциите на търсене и предлагане в условия на пазарно неравновесие, като комбинират метода на максималното правдоподобие *(maximum likelihood method)* с информацията за посоката на изменение на цената. В неравновесните модели се прави допускането, че е налице несъответствие между търсенето и предлагането на пазара. Този подход изисква отделно моделиране на двете функции и оценка на вероятността, с която наблюдаваният пазарен обем принадлежи на кривата на търсене или предлагане.

В случая на кредитния пазар изследванията се фокусират главно върху наличието на евентуално превишение на търсенето над предлагането на кредити и съответно количествено ограничаване на кредита от страна на банките, което се обяснява с негъвкавост на лихвения процент в краткосрочен план, обусловена от наличието на асиметрия в информацията между банките и потенциалните креди-

тополучатели.[1] Същевременно в литературата не се среща конкретна интерпретация на ситуациите, в които предлагането на заемни ресурси надхвърля тяхното търсене.

Първите опити за използване на неравновесните модели по отношение на кредитния пазар са от края на 70-те години. *Laffont* и *Garcia* (1977) оценяват с месечни данни такъв модел за търсенето и предлагането на корпоративни кредити в Канада. Те установяват негъвкавост на лихвения процент по кредитите в низходяща посока и стигат до заключението, че като цяло наблюдаваните обеми на кредитите зависят и се определят от нивото на кредитното търсене. *Sealey* (1979) проверява хипотезата за наличие на неравновесие на пазара на корпоративни кредити в САЩ с данни за периода от второто тримесечие на 1952 г. до третото тримесечие на 1977 г., като резултатите му показват, че в отделни периоди от извадката е налице значително несъответствие между търсенето и предлагането на кредити.

През последните двадесет години неравновесните модели са често използвано средство за анализ на кредитния пазар и на конкретните причини за потиснатата кредитна активност в периоди след финансови кризи. Например, *Pazarbaşioğlu* (1997) оценява такъв модел за пазара на корпоративни кредити във Финландия, за да провери хипотезата дали свиването на кредита след банковата криза от 1991–1992 г. се дължи на намаление в предлагането на кредити. Тя установява, че макар склонността на банките да кредитират е отслабнала в периода на влошено качество на активите, основната причина за свиването на кредита е резултат от цикличното намаление на търсенето на кредити, обусловено от високата задлъжнялост на кредитополучателите. *Ghosh* и *Ghosh* (1999) използват неравновесен модел на кредитния пазар в Индонезия, Корея и Тайланд с месечни данни за периода от началото на 1992 г. до средата на 1998 г., за да проверят дали свиването на кредита в реално изражение в тези страни през периода 1998–1999 г. се дължи на ограничено предлагане на кредити. Според получените резултати търсенето на кредит в реално изражение в разглежда-

[1] *Stiglitz* и *Weiss* (1981) предлагат възможно обяснение за негъвкавостта на лихвеното ниво по кредитите във възходяща посока. Те показват, че повишението на лихвения процент по кредитите над определено равнище би намалило очакваната възвръщаемост на банките, тъй като потенциалните кандидати за кредит, които са склонни да поемат по-високо лихвено бреме (по-голяма лихвена тежест), вероятно имат по-лоши от средното рискови характеристики, а настоящите длъжници биха избрали по-рискови проекти, когато са изправени пред по-висок лихвен процент по кредитите. За да не се отклонят от оптималното равнище, банките не биха предприели повишение на лихвения процент по кредитите дори и при нарастваща цена на привлечените ресурси, като вместо това биха предпочели да ограничат отпускането на кредити. В такова равновесие лихвеният процент по кредитите се характеризира с негъвкавост във възходяща посока.

ните страни е намаляло в по-голяма степен от реалното предлагане на кредит, поради което хипотезата за количествено ограничение на кредита от страна на предлагането не се потвърждава. *Barajas* и *Steiner* (2002) анализират значителното забавяне на растежа на кредита за неправителствения сектор в Колумбия, Мексико и Перу в годините след мексиканската криза от 1994–1995 г. Според получените резултати забавянето на кредитния растеж в трите страни се дължи на различни фактори: в Колумбия търсенето и предлагането на кредит са намалели приблизително в еднаква степен, в Мексико е налице превишение на предлагането над търсенето на кредити, а в Перу търсенето на кредити е било по-високо от предлагането.

На базата на неравновесен модел за пазара на корпоративни кредити, оценен с тримесечни данни за периода 1992–2002 г., *Nehls* и *Schmidt* (2003) проверяват хипотезата за количествено ограничение на кредита от страна на банките в Германия след рецесията през 2001–2002 г. Авторите установяват, че търсенето на кредити е превишавало предлагането през 2002 г., като причината за по-ниското предлагане на кредити е свързана със спада на цените на акциите на капиталовия пазар, отразяващ влошаването на финансовите резултати на предприятията.

Allain и *Oulidi* (2009) анализират динамиката на кредита в Мароко с тримесечни данни за периода от средата на 1997 г. до края на 2007 г. С помощта на неравновесен модел на кредитния пазар авторите стигат до извода, че причината за слабия кредитен растеж в страната в първите години след 2000 г. не е свързана с предлагането на кредити и че за ускоряването на растежа на кредита през следващите години значителен принос има бързото нарастване на цените на недвижимите имоти, което е стимулирало както търсенето, така и предлагането на кредити.

Poghosyan (2010) оценява неравновесен модел на кредитния пазар в Йордания с данни от декември 1999 г. до януари 2010 г., за да определи относителната роля на търсенето и предлагането на кредити за забавянето на кредитния растеж в страната. Според резултатите причината за слабия растеж на кредита е свързана главно с ограниченото предлагане на кредити, въпреки че търсенето на кредити също е намаляло заради отслабването на икономическата активност.

Čeh, Dumičić и *Krznar* (2011) прилагат неравновесен модел по отношение на кредитния пазар в Хърватия и обособяват три периода в рамките на използваната от тях извадка. В рамките на първия период (2000–2002 г.) те установяват равновесие между търсенето и предлагането на кредити в страната; от 2002 г. до средата на 2008 г. под

влияние на силния приток на чуждестранни капитали предлагането на кредити надвишава търсенето на заемни средства, докато по време на глобалната финансова и икономическа криза авторите отчитат недостиг в предлагането на кредити в Хърватия.

Въпреки че преобладаващата част от изследванията с неравновесен модел на кредитния пазар са базирани върху използването на агрегирани динамични редове за нивото на кредита и определящите го фактори, в някои от публикациите неравновесният модел е оценен с иконометрични техники за панелни данни, отнасящи се за извадка от кредитополучатели. Например, *Atanasova* и *Wilson* (2004) прилагат този тип анализ в контекста на изследването на кредитния канал в предавателния механизъм на паричната политика и по-конкретно на въпроса, дали предлагането на кредити за по-малките предприятия намалява при затягане на паричната политика. Авторите оценяват неравновесния модел за периода 1989–1999 г. по отношение на извадка от 639 малки и средни предприятия във Великобритания. В индивидуалните индикатори за тях са включени размерът на активите (като измерител на потенциалното обезпечение по кредита), паричните потоци и обемът на продажбите на индивидуалните предприятия. *Farinha* и *Felix* (2014) прилагат подобна аналитична рамка за извадка от индивидуални данни за малки и средни предприятия в Португалия, като установяват, че през периода 2010–2012 г. банките са ограничавали предлагането на кредит по отношение на значителна част от тях.

Независимо от разнообразието на използваните спецификации и конкретните променливи, които отразяват влиянието на разглежданите фактори, търсенето на кредити обикновено се представя като функция на променливи за икономическата активност, очакваната инфлация, лихвения процент по кредитите, възможностите за финансиране с вътрешно генерирани парични потоци от дейността и за цената на алтернативните източници за привличане на заемни ресурси.

Състоянието на икономическата активност влияе върху очакванията на икономическите агенти за размера на бъдещите им доходи и парични потоци и съответно върху склонността им да финансират потреблението и инвестициите си със заемни средства. Поради тази причина се смята, че по-високата икономическа активност обикновено води до нарастване на търсенето на кредити. Според *Melitz* и *Pardue* (1973) обаче е необходимо да се прави разграничение между дългосрочното равнище и цикличния компонент на доходите, тъй като тяхното влияние върху търсенето на кредити може да бъде разнопо-

сочно: докато нарастването на дохода в дългосрочен план стимулира търсенето на заемни ресурси, то размерът на цикличния компонент на дохода може да се намира в обратнопропорционална зависимост с кредитното търсене. Например, за да изгладят динамиката на потреблението и инвестициите, икономическите агенти ще увеличат търсенето на заемни средства при временен спад на доходите, докато в ситуация на циклично по-високи доходи ще се стремят да намалят своята задлъжнялост. В емпиричната литература състоянието на икономическата активност обикновено се представя с лаговите стойности на индекса на промишленото производство (*Laffont, Garcia* 1977; *Sealey* 1979; *Pazarbaşioğlu* 1997; *Ghosh, Ghosh* 1999) или с реалния БВП (*Nehls, Schmidt* 2003; *Čeh et al.* 2011). При оценяването на неравновесен модел за кредитния пазар в разглежданите от тях страни *Barajas* и *Steiner* (2002) използват индекса на промишленото производство в някои от спецификациите, докато в други състоянието на икономическата активност е измерено с динамиката на реалния БВП. Хипотезата на *Melitz* и *Pardue* (1973) за наличието на обратнопропорционална зависимост между цикличния компонент на дохода и търсенето на заемни средства е възприета в анализите на *Ghosh* и *Ghosh* (1999) и *Čeh, Dumičić* и *Krznar* (2011), които в уравнението за търсенето на кредити включват както променлива за текущото състояние на икономическата активност, така и отклонението от дългосрочната тенденция. *Ghosh* и *Ghosh* (1999) и *Barajas* и *Steiner* (2002) апроксимират състоянието на икономическата среда и с динамиката на индекса на цените на акциите, но отбелязват, че неговата зависимост с търсенето на кредити може да се окаже обратнопропорционална, ако този индекс отразява по-скоро алтернативните възможности за набиране на средства за финансиране на дейността на предприятията.

По отношение на влиянието на инфлацията върху търсенето на кредити обикновено се смята, че зависимостта между тях е правопропорционална. Допускането е, че по-високата текуща инфлация създава очаквания за по-бързо повишаване на ценовото равнище в бъдеще, при реализацията на които би се облекчило обслужването на номиналния размер на дълга. Съществува обаче и алтернативен възглед, според който високата инфлация създава несигурност в макроикономическата среда, поради което се отразява негативно върху търсенето на кредити (*Ghosh, Ghosh* 1999).

Цената на заемните ресурси е един основните фактори, който ограничава търсенето на кредити от страна на потенциалните кредитополучатели. Нейното влияние се отчита, като в уравнението за търсене на кредити се включи лихвеният процент по кредитите

(*Pazarbaşioğlu* 1997; *Ghosh, Ghosh* 1999) или относителната цена на банковия кредит, измерена със спреда спрямо цената на алтернативните източници на финансиране, най-често лихвеното ниво по корпоративните дългови книжа (*Laffont, Garcia* 1977; *Sealey* 1979). По отношение на цената на алтернативните източници за набиране на заемни ресурси се смята, че нейното повишаване води до нарастване на търсенето на банкови кредити. *Čeh, Dumičić* и *Krznar* (2011) използват ЮРИБОР и *JP Morgan EMBI* спреда като измерители за цената на алтернативните източници за финансиране на предприятията, но посочват, че зависимостта на този спред с търсенето на кредити може да се окаже и обратнопропорционална, ако той отразява по-скоро оценката на чуждестранните инвеститори за риска в страната.

Възможностите за финансиране със собствени средства чрез използване на неразпределените печалби на предприятията обикновено се включват в уравнението за търсене на кредит, без да се формулират категорични твърдения относно посоката им на влияние. В краткосрочен план неразпределените печалби са алтернатива на банковия кредит като средство за финансиране на дейността, поради което тяхното нарастване би намалило потребностите от набиране на ресурси от външни източници. В дългосрочен план обаче е нормално зависимостта между размера на вътрешно генерираните средства и търсенето на банкови кредити да е правопропорционална, тъй като по-високите печалби дават възможност на предприятията за повече инвестиционни проекти, при финансирането на които банките изискват от потенциалните кредитополучатели да участват и със собствени средства (*Laffont, Garcia* 1977; *Sealey* 1979).

Ролята на цените на недвижимите имоти като фактор за търсенето на кредит е предмет на редица изследвания (вж. например *Hofmann* 2001). Нарастването на цените на недвижимите имоти създава ефект на богатството и води до увеличение на стойността на потенциалното обезпечение по ипотечните кредити, което от своя страна засилва склонността на кредитополучателите да поемат нови задължения. Наред с това правопропорционалната зависимост произтича и от необходимостта от допълнителни средства за покупките на недвижими имоти в условията на повишение на техните цени.

При моделирането на функцията на предлагане на банкови кредити се изхожда от разбирането, че размерът на предлаганите от банковата система заемни ресурси зависи от: 1) обема на ликвидните средства, с които банките разполагат; и 2) склонността на банките да отпускат кредити, разглеждана в контекста на портфейлната теория за инвес-

тициите, при която се търси оптимално съотношение между риск и доходност.

Включването на променливата за разполагаемите ресурси в банковата система служи за идентификацията на функцията на предлагането на кредити. Тъй като тази променлива присъства във функцията на предлагане, но не и в уравнението за търсене на кредити, с нейна помощ е възможно разграничаването на двете функции. В емпиричната литература се използват различни измерители за наличните ликвидни ресурси в банковата система, например депозитната маса (*Sealey* 1979), сумата на депозитите и капитала на банковия сектор (*Pazarbaşioğlu* 1997; *Nehls, Schmidt* 2003), обемът на депозитите, намален със задължителните минимални резерви (ЗМР), (*Laffont, Garcia* 1977; *Allain, Oulidi* 2009; *Poghosyan* 2010). *Ghosh* и *Ghosh* (1999) определят като показател за капацитета на банковата система да кредитира по-малката от две величини: максималната стойност на активите, която може да се поддържа при текущите изисквания за капиталова адекватност; и сумата на пасивите на банковата система, намалени със ЗМР, ликвидните изисквания и касовите наличности. Като допълнителна променлива за способността на банките да наберат допълнителен капитал, с който да финансират кредитирането, *Pazarbaşioğlu* (1997) включва и пазарната оценка за относителната доходност на банковата дейност, измерена със съотношението между цената на банковите акции и средната цена на акциите на капиталовия пазар.

Сред факторите, влияещи върху склонността на банките да разширяват кредитната си дейност, е нивото на лихвения процент по кредитите. Влиянието на този фактор върху предлагането на кредити се взема предвид посредством включването на директен измерител за нивото на лихвения процент по кредитите (*Laffont, Garcia* 1977; *Pazarbaşioğlu* 1997; *Nehls, Schmidt* 2003) или на неговия спред спрямо възвръщаемостта на алтернативните инвестиции, най-често доходността на държавните ценни книжа (*Laffont, Garcia* 1977; *Sealey* 1979). *Ghosh* и *Ghosh* (1999) използват като алтернативен показател за доходността от кредитната дейност спреда между лихвения процент по кредитите и лихвения процент по депозитите.

В емпиричната литература влиянието на свързания с кредитната дейност риск върху склонността на банките да кредитират се отчита най-често с включването на няколко показателя. Нивото на икономическата активност е правопропорционално свързано с предлагането на кредит, тъй като подобряването на нейното състояние влияе благоприятно върху оценката на банките за способността на кредитополучателите да обслужват своите задължения. Най-често

използваните показатели за икономическата активност в литературата за моделирането на търсенето и предлагането на кредити са индексът на промишленото производство (*Laffont, Garcia* 1977; *Sealey* 1979; *Pazarbaşioğlu* 1997; *Ghosh, Ghosh* 1999; *Poghosyan* 2010) и реалният БВП (*Nehls, Schmidt* 2003; *Čeh et al.* 2011). Относно влиянието на инфлацията върху предлагането на кредити се смята, че по-бързите темпове на изменение на ценовото равнище създават по-голяма несигурност, което прави банките по-внимателни и проявяващи по-малка склонност да разширяват кредитната си дейност (*Pazarbaşioğlu* 1997).

Кредитният риск ограничава предлагането на кредити, като в литературата се срещат различни измерители за нивото на този риск. Като такъв показател се използва цикличната рискова премия, измерена със спреда между лихвения процент по кредитите и лихвеното ниво на междубанковия паричен пазар (*Pazarbaşioğlu* 1997), или със спреда между лихвения процент по кредитите и лихвения процент по депозитите (*Laffont, Garcia* 1977; *Nehls, Schmidt* 2003; *Čeh et al.* 2011). *Barajas и Steiner* (2002) и *Čeh, Dumičić и Krznar* (2011) измерват кредитния риск с дела на необслужваните кредити и съотношението между обезценките за кредитен риск и необслужваните кредити, въпреки че тези показатели отразяват вече материализиран, а не очакван кредитен риск. Потенциална защита на банките срещу загуби от кредитен риск са обезпеченията, предоставени от кредитополучателите. Нарастването на стойността на тези обезпечения увеличават склонността на банките да разширяват предлагането на кредити, като в някои публикации влиянието на този фактор е отчетено с включването на променлива за пазарната капитализация на предприятията (*Pazarbaşioğlu* 1997; *Nehls, Schmidt* 2003) или за цените на недвижимите имоти (*Allain, Oulidi* 2009).

Недостатък на по-ранните изследвания с неравновесни модели е, че не обръщат внимание на това, дали разглежданите променливи са нестационарни, и оценяват моделите в нива. Въпреки че при използването на разлики се губи твърде много информация (например, ако темпът на растеж на търсенето е по-висок от растежа на предлагането, това не означава непременно, че нивото на търсене е по-високо от това на предлагането), работата с нестационарни данни изисква внимание. *Ghosh* и *Ghosh* (1999) посочват, че когато променливите имат единичен корен, оценката в нива е оправдана, само ако те формират коинтеграционен вектор. За целта тези автори предлагат да се провери дали оценките за търсенето и предлагането са коинтегрирани със серията за наблюдавания пазарен обем. В този случай уравненията

за търсенето и предлагането могат да се интерпретират като коинтеграционни зависимости.

Сред изследванията за България, в които е направена оценка на отделни функции на търсене и предлагане на кредити, са разработките на Христов и Михайлов (2002) и Пешев (2015). Христов и Михайлов (2002) оценяват функции на търсенето и предлагането на кредити за предприятия, за да изследват причините за слабата кредитна активност в периода след кризата от 1996–1997 г. В оцененото уравнение за предлагането на кредит кредитният капацитет на банковата система и лихвеният процент по кредитите са статистически значими и имат очаквания положителен знак, а рисковата премия, измерена с разликата между цената на кредита и лихвеното ниво на междубанковия паричен пазар, влиза в уравнението с очаквания отрицателен знак. Търсенето на кредит се влияе правопропорционално от обема на инвестициите и очакваната инфлация, а полученият положителен знак на коефициента пред лихвения процент в уравнението за търсене на кредити е интерпретиран като свидетелство за това, че по-високите лихвени проценти са привличали кредитоискатели, готови да поемат по-голям риск. Въз основа на резултатите от оценените уравнения са генерирани серии за търсенето и предлагането на кредити от предприятия, съпоставката на които показва, че от 1997 г. до началото на 2000 г. има по-голямо предлагане на кредит в икономиката, отколкото търсене от страна на предприятията. През периода 2000-2001 г. се наблюдава промяна в кредитната среда, като от ситуация на недостатъчно търсене на кредит се преминава във фаза на недостатъчно предлагане.

Пешев (2015) използва няколко алтернативни спецификации на модели с корекция на грешката при оценяването на функциите на търсенето и предлагането на банкови кредити в България. Основните резултати от неговия анализ показват, че търсенето на банкови кредити в дългосрочен период зависи главно от икономическата активност, входящите капитали, банковите депозити и съотношението между депозитите и кредитите на нефинансовия частен сектор. В краткосрочен период тези детерминанти се допълват и от вноса на стоки и услуги. Предлагането на кредити зависи в дългосрочен и краткосрочен период от промяната в банковата ликвидност, нетния лихвен доход, капиталовата адекватност и динамиката в производствените цени, но само в дългосрочен период потребителските цени, пазарната концентрация и чуждестранната собственост допълват факторите на предлагане на кредити. Въпреки че оценява функциите на търсене и предлагане на заемни ресурси в България, Пешев (2015) не

анализира относителната им роля при формирането на реално наблюдаваните обеми на кредита.

III. Използвани данни

Анализът е базиран на динамични редове с тримесечна честота за периода от началото на 2000 г. до края на 2016 г.

По отношение на данните за обемите на кредитния пазар е необходимо да се направи уточнението, че те се отнасят за потоците от новоотпуснати кредити. Поначало изменението на кредитните салда включва четири компонента: 1) потока на новоотпуснатите през съответния период кредити; 2) потока на погашенията по съществуващите кредити; 3) размера на отписаните кредити; и 4) обема на нетните продажби на кредити през съответния период. Фокусирането върху общото изменение на кредитните салда изисква наличието на достатъчно подробна информация както за тези компоненти, така и за факторите, влияещи върху обемите на всеки един от тях. Тъй като от изброените четири елемента на изменението на кредитните салда детайлна публична информация е налична единствено за потоците новоотпуснати кредити, техният размер е използван като индикатор за обемите на кредитния пазар.

Данните за обемите и лихвените проценти по новоотпуснатите кредити са с източник лихвената статистика на БНБ. Тази статистика дава възможност да бъдат използвани няколко критерия за разграничение на кредитите, които са обект на анализ. Първо, от гледна точка на секторната принадлежност на длъжниците кредитите могат да бъдат обособени на такива, отпуснати на нефинансови предприятия или на домакинства. Спрямо кредитите за нефинансови предприятия допълнително е приложено разграничение в зависимост от оригиналния матуритет на кредита. За тази цел е използвана разбивката, осигурявана от лихвената статистика на БНБ до края на 2006 г., т.е. разделение на кредитите на краткосрочни (матуритет до една година) и дългосрочни (над една година). От началото на 2007 г. е налице и по-подробна матуритетна разбивка, но необходимостта от максимално дълги динамични редове и съпоставимост наложи данните за лихвените проценти по кредитите след 2007 г. да бъдат групирани според съществуващото дотогава разделение. По отношение на домакинствата са използвани данни както за общия обем на отпуснатите им кредити, така и отделни редове за потребителските и жилищните кредити (поради различния обхват на лихвената статистика преди 2007 г. информацията за потребителските и жилищните кредити

до края на 2006 г. обхваща само левовия им компонент). За целите на изследването месечните данни от лихвената статистика на БНБ са трансформирани в тримесечни: данните за обемите са получени като сума от новоотпуснатите кредити през отделните месеци от тримесечието, а тримесечните стойности на лихвените проценти по кредитите са изчислени посредством претегляне на месечните данни за лихвените проценти с обемите на новоотпуснатите през съответните месеци кредити. Въпреки че може да е свързано с известна загуба на информация, преминаването към по-ниска честота на данните дава възможност да се неутрализира влиянието на екстремните стойности в лихвените условия и при обемите по отделни кредитни споразумения, което ясно се проявява при сравнително малък обем на пазара като този в България (значителната вариация в месечните данни за новоотпуснатите кредити от лихвената статистика се дължи именно на това, което потвърждава необходимостта от преминаване към по-ниска честота). Другата причина за трансформирането на месечните данни за обемите и лихвените проценти по новоотпуснатите кредити в тримесечни серии е необходимостта от уеднаквяване на периодичността им с тази на информацията от системата на националните сметки. Обемите на новоотпуснатите кредити са изразени в постоянни цени от 2010 г., като са дефлирани с хармонизирания индекс на потребителските цени (ХИПЦ). Причините този индекс да бъде избран за дефлатор са широкото му използване като показател за ценовата динамика, малкият лаг при неговото публикуване и релевантността му по отношение както на домакинствата, така и на предприятията (във връзка с влиянието на ценовото равнище върху номиналния обем на генерираните парични потоци от продажби).

Наред с лихвените проценти по новоотпуснатите кредити в групата на показателите за лихвените нива са включени тримесечният ЮРИБОР (като приблизителен измерител на доходността за банките на алтернативни на кредитната дейност финансови инструменти) и имплицитният лихвен процент по брутния външен дълг на частния небанков сектор (като фактор, измерващ цената на алтернативните източници на финансиране на икономическите агенти в страната). Динамичният ред за ЮРИБОР е с източник базата данни на ЕЦБ, като той отразява официалната информация за лихвеното ниво по междубанковите депозити на паричния пазар в еврозоната. Поради широкото използване на дългово финансиране от чужбина от страна на частния небанков сектор е необходимо при анализа на кредитния пазар в страната да бъде взета предвид и цената, по която се привличат

ресурсите от нерезиденти. Тъй като динамични редове за този индикатор не се публикуват, за целите на изследването цената на алтернативните източници на финансиране е измерена косвено въз основа на публикуваната от БНБ информация за размера на брутния външен дълг и неговото обслужване, като номиналните обеми на лихвените плащания са отнесени към средния размер на брутния външен дълг на частния небанков сектор за съответното тримесечие.

За целите на моделирането на кредитното предлагане е конструиран индикатор за обема на свободните ресурси, които банковата система може да насочи за разширяване на кредитната дейност (т.нар. кредитен капацитет на банковата система). Динамичният ред за него е получен въз основа на аналитичната отчетност на сектор „други парично-финансови институции" от паричната статистика на БНБ, като към депозитите на неправителствения сектор са прибавени чуждестранните пасиви на банковата система и от получената сума е изваден размерът на средствата, които банките са длъжни да поддържат като резервни активи във връзка с регулацията на ЗМР. С включването на чуждестранните пасиви на банките в рамките на кредитния капацитет на системата се отчита ролята на привлеченото от банковия сектор дългово финансиране от чужбина, използването на което бе обичайна практика сред кредитните институции по време на периода с бърз растеж на кредита. Показателят за кредитния капацитет на банковата система е изразен в постоянни цени, като е дефлиран с ХИПЦ.

Състоянието на икономическата активност е апроксимирано с динамиката на брутния вътрешен продукт, изразен по постоянни цени от 2010 г. Източник на тези данни е системата на националните сметки, публикувани от Националния статистически институт (НСИ).

Сред индикаторите, които ще бъдат използвани в количествения анализ на факторите, влияещи върху обемите на кредитния пазар, са годишният темп на инфлация и годишният темп на растеж на цените на жилищата. Динамиката на инфлацията е измерена с годишния растеж на ХИПЦ, публикуван от НСИ. Необходимостта от динамичен ред за динамиката на цените на жилищата, обхващащ целия период на извадката, е причината такъв ред да се конструира чрез комбинирането на информация от настоящата и вече преустановената статистика на НСИ в тази област. Тъй като данни за общия индекс на цените на жилищата (обхващащ съществуващите и новите жилища) са налични само за времето след първото тримесечие на 2005 г., годишните темпове на изменение на цените на жилищата

за предходните периоди са изчислени въз основа на статистиката за пазарните цени на съществуващите жилища, публикуването на която НСИ преустанови от началото на 2015 г. Конструиран по този начин, показателят за изменението на цените на жилищата в периода до края на 2004 г. отразява само ценовата динамика на съществуващите жилища, а от началото на 2005 г. в обхвата му наред с тях присъстват и новите жилища.

С изключение на променливите за лихвените нива и годишните темпове на изменение на потребителските цени и на цените на жилищата, които са изразени в проценти, всички останали индикатори са логаритмувани, за да може оценените за тях коефициенти в иконометричния анализ да бъдат интерпретирани като еластичности. Поради наличието на силно изразена сезонност в обемите на новоотпуснатите кредити, кредитния капацитет на банковата система и показателите за икономическата активност, тези променливи са сезонно изгладени по метода *TRAMO/SEATS*. При показателите за лихвените нива не се наблюдава сезонност и за тях са използвани данните от изходните серии.

Преди да се пристъпи към оценяването на модела, най-напред е необходимо да се определи степента на интегрираност на разглежданите променливи. В таблица 3 в приложението са представени резултатите от *Augmented Dickey-Fuller (ADF)* теста за единичен корен, откъдето се вижда, че като цяло нулевата хипотеза за наличие на единичен корен не може да бъде отхвърлена за нивата на променливите, но се отхвърля при прилагането на *ADF* теста към първите им разлики. За да се провери устойчивостта на тези изводи беше приложен и тест за единичен корен, в който нулевата хипотеза предполага стационарност на изследваната променлива. В таблица 4 са представени резултатите от *Kwiatkowski–Phillips–Schmidt–Shin (KPSS)* тест за стационарност. Като цяло *KPSS* тестът потвърждава заключенията относно степента на интегрираност на променливите. При такива статистически характеристики на динамичните редове оценяването на моделите в нива би имало смисъл, само ако разглежданите променливи формират коинтеграционен вектор. В съответствие с по-новата литература за неравновесните модели това условие е проверено по-нататък в анализа.

Таблица 1. **Означения на използваните променливи**

Означение	Дефиниция
leh	Логаритмувана стойност на кредитите за нефинансови предприятия и домакинства в реално изражение
leh	Логаритмувана стойност на кредитите за нефинансови предприятия в реално изражение
lest	Логаритмувана стойност на краткосрочните кредити за предприятия в реално изражение
lelt	Логаритмувана стойност на дългосрочните кредити за предприятия в реално изражение
lh	Логаритмувана стойност на кредитите за домакинства в реално изражение
lcons	Логаритмувана стойност на потребителските кредити в реално изражение
lhousing	Логатимувана стойност на жилищните кредити в реално изражение
il eh	Лихвен процент по кредитите за нефинансови предприятия и домакинства
il e	Лихвен процент по кредитите за нефинансови предприятия
i stle	Лихвен процент по краткосрочните кредити за предприятия
i ltle	Лихвен процент по дългосрочните кредити за предприятия
i h	Лихвен процент по кредитите за домакинства
il consl	Лихвен процент по потребителските кредити
il hl	Лихвен процент по жилищните кредити
blc	Логаритмувана стойност на кредитния капацитет на банковата система в реално изражение
euribor 3m	Тримесечен Юрибор
i ed	Имплицитен лихвен процент по външния дълг на частния небанков сектор
y	Логаритмувана стойност на БВП по цени от 2010 г.
infl	Годишен темп на изменение на хармонизирания индекс на потребителските цени
hpag	Годишен темп на изменение на цените на жилищата

IV. Неравновесни модели за кредитния пазар в България

В тази част от изследването са оценени функции на търсенето и предлагането на кредити при допускане, че стойността на пазарните обеми се определя от по-ниското от нивата на кредитно търсене и предлагане. Тази аналитична рамка е приложена спрямо всеки един от сегментите на кредитния пазар в страната, като са идентифицирани и периодите, в които е налице по-значително несъответствие между търсенето и предлагането на кредити.

Търсенето на кредити е представено като функция на променливата за реалния БВП, лихвения процент по съответния вид кредити и годишния темп на изменение на ХИПЦ, като наред с тези обяснителни променливи в моделите за търсенето на корпоративни кредити е добавен имплицитният лихвен процент по външния дълг на частния небанков сектор, а в моделите за търсенето на кредити от домакинства – годишният темп на растеж на цените на жилищата.

Променливата за реалния БВП е използвана като измерител на икономическата активност, като в общия случай се очаква нейната връзка с търсенето на кредити да е правопропорционална, тъй като по-високите текущи нива на икономическата активност се пренасят в очаквания за увеличение на доходите и паричните потоци в бъдеще, което води до засилване на склонността на икономическите агенти да използват заемни ресурси за финансиране на потреблението и инвестициите. Въпреки това обратнопропорционална зависимост също не може да бъде изключена напълно, като тя би била установена, в случай че икономическите агенти използват заемните ресурси като средство за покриване на временен недостиг на доходите и паричните потоци, с което да си осигурят по-равномерна динамика на потреблението и инвестициите в различните фази на икономическия цикъл (вж. *Melitz, Pardue* 1973).

Разгледани заедно, номиналният лихвен процент по съответния вид кредити и темпът на инфлацията отразяват реалната цена на обслужването на кредитите. Увеличението на лихвения процент по кредитите се очаква да ограничава търсенето на заемни ресурси, докато хипотезата по отношение на влиянието на отчетната инфлация е, че тя влияе стимулиращо върху търсенето на кредити, тъй като нейното увеличение е съпътствано със засилване на очакванията за повишаване на ценовото равнище в следващите периоди и съответно за облекчаване на номиналната тежест по обслужването на задълженията. Във функциите на общото търсене на кредити и търсенето на различните видове кредити за предприятия като измерител на цената на алтерна-

тивните източници на финансиране присъства имплицитното лихвено ниво по външния дълг на частния небанков сектор. С включването на тази променлива се отчита ролята на използваното от корпоративния сектор дългово финансиране от чужбина. Повишението на лихвеното равнище по този вид задължения се очаква да влияе за засилване на търсенето на кредити от местната банкова система, тъй като относителната цена на банковите кредити спрямо тази на финансирането от чужбина се понижава.

Включването на годишния темп на изменение на цените на жилищата във функциите на търсене на кредитите от домакинства е направено с хипотезата, че нарастването на цените на жилищата води до повишаване на търсенето на кредити, тъй като е свързано както с потребност от допълнителни ресурси с трансакционна цел, така и с увеличение на стойността на недвижимите имоти, които могат да бъдат предоставени като обезпечение по ипотечните кредити. В процеса на оценяване на модела като измерител на това влияние бяха използвани както нивото на индекса на цените на жилищата, така и годишният му темп на растеж. Въпреки че нивото на индекса на цените на жилищата също се оказа статистически значимо и с очаквана посока на влияние, в окончателния си вид моделът съдържа годишния му растеж, тъй като при използване на тази трансформация се постига по-голяма обяснителна сила.

Предлагането на кредити може да се разглежда като функция на кредитния капацитет на банковата система (т.е. сумата на депозитите на неправителствения сектор и чуждестранните пасиви, намалена с размера на ЗМР), степента на риск, свързана с кредитирането (измерена със спреда между лихвения процент по съответния вид кредити и ЮРИБОР) и доходността по алтернативните инвестиции на банковата система (представена от нивото на ЮРИБОР), като наред с тези променливи във функцията за общото предлагане на кредити и предлагането на кредити за домакинства е добавен и годишният темп на изменение на цените на жилищата.

Увеличаването на обема на разполагаемите ресурси на банковата система се очаква да допринася за нарастващо предлагане на кредити. Както бе посочено при описанието на източниците на данни, в обхвата му попадат както привлечените средства от резиденти, така и чуждестранните пасиви на банките. По този начин се отчита ролята на привлеченото от банковия сектор дългово финансиране от чужбина, което бе широко използвано от кредитните институции по време на периода с бърз растеж на кредита.

Един от основните фактори, които се очаква да оказват влияние върху предлагането на кредити, е оценката на банките за кредитния риск. Тази оценка може да бъде измерена най-добре със спреда между лихвения процент по съответния вид кредити и ЮРИБОР, като този показател отразява размера на рисковата премия, която банките изискват при предоставянето на заемни средства. Рисковата премия варира в рамките на икономическия цикъл – тя е най-ниска в условията на силна икономическа активност (когато очакванията са оптимистични) и най-висока в низходящата фаза на цикъла. Включването на спреда между лихвения процент по съответния вид кредити и ЮРИБОР е направено с цел да се провери хипотезата, че по-високият риск ограничава склонността на банковата система да разширява кредитната си дейност. Съществено предимство на този показател в сравнение с променливите за качеството на кредитния портфейл е, че той отразява очакваното ниво на кредитния риск, докато делът на необслужваните кредити е измерител на вече материализиран кредитен риск. Друга причина в модела да не бъдат включени показатели за качеството на активите е, че необслужваните кредити в българската банкова система се отличават с високо покритие с обезценки и обезпечения, което наред с наличието на капиталов излишък над минималните капиталови изисквания показва, че нивото на необслужваните кредити не е ограничителен фактор по отношение на предлагането на кредити.

Освен от нивото на кредитния риск склонността на банките да разширяват кредитната си дейност зависи и от динамиката на доходността, която кредитните институции могат да получат по финансови инструменти, различни от предоставянето на кредити на предприятия и домакинства. Хипотезата е, че повишаването на възвръщаемостта на алтернативните инвестиции при равни други условия ограничава предлагането на кредити. За целите на изследването влиянието на този фактор е отчетено с включването на нивото на лихвения процент по междубанковите депозити в еврозоната. Въпреки че поначало доходността на българските ДЦК също може да се използва като подходящ приблизителен измерител на възвръщаемостта от алтернативните на кредитната дейност инвестиции, отсъствието на достатъчно дълги динамични редове за нея бе причината тя да не бъде включена в модела.

Подобно на функцията на търсене на кредити, в уравнението за предлагането също присъства показател за динамиката на цените на жилищните недвижими имоти. Хипотезата по отношение на влиянието на променливата за цените на жилищата е, че тяхното нарастване засилва склонността на банките да разширяват кредит-

ната си дейност и води до повишаване на предлагането на кредити, тъй като е съпътствано от формиране на оптимистични нагласи за увеличаване на стойността на предоставените обезпечения по ипотечните кредити.

За всеки един сегмент от кредитния пазар е построен модел в следния вид:

$$L_t^D = X_{1t}'\alpha + \varepsilon_{1t}$$

$$L_t^S = X_{1t}'\beta + \varepsilon_{2t}$$

$$L_t = \min(L_t^D, L_t^S)\ ,$$

където L_t^D е търсенето на кредити, L_t^S е предлагането на кредити, X_{1t} и X_{2t} са обяснителните променливи съответно в уравненията за търсене и предлагане на кредити, α и β са оценяваните коефициенти на обяснителните променливи, а ε_{1t} и ε_{2t} са остатъците в тези уравнения. L_t е наблюдаваният пазарен обем.

При допускане, че остатъците са независими и имат нормално разпределение с вариация съответно σ_1^2 и σ_2^2, анализът може се осъществи на базата на използването на следните функции:

$$f_1(L_t) = \frac{1}{\sqrt{2\pi}\sigma_1} exp\left[-\frac{1}{2\sigma_1^2}(L_t - \alpha X_{1t})^2\right]$$

$$f_2(L_t) = \frac{1}{\sqrt{2\pi}\sigma_2} exp\left[-\frac{1}{2\sigma_2^2}(L_t - \beta X_{2t})^2\right]$$

$$F_1(L_t) = \frac{1}{\sqrt{2\pi}\sigma_1}\int_{L_t}^{\infty} exp\left[-\frac{1}{2\sigma_1^2}(L_t^D - \alpha X_{1t})^2\right] dL_t^D$$

$$F_2(L_t) = \frac{1}{\sqrt{2\pi}\sigma_2}\int_{L_t}^{\infty} exp\left[-\frac{1}{2\sigma_2^2}(L_t^S - \beta X_{2t})^2\right] dL_t^S\ ,$$

където $f_1(L_t)$ описва маргиналното разпределение на търсенето на кредити, $f_2(L_t)$ е функцията на маргиналното разпределение на предлагането на кредити, $F_1(L_t)$ е кумулативното разпределение на търсенето, а $F_2(L_t)$ – кумулативното разпределение на предлагането на кредити.

Вероятността дадено наблюдение за пазарния обем да се определя от равнището на търсенето на кредити(L_t=L_t^D<L_t^S) може да се опише с условната плътност

$$f(L_t|L_t = L_t^D < L_t^S) = \frac{\int_{L_t}^{\infty} f(L_t, L_t^S)dL_t^S}{Prob(L_t^D < L_t^S)} = \frac{f_1(L_t)\int_{L_t}^{\infty} f_2(L_t^S)dL_t^S}{Prob(L_t^D < L_t^S)} = \frac{f_1(L_t) * F_2(L_t)}{Prob(L_t^D < L_t^S)},$$

където $f(L_t, L_t^S)$ е съвместната плътност на L_t и L_t^S.

Обратно, ако в даден период търсенето превишава предлагането на кредити (L_t=L_t^S<L_t^D), условната плътност L_t е

$$f(L_t|L_t = L_t^S < L_t^D) = \frac{\int_{L_t}^{\infty} f(L_t, L_t^D)dL_t^D}{1 - Prob(L_t^D < L_t^S)} = \frac{f_2(L_t)\int_{L_t}^{\infty} f_1(L_t^D)dL_t^D}{1 - Prob(L_t^D < L_t^S)} = \frac{f_2(L_t) * F_1(L_t)}{1 - Prob(L_t^D < L_t^S)}.$$

Тъй като в даден период t пазарният обем се определя или от равнището на търсенето, или от това на предлагането,

$$f(L_t|X_{1t}, X_{2t}) = Prob(L_t^D < L_t^S)f(L_t|L_t = L_t^D) + \left(1 - Prob(L_t^D < L_t^S)\right)f(L_t|L_t = L_t^S)$$
$$= f_1(L_t).F_2(L_t) + f_2(L_t).F_1(L_t).$$

Въз основа на това целта е стойностите на параметрите α и β да се определят така, че да се максимизира функцията

$$L(\alpha, \beta|X_{1t}, X_{2t}) = \sum_{t=1}^{n} \log[f_1(L_t) * F_2(L_t) + f_2(L_t) * F_1(L_t)].$$

Тъй като максимизирането на функцията на вероятността посредством метода на максималното правдоподобие изисква задаването на начални нива на параметрите α и β, като такива са използвани стойностите на коефициентите, получени при оценяване на функциите на търсене и предлагане на кредити с двустъпковия метод на най-малките квадрати *(TSLS)*. Причината при определянето на началните стойности на коефициентите да бъде избран този метод за оценка на уравненията е да се избегне евентуално изкривяване на резултатите, произтичащо от потенциална ендогенност на обяснителните променливи. Самата итеративна процедура е проведена по алгоритъма на *Broyden–Fletcher–Goldfarb–Shanno (BFGS)*.

В таблица 2 са обобщени резултатите за оценените с метода на максималното правдоподобие функции на търсене и предлагане на различните видове кредити.

Таблица 2. **Оценки на функциите на търсене и предлагане на кредити по метода на максималното правдоподобие**

		LEH	LE	LEST	LELT	LH	LCONS	LHOUSING
Функция на търсене								
C	coef.	-12.5477***	-10.8852 **	19.5615 **	-20.8479 **	-7.6306	-2.8558	-53.2053 **
	z-stat	-2.75	-2.04	2.40	-2.22	-0.36	-0.08	-2.07
y	coef.	2.1548 ***	1.9651 ***	-1.2984 *	2.9913 ***	1.6212	1.1008	6.0412 **
	z-stat	5.02	3.86	-1.65	3.37	0.79	0.34	2.41
Interest rate	coef.	-0.0929 **	-0.1078 *	-0.1153	-0.1689 *	-0.1401	-0.1369	-0.0726
	z-stat	-2.04	-1.88	-1.32	-1.91	-1.16	-0.62	-0.57
infl	coef.	0.0054	0.0122	0.0385	0.0045	0.0221	0.0009	0.0559 ***
	z-stat	0.70	0.92	1.47	0.30	1.04	0.07	2.87
i ed	coef.	0.2823 ***	0.2553	0.1744	0.3710	-	-	-
	z-stat	2.60	1.56	1.10	1.58			
hpag	coef.	0.0092 ***	-	-	-	0.0169 ***	0.0179	0.0165 ***
	z-stat	4.33				2.83	1.27	3.25
SIGMA (D)	coef.	0.1155 ***	0.1600 ***	0.2407 ***	0.2207 ***	0.2249 *	0.1959	0.2923 ***
	z-stat	5.09	2.89	3.00	3.01	1.71	0.53	5.83
Функция на предлагане								
C	coef.	2.5085	2.7063	10.2941 ***	-0.5260	3.7914	4.5197	-10.6336 *
	z-stat	1.29	1.39	4.10	-0.14	0.56	0.31	-1.94
blc	coef.	0.6034 ***	0.5593 ***	-0.3145	0.8737 ***	0.4357	0.3240	1.5905 ***
	z-stat	4.29	3.79	-1.57	3.08	0.83	0.29	3.57
Interest rate - EURIBOR	coef.	-0.0962 *	-0.1058 *	-0.1015	-0.1506	-0.1658	-0.1444	-0.1349
	z-stat	-1.77	-1.78	-1.31	-1.54	-1.58	-0.63	-1.55
EURIBOR	coef.	-0.0163	-0.0216	0.0284	-0.0685	-0.0641	-0.1146	0.0936
	z-stat	-0.28	-0.30	0.37	-0.60	-0.40	-0.34	1.07
hpag	coef.	0.0088 ***	-	-	-	0.0160 ***	0.0174	0.0164 ***
	z-stat	3.97				2.75	1.26	5.36
SIGMA (S)	coef.	0.1352 ***	0.1742 ***	0.2394 ***	0.2411 **	0.2078 **	0.1967	0.2483 ***
	z-stat	4.18	2.62	2.90	2.53	2.01	0.53	5.27
Log likelihood		89.8813	50.4847	1.1072	6.5140	15.3310	28.4209	-14.6013

Бележка: Със знак ***/**/* е означена статистическа значимост при 1%, 5% и 10%.

Според получените резултати засилването на икономическата активност води до увеличение на общото търсене на кредити, като ефектът е положителен и статистически значим при общия обем на корпоративните кредити, дългосрочните кредити за предприятия и жилищните кредити. Това е в съответствие с предварително направената хипотеза, че по-високото равнище на текущата икономическа активност създава оптимистични очаквания за капацитета на икономическите агенти да поемат нови дългове и води до повишаване на търсенето на кредити. Единствено в случая на кредитите за предприятия със срок до една година връзката е обратнопропорционална, което означава, че предприятията разглеждат този вид финансиране главно като средство за покриване на временни нужди, предизвикани от отслабване на бизнес активността.

Коефициентът пред лихвения процент по кредитите има очаквания отрицателен знак, отразявайки ограничителното влияние на цената на кредитите върху тяхното търсене, като този ефект е статистически значим в уравнението за общото търсене на кредит, както и във функцията на търсенето на общите и дългосрочните кредити за предприятия. Ускоряването на годишния темп на инфлация води до засилване на търсенето само на жилищни кредити, докато при другите пазарни сегменти не е статистически значим фактор. Това може да се интерпретира като свидетелство, че в повечето случаи при вземането на решения за евентуално поемане на нови кредитни задължения икономическите агенти се ръководят главно от номиналната, а не от реалната цена на кредита. Положителният знак на коефициента пред имплицитния лихвен процент по външния дълг на частния небанков сектор потвърждава хипотезата, че относителното поевтиняване на заемните средства от местната банкова система стимулира общото търсене на отпусканите от нея кредити, като този ефект е статистически значим в модела за общия кредит. Коефициентът пред годишния темп на изменение на цените на жилищата има очаквания положителен знак, като оказва статистически значимо влияние в случая на общото търсене на кредити, общия обем на кредитите за домакинства и жилищните кредити. Този резултат е в съответствие с допускането, че повишаването на цените на жилищата формира оптимистични очаквания за стойността на потенциалните обезпечения по ипотечните заеми и за капацитета на домакинствата да поемат нови кредитни задължения, което стимулира търсенето на кредити.

Резултатите за функцията на предлагане на заемни ресурси показват, че увеличението на кредитния капацитет на банковата система

води до повишаване на общото предлагане на кредити, като този ефект се наблюдава и при предлагането на общия обем на корпоративните кредити, дългосрочните кредити за предприятия и жилищните кредити. Коефициентът пред спреда между лихвения процент по кредитите и ЮРИБОР е оценен с очаквания отрицателен знак, отразявайки ограничаващото влияние на възприятието за риск върху предлагането на кредити, като този ефект е статистически значим в уравненията за общото предлагане на кредит и предлагането на кредити за предприятия. Според получените резултати нивата на ЮРИБОР не оказват статистически значимо влияние върху предлагането на кредити. Въпреки че може да се разглежда като следствие от несъвършенствата на избрания начин за отчитане на възвръщаемостта на алтернативните на кредитната дейност инвестиции, този резултат би могъл да се интерпретира и като свидетелство, че банките не възприемат вложенията в нискорискови финансови инструменти като дългосрочен заместител на кредитната дейност. Накрая, коефициентът пред променливата за растежа на цените на жилищата е положителен и статистически значим в моделите за общото предлагане на заемни ресурси, както и за предлагането на общия обем на кредити за домакинства и жилищните кредити. Това е в съответствие с хипотезата, че повишаването на цените на недвижимите имоти увеличава склонността на банките да отпускат ипотечни кредити заради формирането на очаквания за повишаване на стойността на предоставените обезпечения.

Въз основа на динамиката на обяснителните променливи и получените оценки на техните коефициенти са генерирани динамични редове за търсенето и предлагането на съответния вид кредити. В таблица 5 в приложение 1 са представени резултатите от теста на *Johansen* за наличие на коинтеграция между тях и наблюдаваните пазарни обеми. Тъй като във всички случаи е налице поне един коинтеграционен вектор, оценените функции на търсенето и предлагането на кредити могат да се интерпретират като дългосрочни връзки.

Една от задачите на изследването е да се идентифицират периодите с по-значително несъответствие между търсенето и предлагането на заемни ресурси. Този анализ е приложен както спрямо общия обем на новоотпуснатите кредити, така и по отношение на различните сегменти на кредитния пазар. За всеки от тях са построени по две графики, едната от които сравнява номиналните стойности на търсенето и предлагането на кредити, а другата представя степента на несъответствие между търсенето на заемни ресурси и кредитното предлагане.

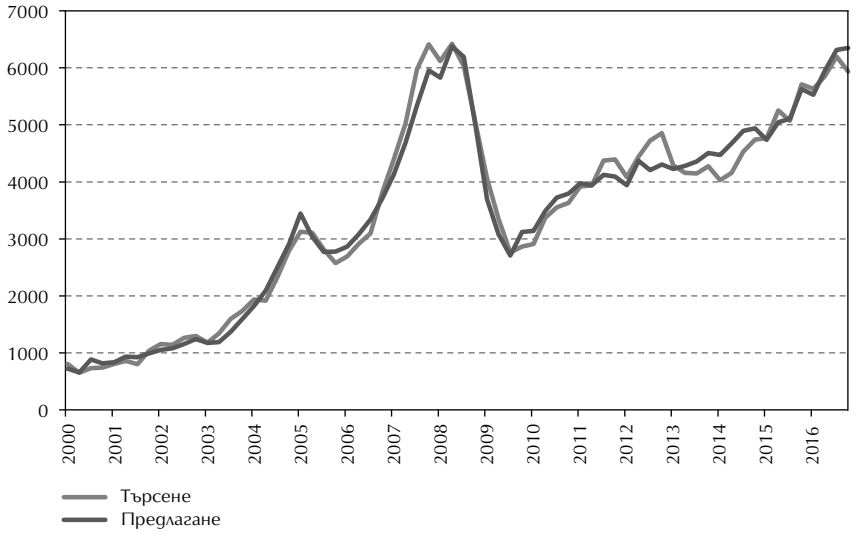
Като цяло резултатите показват, че въпреки някои временни несъответствия е налице висока степен на корелация между търсенето и предлагането на кредити (графика 9). Бързото нарастване на кредита в България през периода 2004–2008 г. е следствие в еднаква степен от активизирането на търсенето и предлагането на заемни ресурси. Аналогично резкият спад в обема на новоотпуснатите кредити в периода на глобалната криза и на последвалото бавно икономическо възстановяване е обусловен от отслабване както на търсенето, така и на предлагането на кредити.

Независимо че като цяло резултатите показват висока степен на корелация между търсенето и предлагането на кредити, интерес представляват периодите, в които търсенето на заемни ресурси се разминава по-съществено с нивото на предлагането им от страна на банките. Графика 10 илюстрира разликата между общото търсене на кредити от предприятията и домакинствата, от една страна, и предлагането на заемни ресурси от банките. Несъответствието е представено както в номинално изражение, така и като относителен дял от предлагането на кредити за тези сектори. Там могат да се различат няколко периода, в които предлагането на кредити е било по-ниско от търсенето на заемни ресурси и съответно е оказвало ограничаващо влияние върху пазарните обеми. Първият от тях е периодът 2002–2004 г., когато въпреки започналото активизиране на кредитната дейност поведението на банките все още се отличава с предпазливата кредитна политика, възприета след кризата от 1996–1997 г., както и с нагласата голяма част от привлечените ресурси да бъдат инвестирани в чуждестранни активи заради сравнително високата им доходност. Следващият епизод, в който търсенето на заемни ресурси надхвърля предлагането им, е времето от началото на 2007 г. до края на 2009 г. Този период обхваща както години на икономически бум (до края на 2008 г.), така и времето, в което глобалната и икономическа криза обуслови спада на икономическата активност в българската икономика. В първия случай бързият икономически растеж подхранваше търсенето на кредити, създавайки прекомерно оптимистични нагласи в предприятията и домакинствата, докато кредитните институции бяха обект на мерки от страна на БНБ, целящи по-умерено разширяване на кредитната дейност. По време на кризата настъпи съществено свиване и на търсенето, и на предлагането на кредити, като въпреки това по-силно ограничаващо влияние върху пазарния обем имаше поведението на банките и по-конкретно изострената им чувствителност към кредитния риск. По-нататък превишение на търсенето над предлагането на кредити се наблюдава през още два епизода от извадката:

от средата на 2011 г. до средата на 2013 г. (заради нарасналите нужди на предприятията от заемни ресурси в условията на започналото плавно икономическо възстановяване и наблюдаваното тогава временно увеличение на цената на финансирането от чужбина), както и от началото на 2015 г. до средата на 2016 г. В последния случай обемът на предлагането на кредити бе повлиян в известна степен от подготовката на предстоящите тогава преглед на качеството на активите и стрес тест на банковата система, провеждането на които накара банките да променят фокуса на внимание от отпускането на кредити към предприемането на по-активни действия по отношение на състоянието на балансите и управлението на кредитния риск.

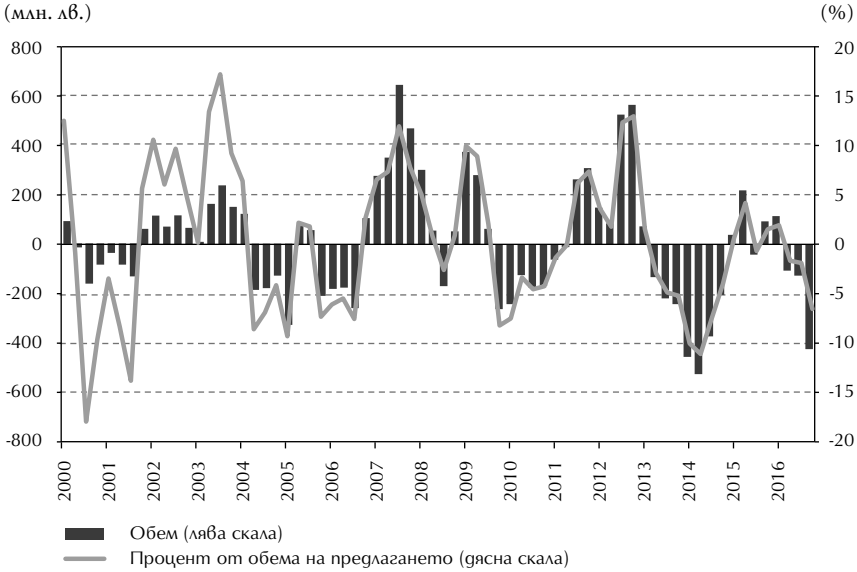
Графика 9. **Търсене и предлагане на кредити за предприятия и домакинства**

(млн. лв.)



Източници: БНБ, собствени изчисления.

Графика 10. **Превишение на търсенето над предлагането на кредити за предприятия и домакинства**



Източници: БНБ, собствени изчисления.

Наред с това особено внимание заслужават периодите, в които обемът на предлаганите от банките заемни ресурси надхвърля тяхното търсене. Анализирането на такива ситуации е наложително заради потенциалната възможност те да се дължат на повишена склонност към поемане на риск от страна на кредитните институции и обусловената от това склонност към бързо и несъобразено с капацитета на кредитополучателите разширяване на кредитната дейност. Сред епизодите, в които предлагането на кредити надхвърля търсенето на заемни ресурси, е периодът 2004–2006 г., когато бързото нарастване на привлечените средства в банковата система и стремежът на банките за привличане на нови кредитополучатели в условията на прекомерно оптимистични оценки и очаквания за състоянието на икономическата среда се отразяваха в значително нарастване на предлагания обем кредити, независимо от ограничаващия ефект на взетите от БНБ мерки. Тъй като епизоди с превишение на предлагането над търсенето на кредити има и в годините на бавно възстановяване след глобалната криза (периодите от началото на 2010 г. до средата на 2011 г. и от средата на 2013 г. до края на 2014 г.), това подчертава

необходимостта от засилено внимание към адекватността на прилаганите кредитни стандарти не само при кредитна експанзия, но и в условията на сравнително потисната кредитна дейност.

Заключение

Значението на кредитната дейност на банките за общото състояние на икономическата среда произтича от ролята на кредита в предавателния механизъм на паричната политика, влиянието на финансовото посредничество върху степента на икономическо развитие и наличието на зависимост между динамиката на кредитната дейност и състоянието на финансовата стабилност. Целта на това изследване е в рамките на неравновесен модел на кредитния пазар да бъдат оценени функции на търсенето и предлагането на отделните видове кредити в България, като се даде количествена оценка на влиянието, което различни фактори оказват върху тяхната динамика. Друга цел на анализа е да се определи относителната роля на търсенето и предлагането на кредити при формирането на обемите на кредитния пазар в България, като се идентифицират периодите с по-значително несъответствие между търсенето и предлагането на заемни ресурси.

Като цяло резултатите показват, че влиянието на разглежданите фактори за търсенето и предлагането на кредит съответства на предварително направените хипотези и е сходно с установените в литературата по темата оценки. Друг извод от анализа е, че в рамките на разглеждания период са налице някои временни несъответствия между търсенето и предлагането на кредити, но въпреки това има висока степен на корелация между търсените и предлаганите обеми. Бързото нарастване на кредита в България през периода 2004–2008 г. е следствие в еднаква степен от активизирането на търсенето и предлагането на заемни ресурси. Аналогично, резкият спад в обема на новоотпуснатите кредити в периода на глобалната криза и последвалия период на бавно икономическо възстановяване е обусловен от отслабване както на търсенето, така и на предлагането на кредити.

Оценяването на относителната роля на търсенето и предлагането на кредити при формирането на обемите на пазара на заемни ресурси е важно по две причини. Първо, то дава възможност да се обогати представата за факторите,определящи динамиката на кредита, което би позволило по-задълбочено изследване на ролята на кредита в предавателния механизъм на паричната политика. Например, предпоставка за оценяването на ролята на кредитния канал е разграничаването на динамиката на предлагането от тази на търсенето на кредити. Дру-

гата потенциална насока на приложение на резултатите от подобен тип анализ е свързана с провеждането на макропруденциалната политика и по-конкретно с избора на такива инструменти за ограничаване на системния риск чрез влияние върху динамиката на кредита, които да са съобразени с това, дали наблюдаваните тенденции на кредитния пазар са обусловени от търсенето или предлагането на заемни ресурси. По този начин оценките за относителната роля на търсенето и предлагането на кредити за формирането на кредитната динамика в даден период биха могли да осигурят допълнителна информация в процеса на определяне на дизайна на макропруденциалните инструменти, използвани за ограничаване на цикличния системен риск.

Приложения

1. Тестове за единичен корен и коинтеграция

Таблица 3. ***ADF* тест за единичен корен**

Променлива	Нива		Първи разлики		Степен на интеграция
	Лагове	ADF статистика	Лагове	ADF статистика	
leh	0	-1.69	0	-8.99 ***	I(1)
le	0	-2.02	0	-10.74 ***	I(1)
lest	1	-2.27	0	-12.66 ***	I(1)
lelt	1	-2.86 *	0	-11.23 ***	I(1)
lh	1	-2.16	0	-5.40 ***	I(1)
lcons	1	-2.07	0	-5.65 ***	I(1)
lhousing	1	-1.51	0	-4.67 ***	I(1)
il eh	0	-0.71	0	-11.69 ***	I(1)
il e	0	-1.09	0	-12.41 ***	I(1)
i_stle	0	-1.11	0	-11.95 ***	I(1)
i_ltle	1	-1.18	0	-11.85 ***	I(1)
il h	0	-0.78	0	-6.83 ***	I(1)
il_consl	0	-1.37	0	-7.42 ***	I(1)
il_hl	0	-1.62	0	-7.30 ***	I(1)
i ed	1	-1.54	0	-7.60 ***	I(1)
euribor 3m	1	-1.55	0	-4.52 ***	I(1)
blc	0	-2.73 *	0	-8.34 ***	I(1)
y	3	-1.88	2	-2.02	I(2)
infl	4	-1.46	3	-6.00 ***	I(1)
hpag	2	-3.09 **	2	-5.37 ***	I(0)

Забележка: Нулевата хипотеза е за наличие на единичен корен. Знакът ***/**/* означава отхвърляне на нулевата хипотеза при ниво на значимост съответно 1%, 5% и 10%. Спецификацията включва константа. Броят на лаговете е избран на базата на *Schwarz Info Crirterion.*

Таблица 4. ***KPSS* тест за единичен корен**

Променлива	LM статистика		Степен на интеграция
	Нива	Първи разлики	
leh	0.73 **	0.12	I(1)
le	0.85 ***	0.07	I(1)
lest	0.75 ***	0.19	I(1)
lelt	0.92 ***	0.45 *	I(1)
lh	0.59 **	0.25	I(1)
lcons	0.51 **	0.22	I(1)
lhousing	0.77 ***	0.19	I(1)
il eh	0.96 ***	0.10	I(1)
il e	0.80 ***	0.12	I(1)
i stle	0.86 ***	0.10	I(1)
i ltle	0.93 ***	0.19	I(1)
il h	0.93 ***	0.11	I(1)
il consl	0.81 ***	0.14	I(1)
il hl	0.99 ***	0.22	I(1)
i ed	0.78 ***	0.07	I(1)
euribor 3m	0.79 ***	0.07	I(1)
blc	1.05 ***	0.63 **	I(2)
y	0.98 ***	0.40 *	I(1)
infl	0.71 **	0.03	I(1)
hpag	0.25	0.07	I(0)

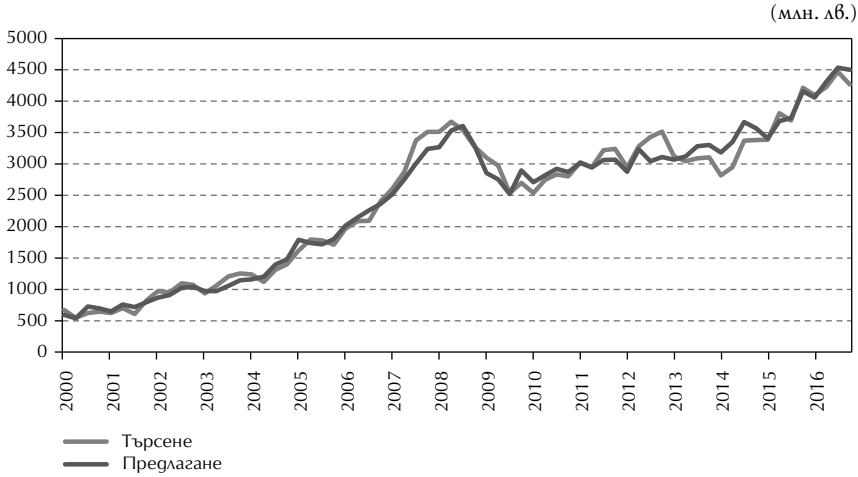
Забележка: Нулевата хипотеза е, че съответната променлива е стационарна. Знакът ***/**/* означава отхвърляне на нулевата хипотеза при ниво на значимост съответно 1%, 5% и 10%.

Таблица 5. **Тест на *Johansen* за коинтеграция на наблюдавания пазарен обем с оценките за търсенето и предлагането на съответния вид кредити**

		Hypothesized No. of CE(s)	Trace Statistic	Prob.	Max-Eigen Statistic	Prob.
Търсене на кредити за предприятия и домакинства	Търсене	r = 0	23.3	0.0027	17.4	0.0155
		r < 1	5.9	0.0152	5.9	0.0152
	Предлагане	r = 0	18.9	0.0147	14.9	0.0396
		r < 1	4.0	0.0453	4.0	0.0453
Кредити за предприятия	Търсене	r = 0	14.9	0.0618	9.7	0.2348
		r < 1	5.2	0.0224	5.2	0.0224
	Предлагане	r = 0	14.9	0.0612	10.8	0.1636
		r < 1	4.1	0.0433	4.1	0.0433
Краткосрочни кредити за предприятия	Търсене	r = 0	20.0	0.0096	13.0	0.0788
		r < 1	7.1	0.0079	7.1	0.0079
	Предлагане	r = 0	15.5	0.0496	10.3	0.1917
		r < 1	5.2	0.0226	5.2	0.0226
Дългосрочни кредити за предприятия	Търсене	r = 0	11.4	0.1872	6.7	0.5285
		r < 1	4.7	0.0294	4.7	0.0294
	Предлагане	r = 0	14.6	0.0679	7.9	0.3845
		r < 1	6.7	0.0099	6.7	0.0099
Кредити за домакинства	Търсене	r = 0	26.2	0.0009	18.4	0.0103
		r < 1	7.7	0.0054	7.7	0.0054
	Предлагане	r = 0	29.1	0.0003	22.1	0.0024
		r < 1	7.0	0.0082	7.0	0.0082
Потребителски кредити	Търсене	r = 0	27.1	0.0006	18.4	0.0105
		r < 1	8.7	0.0031	8.7	0.0031
	Предлагане	r = 0	28.0	0.0004	19.2	0.0077
		r < 1	8.8	0.0030	8.8	0.0030
Жилищни кредити	Търсене	r = 0	20.7	0.0076	16.8	0.0197
		r < 1	3.9	0.0488	3.9	0.0488
	Предлагане	r = 0	18.8	0.0155	14.4	0.0480
		r < 1	4.4	0.0363	4.4	0.0363

2. Оценки на търсенето и предлагането на различните видове кредити

Графика 11. **Търсене и предлагане на кредити за предприятия**



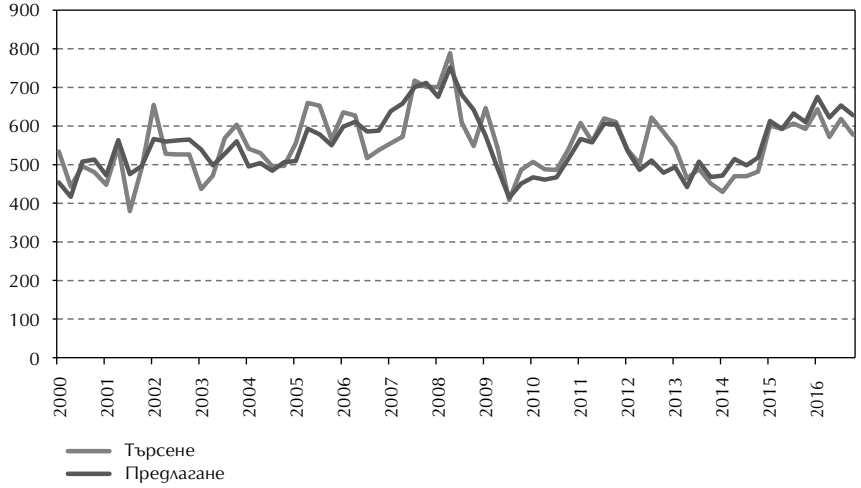
Източници: БНБ, собствени изчисления.

Графика 12. **Превишение на търсенето над предлагането на кредити за предприятия**

Източници: БНБ, собствени изчисления.

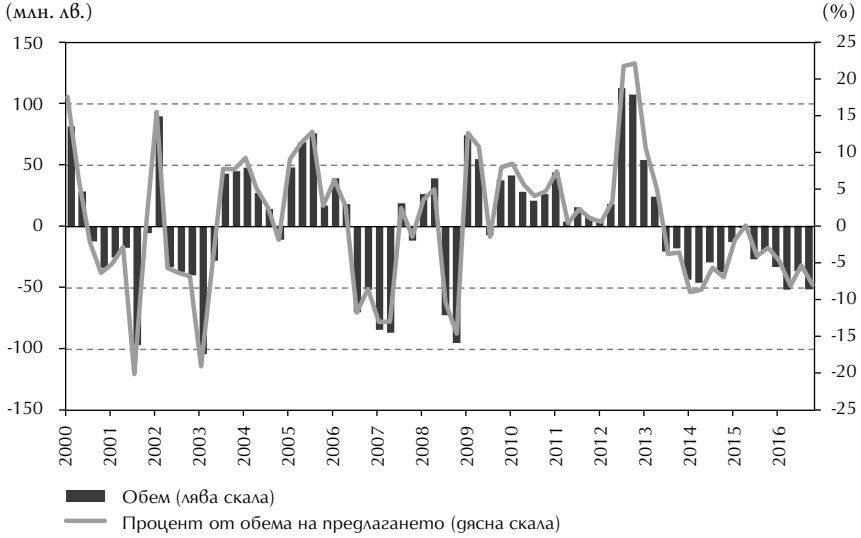
Графика 13. **Търсене и предлагане на краткосрочни кредити за предприятия**

(млн. лв.)



Източници: БНБ, собствени изчисления.

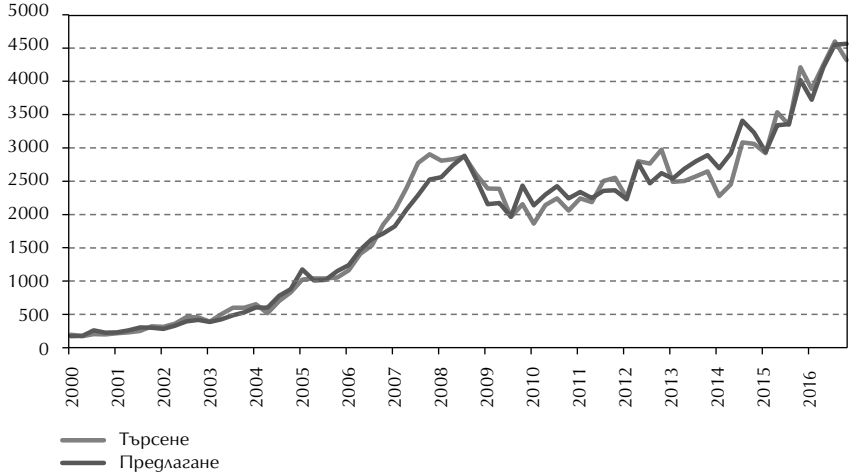
Графика 14. **Превишение на търсенето над предлагането на краткосрочни кредити за предприятия**



Източници: БНБ, собствени изчисления.

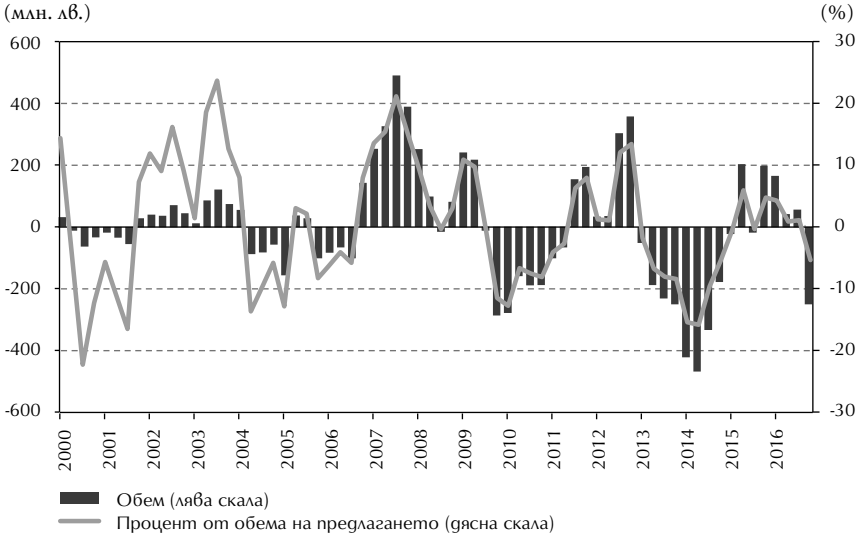
Графика 15. **Търсене и предлагане на дългосрочни кредити за предприятия**

(млн. лв.)



Източници: БНБ, собствени изчисления.

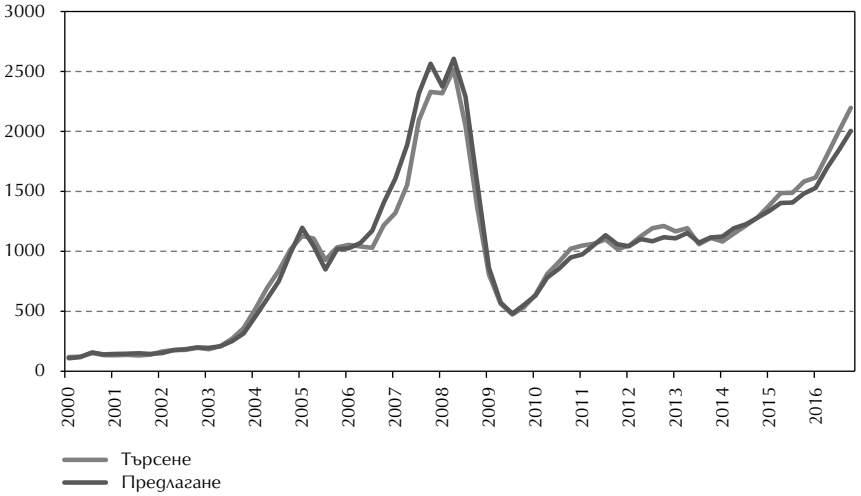
Графика 16. **Превишение на търсенето над предлагането на дългосрочни кредити за предприятия**



Източници: БНБ, собствени изчисления.

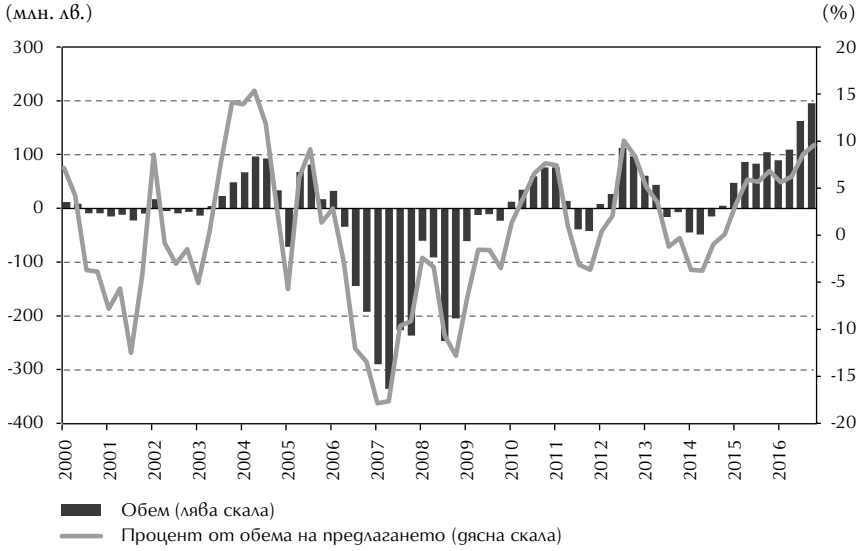
Графика 17. **Търсене и предлагане на кредити за домакинства**

(млн. лв.)



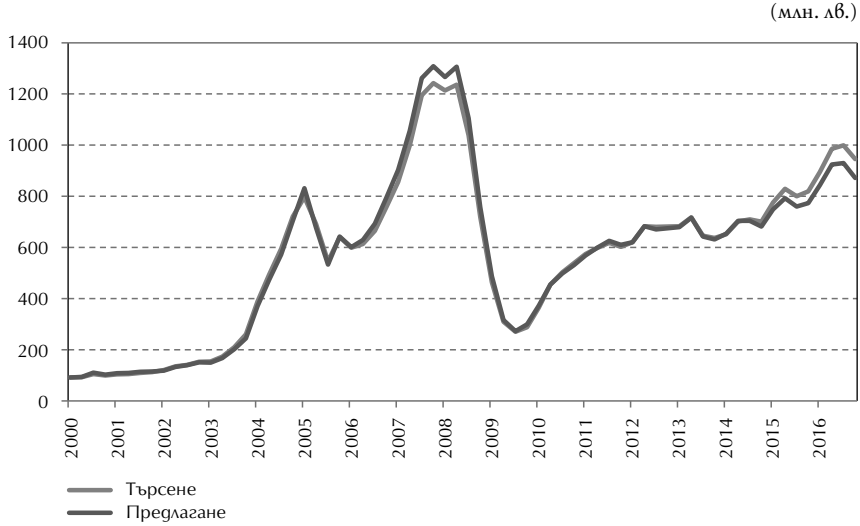
Източници: БНБ, собствени изчисления.

Графика 18. **Превишение на търсенето над предлагането на кредити за домакинства**



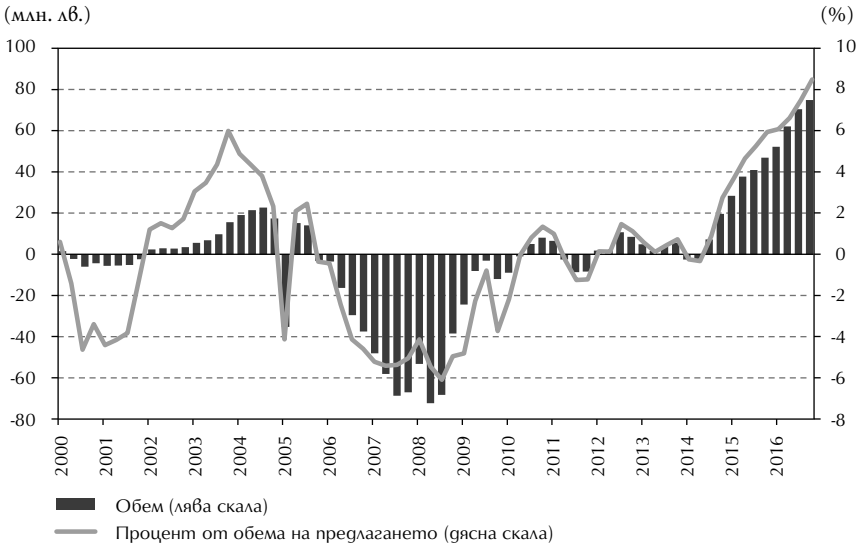
Източници: БНБ, собствени изчисления.

Графика 19. **Търсене и предлагане на потребителски кредити**



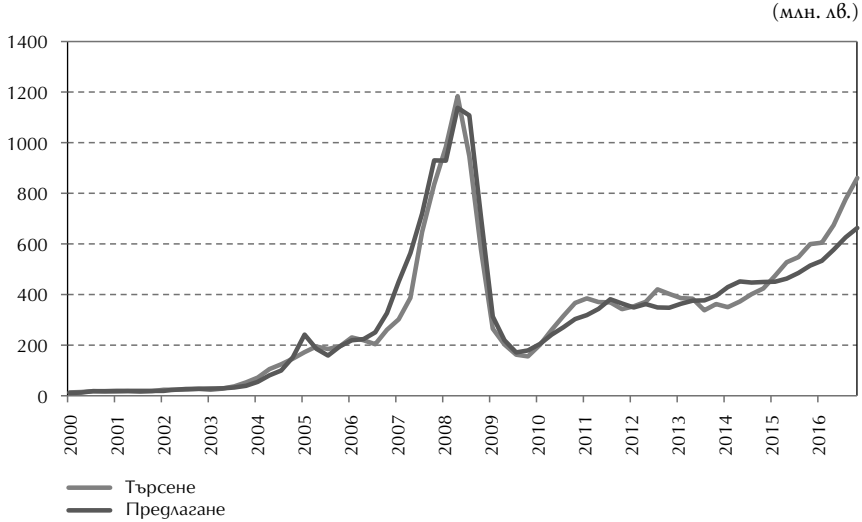
Източници: БНБ, собствени изчисления.

Графика 20. **Превишение на търсенето над предлагането на потребителски кредити**



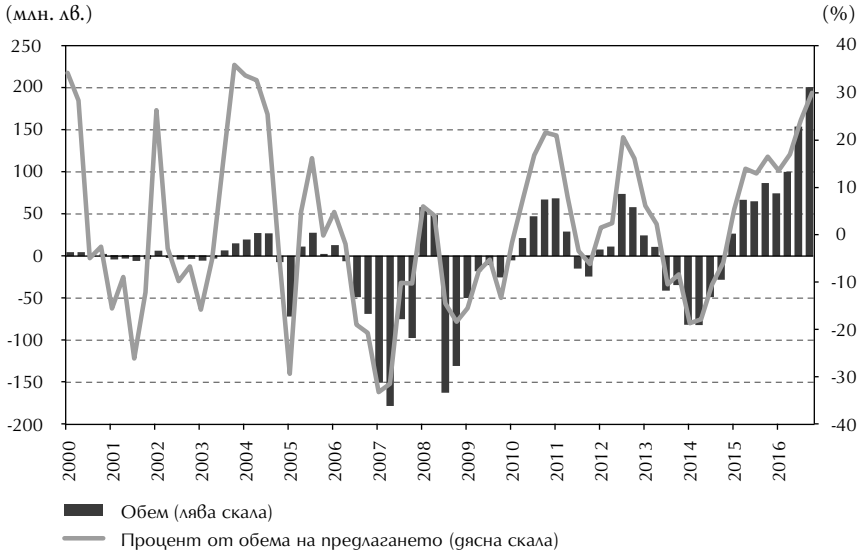
Източници: БНБ, собствени изчисления.

Графика 21. **Търсене и предлагане на жилищни кредити**



Източници: БНБ, собствени изчисления.

Графика 22. **Превишение на търсенето над предлагането на жилищни кредити**



Източници: БНБ, собствени изчисления.

Използвана литература

Пешев, П. Моделиране на търсенето и предлагането на кредити в България. – В: Икономическа мисъл, 2015, № 5.

Христов, К., М. Михайлов. Кредитна активност на търговските банки и рациониране на кредитния пазар в България. София: БНБ, 2002, Дискусионни материали № 23, 52 с.

Allain, L. and N. Oulidi. Credit Market in Morocco: A Disequilibrium Approach, IMF Working Paper WP/09/53, March 2009.

Atanasova, C. and N. Wilson. Disequilibrium in the UK corporate loan market. *Journal of Banking & Finance*, Vol. 28, с. 595–614.

Barajas, A. and R. Steiner. Credit Stagnation in Latin America, IMF Working Paper WP/02/53, March 2002.

Čeh, A. M., M. Dumičić and I. Krznar. A Credit Market Disequilibrium Model And Periods of Credit Crunch, Croatian National Bank, Working Paper No. 28, January 2011.

Fair, R. C. and D. M. Jaffee. Methods of Estimation for Markets in Disequilibrium. *Econometrica*, Vol. 40, No. 3, May 1972, с. 497–514.

Fair, R. C. and H. H. Kelejian. Methods of Estimation for Markets in Disequilibrium: A Further Study. *Econometrica*, Vol. 42, No. 1, January 1974, с. 177–190.

Farinha, L. and S. Felix. Credit rationing for Portuguese SMEs, Banco de Portugal Financial Stability Papers, No. 3, November 2014.

Ghosh, A. R. and S. R. Ghosh. East Asia in the Aftermath: Was There a Crunch?, IMF Working Paper WP/99/38, March 1999.

Hofmann, B. The determinants of private sector credit in industrialised countries: do property prices matter?, BIS Working Paper No. 108, December 2001.

Laffont, J.-J. and R. Garcia. Disequilibrium Econometrics for Business Loans. Econometrica, Vol. 45, No. 5, July 1977, с. 1187–1204.

Maddala, G. S. and F. D. Nelson. Maximum Likelihood Methods for Models of Markets in Disequilibrium. *Econometrica*, Vol. 42, No. 6, November 1974, с. 1013–1030.

Melitz, J. and M. Pardue. The Demand and Supply of Commercial Bank Loans. *Journal of Money, Credit and Banking*, Vol. 5, No. 2, May 1973, с. 669–692.

Nehls, H. and T. Schmidt. Credit Crunch in Germany?, Rheinisch-Westfälisches Institut für Wirtschaftsforschung, Discussion Papers No. 6, September 2003.

Pazarbaşioğlu, C. A Credit Crunch? Finland in the Aftermath of the Banking Crisis, IMF Staff Papers, Vol. 44, No. 3, September 1997, с. 315–327.

Poghosyan, T. Slowdown of Credit Flows in Jordan in the Wake of the Global Financial Crisis: Supply or Demand Driven?, IMF Working Paper WP/10/256, November 2010.

Sealey, C. W. Credit Rationing in the Commercial Loan Market: Estimates of a Structural Model Under Conditions of Disequilibrium. *The Journal of Finance*, Vol. 34, No. 3, June 1979, c. 689–702.

Stiglitz, J. and A. Weiss. Credit Rationing in Markets with Imperfect Information. *American Economic Review*, Vol. 71, No. 3, c. 393–410.

DP/108/2017

Macroprudential Bank Stress-Testing – Theoretical Fundamentals, Framework, and Challenges

Ventsislav Hristev

Abstract. This study provides a current and comprehensive overview about the theory and practice of macroprudential bank stress-testing. It builds on the existing stress test literature and covers theoretical fundamentals, framework, and challenges. First, a review of underlying aspects such as definitions, characteristics, principles, and mechanism of transmission reveals the art of stress-testing with its benefits and complexities. Next, the main elements for a fitting framework are examined in details: preparation, application, presentation of results, and if applicable, follow-up and feedback. Finally, remaining challenges and possible ways forward are addressed given the stress-testing development is far from complete. The paper argues that macroprudential stress tests are effective for assessment of banks' resilience when appropriate design and expert judgement are well combined. Thus, they enhance monitoring of potential vulnerabilities to the banking system, add valuable input to the process of implementing policy actions, and consequently contribute to ensuring financial stability.

Резюме. Това изследване прави текущ и цялостен преглед на теорията и практиката на провеждане на макропруденциални стрес тестове на банките. То стъпва на съществуващата литература за стрес тестове и обхваща теоретичните основи, рамката и предизвикателствата. На първо място, преглед на базовите аспекти, като дефиниции, характеристики и механизъм на действие, разкрива изкуството на провеждане на стрес тестовете заедно с неговите ползи и сложност. След това подробно са проучени основните елементи на една подходяща рамка: подготовка, прилагане, представяне на резултатите и, ако е приложимо, последващи действия и обратни връзки. Накрая се обсъждат оставащите предизвикателства и възможните пътища напред, като се има предвид, че развитието на стрес тестовете далеч не е завършено. Изследването твърди, че макропруденциалните стрес тестове са ефективни за оценка на устойчивостта на банките при умело съчетаване на подходящ дизайн и експертна преценка. По този начин те усъвършенстват наблюдението върху възможните уязвимости на банковата система, добавят съществен елемент към процеса на прилагане на действия по провежданата политика и така допринасят за поддържането на финансовата стабилност.

Ventsislav Hristev is head of team in the Macroprudential Supervision and Financial Stability Directorate of the Bulgarian National Bank.

E-mail address: hristev.v@bnbank.org

Acknowledgment:

I would like to thank two anonymous referees for their constructive suggestions and insightful comments. The support and guidance of my colleagues from the Macroprudential Supervision and Financial Stability Directorate are also highly appreciated.

Contents

Introduction .....
1. Theoretical fundamentals .....
2. Framework.....
   2.1. Preparation .....
   2.2. Application .....
   2.3. Presentation of results.....
   2.4. Follow-up and feedback .....
3. Challenges.....
Conclusion.....
References .....

Introduction

Macroprudential Stress-Testing (MpSTing) has become a prominent practice in the modern banking world for risk assessment. It emerged widely as a crisis management instrument in the aftermath of the global financial crisis 2007/8, but since then has been established as a regular supervisory tool for financial stability. The Federal Reserve (FED) and the Bank of England (BoE) have issued official frameworks for annual public stress-testing, while the European Banking Authority (EBA) is carrying out an EU-wide stress test biennially. In the European Union, competent authorities are required under the CRD IV[1] to run at least annually supervisory bank stress tests to facilitate the review and evaluation process. More globally, MpSTs play a major role in the International Monetary Fund Financial Sector Assessment Programs (IMF FSAPs). Together with an asset quality review, stress-testing was also an integral part in the recent comprehensive assessments in Ireland (2011), Spain (2012), Greece (2013), Slovenia (2013), Bulgaria (2016), and in the process of establishing the Banking Union by the European Central Bank (ECB) during 2014 - 2016.

At present, there is a vast and on-going body of research about the theory and practice of MpSTing. But in a constantly developing and complex stress test universe, a need for an up-to-date review emerges continuously. In that manner, the goal of this paper is to provide a current and comprehensive analysis about the MpSTing that covers theoretical fundamentals, framework, and challenges. By building on the existing prolific stress test literature and incorporating both the academic and practical side of the matter, the study utilizes "lessons learned" from past experiences, formulates guidelines for effective MpST, and adds another piece to the debate over how best to carry out the macroprudential exercise among regulators, academics, and bankers. Accordingly, this research contributes to the better understanding of MpSTs in several ways. First, the article updates the theory of MpSTing with the most present concepts. It also identifies specific features and presents taxonomy of MpSTs by a rich set of leading indicators. Next, in this study I propose a detailed framework for MpSTing. The main elements for the overall stress test design are thoroughly covered. Moreover, different options for approaching them are outlined together with corresponding strengths and weaknesses. Where an issue for practical implementation arises, methodological guidance is provided. Finally, the current challenges ahead of MpSTing are discussed. The examination is accompanied with recommendations from my point of view in order to cover concrete setups and circumstances for conducting MpSTs.

The current review shows that MpSTing is beneficial for promoting financial stability if appropriate design and expert judgement are well combined. The

[1] Directive 2013/36/EU of the European Parliament and of the Council of 26 June 2013, art. 100 (1).

underlying fundamentals of a stress test reveal its analytical advantages for identifying potential vulnerabilities to the banking system and mechanism for increasing bank resilience. A fitting framework, which consists of clear goals, robust process (assumptions, methodology, scenarios, risk modeling, results, and follow-up), and sufficient transparency ensures the relevance and credibility of the exercise. Present challenges display that work on MpSTs' development is far from complete, as their application differs across jurisdictions and in time. Acknowledging the inherent limitations and imperfections of stress-testing, it serves as a complement, rather than as a substitute, to the existing regulatory and supervisory tools. As a result, an effective and unambiguous MpST enhances the process of assessing banks' resilience, which ultimately provides valuable input for policy actions purposes.

The rest of the paper continues as follows. Section 2 provides an overview of the theoretical fundamentals, which include definitions, characteristics, principles, and mechanism of transmission. Taxonomy for MpST types is also formulated. Section 3 presents the main elements in the stress test framework: preparation, application, presentation of results, and if applicable follow-up and feedback. Section 4 discusses the current challenges about MpSTing and possible ways forward, while Section 5 concludes.

1. Theoretical fundamentals

The perception of MpSTs has evolved over time. In the early stages of application, the concept of stress-testing is described in a broad manner by IMF 2003 and IMF, World Bank 2005 as a "what if" analysis or a "rough estimate" of a given world (portfolio) in which certain major changes in one or more variables (risk factors) are assumed. More narrowly, several studies attribute the macro stress test to "a range of techniques" for quantitatively assessing the resilience of a financial system to "extreme but plausible" macroeconomic and contagion shocks (Sorge 2004, Čihák 2004, and ECB 2006). A decade later, central banks, supervisory authorities, international organizations, and policy bodies view the MpSTs as important "tools for financial stability" both in advanced and emerging countries all over the world (Jobst et al. 2013, Hristev 2014, and Hsiao, Jiang 2015). With regards to the use of stress-testing to manage systemic risk, Henry, Kok 2013 show that it has proved valuable as an early warning device for identifying potential systemic fragilities, while Borio et al. 2012 argue for its effectiveness as a crisis management and resolution tool. Recently, Dent et al. 2016 provide a governance point of view on the matter and describe a "concurrent" stress test as an exercise conducted under the direction of an official body (usually the banking authority) and in which the entire balance sheets of selected group of banks are simultaneously assessed. To date the most comprehensive

definition for MpSTing is given by the IMF 2012 study, which builds on the quantitative perception of the stress test to the view of a whole process:

"A complete stress testing exercise involves choices on the coverage of institutions, risks, and scenarios; the application of a quantitative framework to link various shock scenarios to solvency and liquidity measures; a strategy for the communication of the results; and follow-up measures, if warranted."

From the presented concepts, the main characteristics could be identified. First, the MpST is a forward-looking technique applied on a system-wide level. This instrument determines the current capacity of a relevant set of banks or the system as a whole to absorb hypothetical future losses. Second, stress-testing is a quantitative (numerical) exercise. It provides for a certain awareness of potential shocks by projecting an estimate of their materiality. Third, it is comprehensive by allowing for incorporation of multiple risks. MpSTs are complementary to the Basel concept of "one instrument – one risk" by combining different risks such as credit, market (interest and FX), operational, sovereign, funding, etc. into one framework. Fourth, stress-testing encompasses "tail risk" events. Simulated shocks are extreme by design and drawn from the tail of the distribution. Fifth, stress tests can be public. If mandated, the process could be accompanied with disclosure of results to foster market discipline. Sixth, being used as a supervisory tool, MpSTing is actionable. In response to the exercise, policy measures including capital and liquidity requirements may be taken. Finally, macro stress tests are dynamic and flexible. They incorporate series of approximations, assumptions, and constraints to project the evolution of the balance sheet, credit quality, profitability, liquidity position, interbank exposures, etc. Accordingly, MpSTing loosens the standard (point-in-time) financial analyses by stretching the accounting and regulatory principles and thus contributes to the financial stability assessment by providing additional insights on bank's resilience to negative extremes and on shock propagation channels.

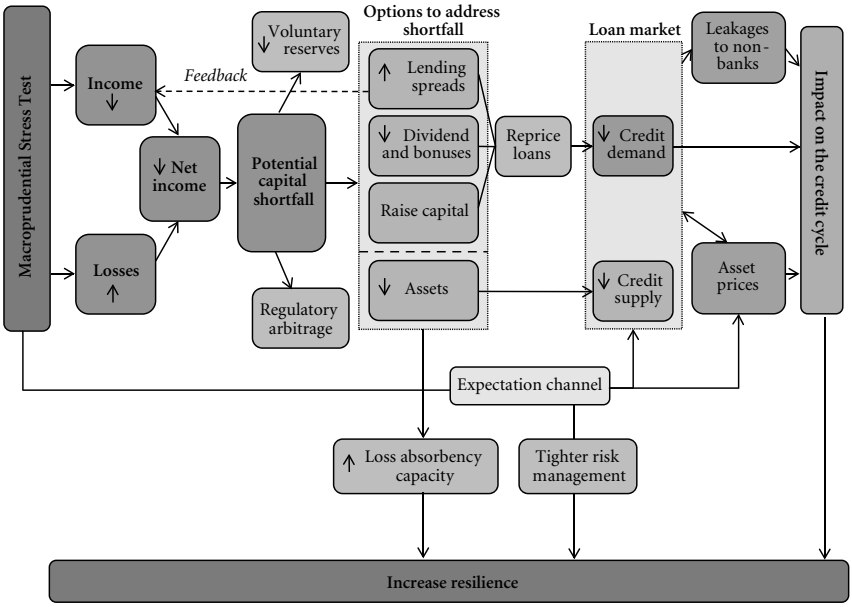
Along with the stress test characteristics expert judgment is applied throughout the process. Expertise is important for validating risk models, but also for fine-tuning and calibrating adversity (Moretti et al. 2008). In Committee on the Global Financial System (CGFS) 2012, it is further elaborated that as modeling uncertainties remain present, sound judgement is required during the conduct of stress-testing as well as for the interpretation of outcomes. As a result, MpSTs inherently go along with the impediment of error and inaccuracy. Hilbers et al. 2004 emphasize on that complexity and points out to the difficulty of regarding the simulation as a "precise tool with scientific accuracy". The works of Quagliariello 2009, Henry, Kok 2013, and Jobst et al. 2013 provide additional caution on the matter of imperfection by raising the issues of associated subjectivity, evolution of risk, and impossible standardization. In that manner, the authors

express the common perception among scholars and practitioners that MpSTing is fundamentally "an art, rather than a science".

The guidelines of EBA 2015 and IMF 2012 reveal the basic principles for designing and conducting a MpST: credibility, relevance, and consistency. The first principle mainly encompasses the corporate governance of the process and its transparency. For an effective stress test that is further linked to supervisory use and actions, there should be a formal program in place with appropriate organization, structure, data support, and IT arrangements. Disclosure of sufficient information regarding the process' assumptions and outcomes should also be considered for achieving reliability. The second principle of relevance is related to the conservativeness of the exercise based on coverage of material risks and propagation channels, adequate scope, suitable scenarios, and comprehensive approach. Accordingly, the focus should be to capture the tail historical events and ultimately to incorporate unthinkable ones (black swans). The third principle of consistency ensures level playing field among participants for common application of the stress test and comparability of results.

The MpST mechanism of transmission as mapped by CGFS 2012 (p. 19) is presented in Figure 1. The direct channel of the exercise is to ultimately promote resilience by assessing bank's capital capacity to scenario shocks. Under stress, banks face constrained income and elevated losses that jointly erode the net income, which has a subsequent effect on the regulatory capital. In the case of a shortfall, a bank could internally turn to its current capital buffers or if manageable to a strategic shift aiming at less risky business. Other options when capital is not available and solvency is called into question by the MpST are to increase lending spreads, retain dividend and bonuses, issue new capital, or reduce assets. The corresponding tightening of stringent credit standards, payment restrictions, costly capital, and risky holdings sales, has a direct impact on the loss absorbing capacity. In addition, resilience could also be enhanced indirectly via the MpST effect on the credit cycle and on the expectations related to market participants' reactions and banks' risk management practices. The former stems from the loan market, in which credit demand and supply are brought down by the tightened banks' businesses ("fasten the seat belts"), while the latter stems from increased market discipline and more robust bank risk management. Kapinos et al. 2015 document that if outcomes are made public and transparent, MpSTs contribute to market discipline by reducing bank opacity and asymmetric information among market participants. Moreover, the authors continue that risk management improves as stress-testing creates incentives for better data collection, richer risk models, and stronger governance practices. The involvement of bank managers in regular exercises keeps the participants alert of potential severity and mindful of their risky behaviour, thereby mitigating "disaster myopia".

Figure 1. **Macroprudential stress test mechanism of transmission**



Source: CGFS 2012, p. 19.

There are different ways for organizing the types of stress-testing on a macro level given a predefined leading preference. In Table 1, I present a specific taxonomy of MpSTs according to their use on the horizontal and their characteristics on the vertical. The former consists of tools by their goal - for supervisory monitoring (surveillance of banks' resilience), crisis management, and calibration of macroprudential instruments (MPIs). The latter categorizes types by scope, methodology, risks, scenarios, approach, balance sheet, transparency, process, and frequency. MpSTs for supervisory purposes encompass elements from across the presented spectrum. The system-wide and macroeconomic, solvency and liquidity, bottom-up and top-down, static and dynamic stress tests fall under this classification. A particular feature of this category is the inclusion of the reverse and market-based type. The crisis management dimension is usually tailored, solvency-focused, bottom-up, accounting-based, and always by design public. For calibration of MPIs the stress-testing is typically internal, low frequent, top-down with sensitivity to a specific risk area analysis.

A vertical cross section of the stress tests' taxonomy reveals further acute links. Solvency stress tests could be static or dynamic, based on sensitivity analysis or macro scenario, and are low frequent. Liquidity stress tests on the other hand are highly or medium frequent ones, but are only structured around accounting (regulatory) data. Market-based stress test is by default tailored, contagion is a top-down type, and the hybrid approach is usually related to solvency. Moreover, complex stress tests are internal and top-down, whereas reverse stress tests involve solvency and liquidity. Public stress tests are run at least annually and incorporate a macroeconomic scenario, direct (traditional) process, and solvency focus.

Table 1. **Taxonomy of macroprudential stress tests**

		Supervisory monitoring	Crisis management	Instruments calibration
Scope	System-wide	✓		✓
	Tailored	✓	✓	
Methodology	Accounting-based	✓	✓	✓
	Market-based	✓		
Risk(s)	Solvency	✓	✓	✓
	Liquidity	✓		
	Contagion	✓		✓
	Complex	✓		
Scenario design	Sensitivity analysis	✓		✓
	Macroeconomic scenario	✓	✓	
Approach	Top-down	✓		✓
	Bottom-up	✓	✓	
	Hybrid	✓		
Balance sheet	Static	✓	✓	✓
	Dynamic	✓	✓	
Transparency	Internal	✓		✓
	Public	✓	✓	
Process	Direct (traditional)	✓	✓	✓
	Reverse	✓		
	Back-testing	✓		
Frequency	High (daily, monthly)	✓		
	Medium (quarterly)	✓		
	Low (annually, biennially)	✓	✓	✓

Note: The table uses insights from IMF 2012 with author contributions on organizing and linking the types of stress tests on the vertical and the horizontal, as well as adding the macroprudential dimension of instruments calibration.

Overall, the solvency macro stress test for supervisory monitoring is to date the most wide-spread and regular form of MpSTing. The exercise is mainly public and goes along with disclosure of either individual bank results (FED CCAR (comprehensive capital analysis and review), EBA EU-wide stress tests, ECB comprehensive assessments, BoE stress tests) or only system aggregated values (central banks' stress tests for financial stability analyses and IMF FSAPs). There are several instances of crisis management stress tests (FED SCAP (supervisory capital assessment program), CEBS EU-wide stress tests, and EU country comprehensive assessments[2]), while the application of stress tests for fine-tuning MPIs is still in early development to be officially formalized.

2. Framework

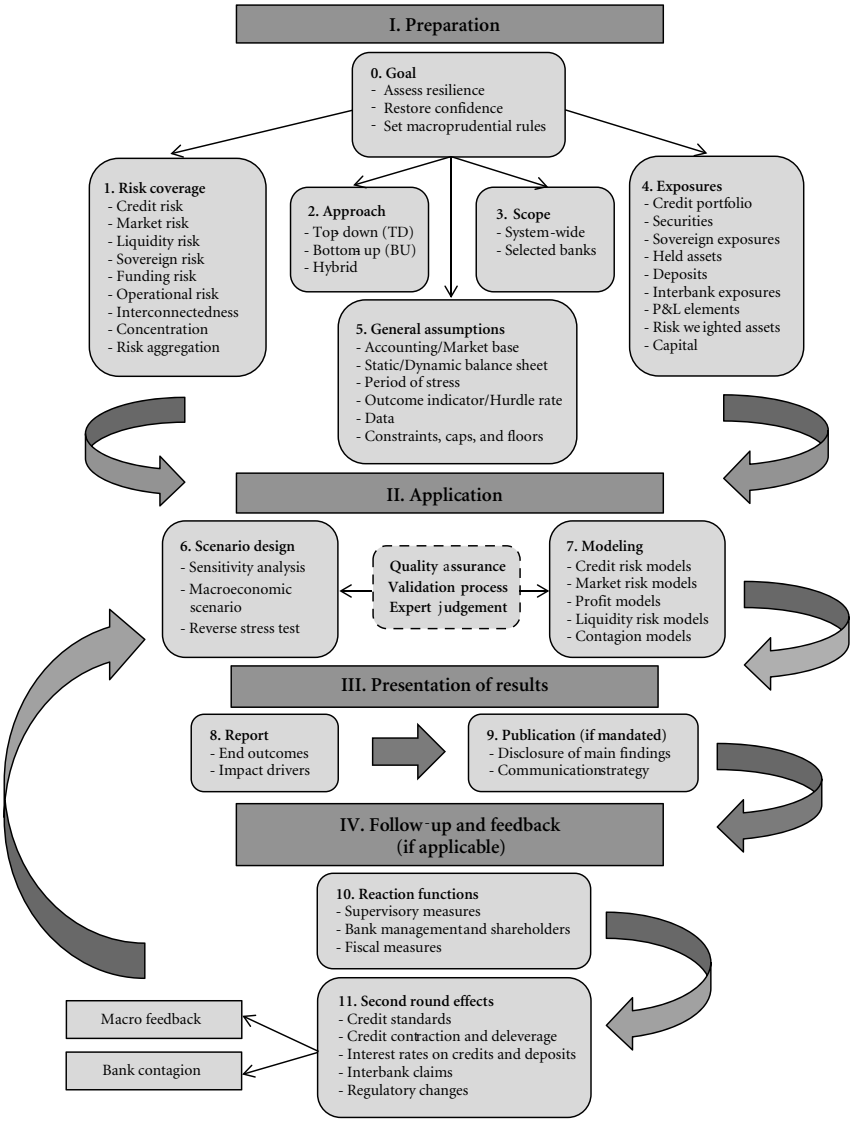
The building blocks of a MpST are preparation, application, and presentation of results (Figure 2). In more details, I propose a framework consisting of the following elements: goal definition, risk coverage, choice of approach, selection of scope and exposures under stress, structuring of general assumptions, scenarios' design, risk modeling, generating a report with assessment of main outcomes, and ways for eventual publication. The process continuous (if applicable) with follow-up measures and a reaction function from banks that could feed back in the stress scenario as dynamic and behavioral second round effects.

2.1. Preparation

In my view, defining a goal is the foundation (ground zero) of macro stress-testing on which all other elements are composed around. It is usually set according to the supervisory perspective of monitoring the potential threats to financial stability by assessing the resilience of banks and the system as a whole to withstand negative shocks. In that manner, the exercise yields thorough insights on banks' capital capacity to adversity and thus can serve as a mean for identifying pockets of vulnerability in the sector. When the MpST is carried out for analytical purposes, the level of severity is stretched, as long as it adheres to the practical golden rule of "severe enough, but plausible". Another objective follows the crisis management concept.

[2] Stress-testing plays an important part of the comprehensive assessments in Europe. To date such exercises on a country-wide (besides the ECB) basis were conducted in Ireland, Portugal, Greece, Cyprus, Spain, Slovenia, and Bulgaria. Although, the nature of assessment is similar, the originating goal and setup is different across the sample. With the exception of Bulgaria, all other comprehensive assessments had a crisis management element reflecting the need for clear capital shortfall measurement. The recapitalization process in Ireland, Portugal, Greece, and Cyprus was part of the financial assistance programmes by the Troika (EC, ECB, and the IMF), while for Spain the scope concerned only a bail-out for the banking sector. In Slovenia, there was not a programme with the Troika, as banks with capital shortfall were recapitalized with own state-aid. The stress test of Bulgarian banks, however, was of a supervisory nature, as the goal was to assess their capital resilience for absorbing negative shocks.

Figure 2. **Macroprudential stress-testing framework**

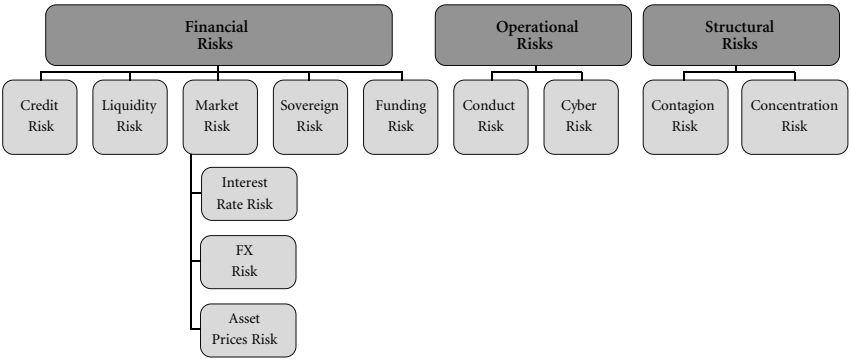


Source: Jobst et al. 2013, Henry, Kok 2013, and author contributions on designing the overall composition, adding goals, quality assurance, follow-up and feedback, and expanding other sections with further details.

The corresponding stress test aims to restore public and market confidence in the banking sector by providing an exact quantification of banks' recapitalization needs. For successful conduct of crisis stress-testing, credibility is of utmost importance (Ong, Pazarbasioglu 2013). It requires a clear objective, sufficient transparency, full commitment, appropriate follow-up actions, and explicit financial backstop. A third goal for MpSTing is to contribute the process of setting macroprudential rules. More specifically stress tests could play additional input for calibration of capital buffer. As the topic is yet to be explored, applying such objective is not that common. There is also another goal that implicitly goes along with stress test conduct - to enhance market discipline (Goldstein, Sapra 2014 and Woo et al. 2014). Even though it is not a primary objective, disclosing bank information with a clear communication on results serves as a complementary to the three main goals by promoting transparency of the risks in banks.

For the next step of the process, the risk coverage is identified. It mainly comprises of financial, operational, and structural risks (Figure 3). Typically for a solvency stress test, credit risk is in the focus, as lending is the center of the traditional banking business. Market risk or the movements in interest rates, currencies, and asset prices (equities, debt securities, commodities, etc.) is also considered together with sovereign risk (haircuts on general government exposures). Additional risks come from higher costs of funding and operational events (including conduct and cyber risk[3]).

Figure 3. **Risk coverage**



Source: Kunghehian 2013, author contributions on adding sovereign and funding risk together with operational and structural risks.

[3] Currently, cyber risk is not part of stress tests but there are strong calls from EBA, ECB, and BoE for including it in the future (Financial Times, 2016).

The liquidity risk is based on the concept of Brunnemeier, Pedersen 2009, which comprises market illiquidity (securities are no longer liquid) and funding illiquidity (bank run). Counterparty structural risk is also important for financial stability that arises from interconnectedness (interbank contagion) or concentration (group of clients or a sector). On aggregate, credit risk and market risk are predominantly combined for solvency purposes with certain additions from sovereign, funding, operational, and concentration risk. As a second order complement, when a bank is insolvent, contagion risk could crystallize (Elsinger et al. 2006). Liquidity risk is normally tested individually but with the works of Puhr, Schmitz 2013, Pierret 2014, and Basel Committee on Banking Supervision (BCBS) 2015 research attention on the solvency-liquidity nexus is growing. Given the systemic importance of limited market access, fire sales and financial contagion, I share the view that an integrated approach is the way forward for incorporating all risks.

There are two alternative approaches in a MpST: top-down (TD) and bottom-up (BU). In the first option authorities run internally the exercise themselves by applying a uniform methodology on aggregate data. In the second, banks evaluate individually the impact on their balance sheets from a centrally pre-defined macroeconomic adversity in a more granular way. The selection of an alternative is central to the defined goal as both approaches have their strengths and weaknesses (Table 2). For instance, the TD approach is more flexible, relatively easy, and cost efficient. In my opinion it also allows for severity calibration, as well as incorporation of contagion analyses and macro feedbacks. By applying common methodology and assumptions, comparability across banks and consistency are ensured. On the downside, TD stress-testing may lack precision, transparency and applicability for banks with internal risk modeling. The BU option tends to be more tailor-made, as it incorporates bank idiosyncrasies, granular data, and risk models specifics. In addition, the involvement of bank management in the process provides better understanding of risk management practices and culture. As results are confirmed by banks, I would argue that a BU approach is more appropriate for individual disclosure and thus it is widely used for public exercises. However, such method comes with a great expense in terms of labor, time, and resources. Moreover, in order to ensure a certain degree of comparability, quality assurance (QA) process is necessary, which in times could be a lengthy and cumbersome procedure. Therefore in practice, a hybrid (constrained BU) option that combines both approaches is present as well[4]. The stress test is run by banks but under pre-defined constraints together with TD challenges and centralized QA. Accordingly, Jobst et

[4] The constrained BU approach is used in the EBA and BoE exercises, as well as in the Bulgarian National Bank (BNB) 2016 stress test.

al. 2013 advocates on behalf of IMF staff for the use of both approaches by considering them as complementary, rather than as substitutes.

Table 2. **Comparison of approaches for macroprudential stress-testing**

	Top-down (TD)	Bottom-up (BU)
Advantages	- Ensures objective comparability across banks by uniform methodology and consistent assumptions - Full aggregation of bank-by-bank results - Relatively easy to implement and resource-effective - Allows for incorporation of second round effects, interconnectedness and dynamic elements - Controlled framework and common environment - Calibration of severity with different scenarios and shocks	- Utilization of granular proprietary data on loan portfolios and exposures - Bank management engaged in the process - Banks' advanced internal models capture risk idiosyncrasies in business models and institutional differences - Confirmed results by bank - Appropriate and credible for individual publication - Widely used in public exercises - Refer banks to severity benchmarks and tail-events - Better understanding of banks' risk models, practices, and risk management culture
Disadvantages	- Low precision and fewer details due to data limitations - Application of simple models may produce overstandartization with inaccurate aggregates - Unsuitable for sample of banks with internal models - Lack of transparency and "black box" to banks and the public - Difficult to account for bank characteristics and institution specific information	- Contagion and second round effects are not accounted for - Requires extensive quality assurance process to challenge results - Labour-intensive and time-consuming - High-level of cooperation and coordination among participants - Possible influence on results by institution-specific assumptions - Conditional aggregation and comparison of results due to banks' risk models - Ex-ante predefined methodology

Source: IMF 2012, Jobst et al. 2013, author contributions on expanding the matrix in terms of usage, calibration, transparency, and management of the process.

The stress test scope follows the selected approach by targeting banks' coverage. Inclusion of all banks in the macroprudential exercise is usually done in a TD manner, while for a BU approach only a number of banks participate, especially the systemically important ones. A specific case is the BNB 2016 stress test, which has a BU approach with system-wide coverage. In terms of exposures, the appropriate level under stress corresponds to the risk materiality but should capture the on- and off-balance sheet dimension, as well as the domestic and foreign one. For credit risk, the focus is on the banking book. Loan exposures are generally aggregated as corporate lending, mortgages, and consumer lending. The trading book comprising of held for trading (HfT) and fair value through profit or loss (FVO) instruments is stressed for market and sovereign risk together with securities (incl. general government exposures) designated as available-for-sale (AfS) and held-to-maturity (HtM). Components of the liquidity stress test are liquid assets (cash, financial assets, bank placements) and funding liabilities from wholesale and retail (deposits) sources. For contagion analyses, a matrix based on interbank claims is typically employed. They can be defined on gross (interbank assets) or net (interbank exposure accounting the difference on the asset and liability side) basis. The effects from stress-testing concern the balance sheet, the profit and loss (P&L) statement, the liquidity position, risk weighted assets, regulatory capital, and pricing of financial instruments.

MpSTs can be generally based either on accounting or market data. The former choice incorporates on- and off-balance sheet positions, as well as supervisory and regulatory items (Čihák 2007 and Schmieder et al. 2011). Frequency of data depends on reporting and is usually at least on a quarterly basis. Balance sheet approach is convenient for designing a macroeconomic scenario, estimation of regulatory risk factors (non-performing loans (NPLs), loss rates), and provision of supervisory capital and liquidity outputs. The latter captures public information embedded in asset prices across financial instruments and bank equities (Acharya, Steffen 2014). Market data is obtained on daily basis and bank specific default frequencies are estimated through sensitivity analyses (for instance a decline in global stock market). The outcome of the market exercise comprises of expected losses and spillover probabilities among banks. IMF 2012 and Demekas 2015 provide extensive overviews of these assumptions and agree on both advantages and disadvantages present. Accounting-based stress test is widely used, as it allows for incorporation of various risks, macroeconomic link to banks' performance, and expert judgment assumptions, but market-based stress-testing steps on data that is forward-looking and easy to update. Weaknesses of the market data method, though, correspond to data availability and possible unrelated to fundamentals market swings. In support of market stress tests, Kapinos et al. 2015 point out that exclusive reliance on book equity may yield benign

results, while Acharya, Steffen 2016 criticize the regulatory approach of measuring capital adequacy (capital definition and risk weights) for being less accurate. In practice, however, for macroprudential and especially policymaking purposes, Constâncio 2016a advocates for the accounting based stress test as the more suitable for application. First, it is linked to a macroeconomic and financial narrative and second it is not subject to substantial procyclical volatility or "whims of the market". Additional limitations stem from the deepness of the market that puts challenges on comprehensiveness. In that manner, Jobst et al. 2013 conclude that the accounting approach remains the "cornerstone" of MpSTing.

The next common assumption posts the dilemma of static or dynamic balance sheet to the extent that accounting base is employed. Under the static balance sheet assumption there are several specifications. First, there is no credit growth and write-offs, as gross loans remain constant. Second, banks maintain the same business model (geographical exposure, currency composition, product mix), there are no management actions (e.g. restructuring, assets sale, one-off effects, risk mitigation techniques), and capital support is disallowed. Finally, maturing securities are replaced by identical instruments in terms of issuer, currency, balance sheet classification, and remaining maturity. Under the dynamic balance sheet assumption, on the other hand, balance positions evolve in line with the scenarios and management actions are allowed to a certain degree. With regards to the evolution of credit, the static approach keeps the total amount of gross loans unchanged, but the quality composition worsens as NPLs increase on the account of performing ones. In the dynamic stress-testing, gross loans are extended with applied credit growth and subsequently for the new lending the general practice is to remain performing, i.e. no credit shocks. The static balance assumption eases the life of stress-testers, as it is simple and straight-forward. It enhances comparability across banks and facilitates communication of results, but there are flaws as well. The stationary balance is regarded as additional constraint and being unrealistic, because banks never stay static. Accordingly, dynamic stress-testing is gaining momentum. Henry, Kok 2013 describe the ECB framework for projecting balance sheet elements, while BCBS 2015 provides an overview of methodologies that incorporate solvency and liquidity dynamics. A general practical rule associates the static balance sheet assumption to a BU stress test and the dynamic elements to a TD stress test, but the vice-versa option is present as well.

Moreover, the horizon of stress is also relevant. As it takes a longer time period for effects to unfold, solvency stress-testing assumes typically at least one year[5]. Liquidity stress tests are shorter-term - between 5 and 30 days, while for market and contagion risk, the shock is instantaneous. When setting the

[5] The stress period in the US, EU, and UK (standard) exercise is 9 quarters, 3 years, and 5 years, respectively.

outcome indicator, the common practice is to follow regulatory measures (risk based capital ratios[6], leverage, liquidity ratios). As far as the corresponding hurdle rates (pass or fail threshold) are concerned there are various experiences. One practice is to set the bar in terms of minimum regulatory requirements (financial stability perspective) or a combination with capital buffers (BoE 2015). The reference levels could also be determined by expert judgment. An alternative approach is to apply no hurdle rate (EBA 2016 and BNB 2016a). The stress test results accordingly serve as an input to the supervisory assessment under the Supervisory Review and Evaluation Process (SREP). For recapitalization purposes (crisis stress test), Ong, Pazarbasioglu 2013 highlight the importance of committing to a pre-defined pass/fail hurdle rate for achieving credible capital shortfall estimates. They in turn affect subsequent capital raising or potential fiscal support. The cut-off period for data is selected as of year-end prior to the stress test on either individual (unconsolidated) or consolidated basis. Semi-annual reference is also possible but less popular. Certain limitations could be assumed on P&L items, RWAs, capital and pricing of instruments. For instance, caps are imposed on dividend income, net fee and commission income, net trading income, and capital while interest expenses and RWAs could be floored by the starting point. One major constraint is the restriction of accruing interest income on NPLs or defaulted assets under adversity, which differs from the usual going concern accounting principles. Another assumption does not allow for optimization of administrative expenses during stress or inclusion of one-off adjustments. In case of assumed funding shock, banks may be restricted to fully pass forward the higher cost of financing to borrowers. Finally, rules for dividend distributions could be set – allowed payout, zero payout, or both.

2.2. Application

The application part of MpSTing begins with the scenario design[7]. In its simplest form sensitivity analysis evaluates the impact of an instantaneous shock of one or a set of risk drivers. Typical examples of sensitivity stress reflect parallel interest rate shifts, currency devaluations, rise of NPLs, and many others. Although it is simple and intuitive, it has no economic justification and therefore has limited suitability. On the other hand, implementing a macroeconomic scenario is more complex, however, the effects are quantified in a comprehensive and continuous manner. Usually in a MpST, a single scenario is constructed in two

[6] A common used outcome indicator is the CET 1 capital ratio. It equates to the division of common equity tier 1 (CET 1) capital as a numerator and risk exposures (RWAs) as a denominator.

[7] Scenarios are designed mainly for solvency stress tests. For liquidity stress-testing, bank runs and haircuts on realized assets are assumed in the short-term.

aspects - baseline and adverse. The baseline scenario mostly follows the current official central bank macroeconomic forecast but alternative sources such as the IMF WEO outlook, EC economic forecast, ECB macroeconomic projections are also possible. The adverse scenario represents hypothetical but plausible negative (extreme) macroeconomic developments. It is based on historical or newly constructed episodes. The techniques for designing an adverse scenario fall into three main categories – model-consistent (Foglia 2009 and Licari, Suárez-Lledó 2013), judgmental (Melecky, Podpriera 2010), and combination of both (Hristev 2014). The first type consists of dynamic stochastic general equilibrium (DSGE), vector autoregressive (VAR), and structural econometric models, as well as statistical approaches like copulas or standard deviations. The second one employs solely expert judgment and is used when modeling is not available or not capable of generating appropriate scenarios. The third category combines a macroeconomic model and expert judgment for fine-tuning. In practice, it is difficult to disentangle a specific technique, as more than one is typically integrated. For instance, the adverse scenario for EBA 2016 EU-wide stress test combines predominantly models[8], multi-country EU-wide stress test elasticities (STEs)[9], and expert judgment (European Systemic Risk Board 2016), while the BNB 2016 stress test utilizes central bank macroeconomic econometric model, EU-wide STEs, statistical approach (standard deviations), and national expertise (BNB 2016b). Regardless of designing techniques, the selected option should be wrapped in a convincing narrative. The story behind the numbers starts with identification of main sources of risk to the economy and subsequently to the banking sector, which materialize as economic (foreign and domestic demand, investment, consumption) and financial (interest rates, volatility, depreciation of local currency, house prices) shocks[10]. Typical macroeconomic variables in scope are real GDP growth rate, inflation (HICP), unemployment rate, short- and long-term interest rates, FX rates, and asset prices (equities, bonds, residential and commercial property prices). Key variables may follow a W-shaped (double dip), V-shaped (fast recovery), U-shaped (slow recovery), or L-shaped (stagnant recovery) recession. Presentation of scenarios is usually against historical time series, crisis episodes, or other scenarios in different stress tests.

For producing a credible and relevant MpST, the level of severity is critical. Breuer et al. 2009 raise the question on credibility that concerns the link between severity and plausibility. The authors provide a sound statistical solution for find-

[8] Among others, there are DSGE models, Bayesian and Global VAR models, and large scale global econometric (NiGEM) models.

[9] The STEs are impulse responses based on forecasting national models in the central banks of the ESCB and integrated by the ECB as an instrument for scenario-generating processes.

[10] In a liquidity stress test, a bank run crisis is designed that incorporates withdrawal shocks and market haircuts.

ing the worst case scenario, which in their view enhances objectivity and makes the "disaster myopia" obsolete. In my opinion, however, it is difficult to link such method with a consistent narrative that captures the "unthinkable" and the next "black swan". Accordingly, a stress-tester may alternatively opt for a combination of scenarios or reverse stress-testing. The multiple scenarios topic attracts supporters, as being a good practice and common sense (Dowd 2015), flexible for exploring new or emerging threats (Dent et al. 2016), and a backstop for banks to "game the test" (Demekas 2015). By using multiple in terms of number and type scenarios, the assessment of bank resilience is enriched with a variety of shocks and the coverage of reledet risks is ensured. Alongside the advantages, there are pitfalls as well. With multiple scenarios, the resource burden is stronger for both banks and regulatory authorities, especially when conducting BU stress tests. Moreover, difficulties with summarizing and publishing stress test outcomes pose a challenge on communication of results. Taking into account both sides, the following question arises: how multiple is optimal (plausible)? I share the pragmatic view of Quagliariello 2017 that at the moment multiple scenarios have limited practical application in a system-wide stress test, but "specific add-ons and sensitivities can add to the severity of the scenario". In a different paradigm, reverse stress-testing is also useful for identifying severe paths in the scenario (Bookstaber et al. 2014). The assessment of tail risks that most likely render a bank unviable reveals additional inputs to the traditional scenario design. Bologna, Segura 2016 bring up the question of scenario relevance regarding severity and cyclicality. Their research calls for procyclical severity generating countercyclical regulatory capital requirements, which in the light of the macroprudential principles I find warranted[11]. It means that during economic expansions, severity should be increased in order to attain higher capital requirements and vice versa during recessions. The proposed mechanism exhibits macroprudential features and helps the setup of Basel III capital buffers along the cycle. In practice, the Bank of England's approach to stress-testing accounts for the issues on severity (BoE 2015). The BoE stress test framework incorporates two types of adverse scenarios – annual cyclical scenario (ACS) and biennial exploratory scenario (BES)[12]. The first one is a standardized macroeconomic scenario with a 5-year horizon and is associated with the financial cycle. Its severity increases as risks buildup and decreases as they materialize or fade away, but the precise calibration integrates judgment from policymakers. The role of ACS is to put focus on capital adequacy and contribute the setting of capital buffers. The second type of scenario explores a wider and mostly unusual from historical perspective set of risks and their

[11] Constâncio 2016b and Brazier 2017 advocate for strong countercyclicality as a fundamental principle when conducting macroprudential policy.

[12] Both scenarios overlap for the first time in the 2017 BoE stress test exercise.

evolution over the period of 7 years. Unlike the ACS, the BES attempts to sketch the future outlook of the financial system by assessing banks' business models prudence, strategies, and resilience. As the subject of scenario severity is yet to be practically and academically settled, expert judgment remains crucial for scenario fine-tuning and calibration.

Risk modeling that translates scenarios into quantifiable effects for banks' balance sheet is an integral part of the MpST. There is a vast universe of models, which could be categorized by risk type. Henry and Kok 2013 provide a detailed overview of various models for different risks, while Foglia 2009 and BCBS 2013 focus exclusively on credit and liquidity risk models, respectively.

For credit risk, the link with the macroeconomic variables has been widely constructed on credit-quality regression (satellite) models. They can be grouped into two main approaches. The first approach employs indicators for credit quality drawn from banks' balance sheet on the left-hand side of the equation and on the right-hand side - macroeconomic indicators coupled with specific for the banking sector factors. A typical dependent variable is the NPL amount or ratio based on supervisory data[13]. When NPL modeling is considered, corresponding credit shocks are generated from the macro scenario and then are applied to the non-working side of the loan portfolio. NPL models could be further specified for corporate and retail credits, thus accounting for their respective propensity to changes in output, interest rates, unemployment, and other macro variables. Finally, the effect for the P&L stems from the impairment costs, which are calculated as the product of the newly formed NPL flows and a provisioning rate under stress. Alternative mechanism is to model directly the loan loss reserves (stock of accumulated impairments) or impairment costs (flow of impairments).

The second approach utilizes the risk parameters from the Basel frameworks. It is currently the most prominent modeling of credit risk and follows the formula:

$$EL = EAD * PD * LGD \qquad (1)$$

Where:

EL is the expected loss;

EAD is the exposure at default, which represents the outstanding amount if the borrower defaults;

PD is the probability of default of a counterparty over a one year period;

LGD is the loss given default, which represents the ratio of the loss on an exposure due to the default of a counterparty to the amount outstanding at default, or the non-recovered part.

[13] Other models could employ different options like write-offs (monetary statistics) or defaults (credit registers).

The *PD* estimation is mainly performed via logistic regressions, as the output is bounded between 0 and 1. Selection of the dependent variable follows a logistic transformation of a bankruptcy rate or a similar deterioration of quality (e.g. loans past due 90 days). The independent variables are of macroeconomic and obligor-specific nature. For the *LGD* factor, the calculation encompasses information on borrower's payments and collateral value. Contrary to the previous approach, *PD* and *LGD* parameters are applied on the performing loans in order to obtain losses. A further modeling consideration expressed by Buncic, Melecky 2012 is the choice of implementing the factors through-the-cycle (TTC) or point-in-time (PIT). TTC risk factors are estimated over the whole cycle and their conservativeness depends on the downturn deepness. If the cycle includes a number of crises, then additional credit risk shocks should be applied with caution. PIT risk factors reflect the current state of business. Usually, they are stressed with higher adversity, but the estimation may pose volatility challenges (or over-conservatism) due to unrelated to macroeconomic conditions or one-off factors. A prudent rule to capture current risks is to apply PIT parameters on newly defaulted exposures (EBA 2016), while for already defaulted exposures a TTC approach is also suitable (BNB 2016c). Additional calibration stems from the starting point of risk parameters, i.e. lower starting points are associated with higher credit shock, and vice versa.

The market models map the stress in the trading book positions to projections of net gains and losses. As the models require high-frequency data, usually they are implemented in a BU stress test with a given set of market risk parameters (stock prices, credit spreads, swap rates, etc.). A TD option is a more simplified version that links the net trading income to its historical volatility. For sovereign exposures, haircuts are derived from the price deviation, which is typically based on discounted cash flow method incorporating a shock to the yield curve. In the profitability models, satellite calculations are translated into the evolution of net interest income, non-interest items, and other P&L elements. The liquidity risk models typically comprise the market and funding part of the shock in a three-step process (van den End 2008). First, a liquidity shortfall is envisaged, which reflects the lack of liquid assets to absorb assumed withdrawals. Second, banks' reaction is modeled via fire sales and third the corresponding feedback effects amplify the initial shock. Contagion models assess the spillover effects from a failing bank. They are rather static and mechanical because are often based on balance interbank claims. Different approaches include a "domino effect" analysis (Espinosa-Vega, Solé 2010), clearing mechanism (Eisenberg, Noe 2001), and post-shock network equilibrium (Gouriéroux et al. 2012).

An important element of the stress test application, especially for BU exercises, is the quality assurance (QA) process and validation of outcomes. The goal

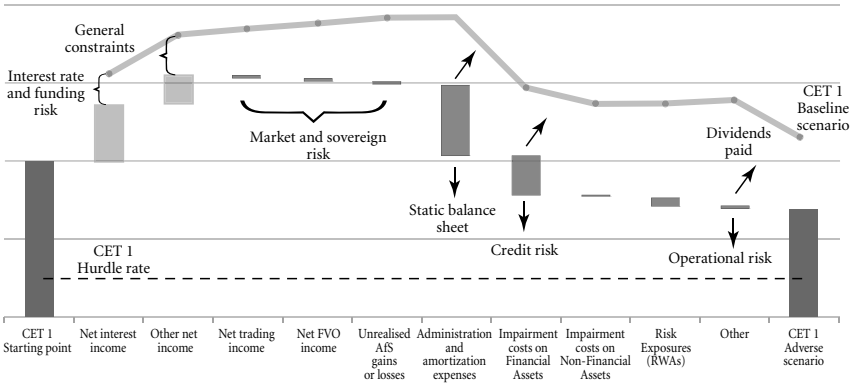
of the QA is to check compliance with the prescribed guidelines, appropriate level of conservativeness, and adherence to the technical constraints. Robust supervisory challenging of results increases relevance and comparability. It also ensures the level playing field, so that more conservative banks and those with less prudent methods are treated equally. For that purpose, stress test outcomes are compared in a TD manner with cross-section benchmarks, historical time-series, and general rules of thumb. Practical guidance on the QA process is provided in the ECB 2014 note. The ECB approach represents a score card that has a Red/Amber/Green (RAG) attributes. Green color corresponds to full compliance and no modification, amber color refers to "comply or explain", and red color stands for non-compliance and further resubmission of results. Traditional areas of examination are input data, credit risk parameters, market and funding shocks, one-off events, inconsistencies with assumptions, sufficient volume of qualitative explanations, etc. TD tools supplement the QA with independent stress test quantification that challenges consistency of results. For additional analyses, qualitative information is employed, as well. As one MpST involves a vast amount of quantitative and qualitative information, the QA process follows the principle of materiality and accordingly focuses on issues that significantly alter the stress test outcome. Moreover, validating results is essential for credibility. By back-testing the models, robustness of shocks is enhanced by not overshooting or underestimating them. Finally, expert judgment is applied for the QA and validation process in order to detect inconsistencies and manage outliers.

2.3. Presentation of results

The MpST ends with a report of the main outcomes. Stress test results can be summarized on a system or bank-by-bank level. Further options are by peer groups or statistical distributions. The evolution of capital or liquidity outcomes is provided for each period of the scenario, as well as the number of banks that fail to meet the pre-defined hurdle rate (if applicable). In that manner, absolute and relative to GDP shortages are given together with a percentage of total system's assets under consideration. The stress test is by design an analytical tool, so even if all banks are above the minimum threshold ("pass the stress test") that does not mean there are no vulnerabilities. Assessment of banks' sensitivities to shocks, propagation channels, and risk amplifications provides valuable insights to policymakers. For example in a solvency framework, the stress test reveals credit risk evolution, market risk sensitivity, capacity of net interest earnings, cost efficiency and sustainability of bank's business model. Moreover, in a liquidity stress test, the liquidity position is analyzed against different patterns of depositors' behavior and corresponding asset management strategies, while in the contagion exercise the

degree of interconnectedness and systemic importance is revealed. Accordingly, stress test results are supplemented with detailed information about credit risk parameters, NPLs and provisioning, liquid assets, P&L items, pricing of loans and deposits, etc. Technical notes describing the scenario design, methodology, models, qualitative summaries, assumptions and limitations also accompany the report for better clarification.

Figure 4. **Stylized contribution of main drivers to the change of CET 1 capital ratio**



Notes: Impact of specified bars is generally colored and it could vary from one bank case to another.

Other net income mainly comprises net income from fees and commissions and net other operating income. Net FVO income is generated from financial assets and liabilities designated at fair value through profit or loss. Unrealized AfS gains or losses correspond to available-for-sale financial assets and liabilities. Impairment costs on financail assets are typically from loans and advances and held-to-maturity investments. Other category consists of operational risk losses, paid dividends and taxes, and capital adjustments.

The figure has an illustrative purpose and does not represent a specific numerical expression of bank results. All impacts are of exemplary nature.

Additional granularity on the impact of different risk factors is included in the report to appropriately convey the stress test findings. In the solvency stress test the contribution of each driver to the change of CET 1 capital ratio is shown via the widely-used format of a waterfall chart (Figure 4). Under the baseline scenario and static balance sheet assumption, net operating income remains intact with additional gains from market instruments. Fixed costs like administration expenses reflect the biggest financial burden for a bank to support its business, whereas credit impairments affect the P&L moderately. End-

capital position depends on the assumed dividend policy - it improves when profits are retained and declines when dividend pay-out exceeds the current profit with reserves from previous financial years. Under the adverse scenario and static balance sheet assumption, net interest income shrinks from interest rate and funding shocks or could even become negative. Next, the economic activity slowdown subdues the net income from fees and commissions, while imposed general constraints like one-offs restrictions affect the net other operating income. Market and sovereign shocks are typically negative, but the net trading income could be positive, however with less impact than the one in the baseline scenario. The static balance sheet assumption keeps the burden of administrative expenses intact[14], while loan losses surpass the baseline figures by a factor of two, three, or more. Impairments on non-financial assets mainly related to lower real estate prices contribute to the CET 1 capital ratio reduction. Final shocks to solvency come from higher risk exposures and operational losses reflecting the deterioration of credit quality and processes (including conduct), respectively. Usually, in the end of the adverse period, CET 1 ratio is lower than the starting point or it could even become negative, i.e. the bank is technically insolvent. In the liquidity stress test, the impact of market illiquidity and funding illiquidity risk drivers is presented for every period via the available liquidity buffers and the cumulative deposit withdrawals. In a contagion simulation, the effect on the interbank network reflects the inflicted capital losses from the insolvency of each bank.

If mandated, a stress test could be publicly disclosed. The decision to communicate the exercise in public together with its granularity and scope relies entirely on the authority's approach and adopted goals. Practices of MpSTs disclosure vary substantially across jurisdictions. Some authorities opt for broad transparency with scenarios, methodologies, and bank-by-bank results, while others limit the publication to aggregate results. There is also the case when stress test outputs are kept private for internal purposes. In the US for example, publication of aggregate and individual results is required by law[15], while in the EU-wide context the EBA founding regulation compels the supervisory authority to disclose bank-specific results if "appropriate to do so"[16]. The UK experience on transparency follows the official stress-testing framework of the Bank of England (BoE 2013, 2015), in which the central bank commits to disclosing its approach, results (both aggregate and individual), and policy actions. Other instances like the IMF FSAPs and com-

[14] Under the dynamic balance sheet assumption (contingency plans) administrative expenses could be limited as banks may respond to a crisis with cost-cutting strategies like lay-offs, business scale downs, remuneration adjustments, etc.

[15] See Sec. 165i (1) and (2) of the Dodd-Franc Act.

[16] See Art. 22 (1a) of the Regulation (EU) No. 1093/2010.

prehensive assessments have unique degree of transparency as each exercise differs from one another. The former always has aggregate results at minimum (Jobst et al. 2013), whereas the latter is traditionally accompanied with bank-by-bank disclosure (ECB practice). From theoretical point of view, a vast body of researchers advocate for publication of results to promote financial stability (Bookstaber et al. 2014, Goldstein, Sapra 2014, and Woo et al. 2014 among others). In their view disclosure of stress test outcomes produces valuable information to the market, reduces banks' opaqueness, strengthens the supervisory position for demanding more capital, improves bank management through market discipline, and induces further voluntary transparency. Empirical papers employing event study and bank performance methods confirm in general the theoretical benefits of stress test transparency. In the US, publication of results generates significant new information to the market, which reduces asymmetry and mitigates bank opacity (Fernandez et al. 2015 and Flannery et al. 2017). The result is affirmed as well for European banks by the market reactions to the EBA 2011 stress test (Petrella, Resti 2013). Candelon, Sy 2015 compare the market reaction to U.S. and EU-wide stress tests in the period of 2009 – 2013 and observe that the qualitative aspects of the stress test governance such as strong institutional framework and credible financial backstop seem to matter more for market participants. Along with the positives of disclosure, stress test publications are also associated with potential costs. Hirtle, Lehnert 2014 alert that there is also the risk of investor panic (self-fulfilling prophecies) or bank managers focusing solely on short-term results (beauty contest). When publishing results, accompanying communication strategy is also key (Enria 2017). It is needed to make sure that stress test nature is understood properly and results are interpreted with caution. While plausible, the MpST is a hypothetical exercise under specific assumptions with outputs that are not perceived by the authority as likely to happen. Moreover, stress test results are intend to illustrate potential vulnerabilities in a quantified manner and are not considered to represent a prediction about the financial stability of the banking sector or an individual bank. Finally, the public should be informed how the stress test results affect the supervisory decision-making process. Taking the costs and benefits together, there is a positive information value of disclosing stress test results, but any decision to go public has to be considered carefully and communicated clearly.

2.4. Follow-up and feedback

The main stress test outcomes feed into the supervisory assessment of banks. They provide for another source of evaluating banks' capital and liquidity position, reviewing their business plan and strategy, and examining the recovery and contingency plans. For solvency purposes, stress test results contribute to determining the adequate level of capital. If deemed necessary a reaction function is possible. Based on the results, supervisors could prompt direct recapitalization or as part of the SREP, they could prescribe capital requirements (P2R) or capital guidance (P2G) under Pillar 2. Other supervisory measures could restrict dividend payout, staff remuneration, and share buybacks. In turn, banks may have to respond with raising capital, adjusting their business model, selling business lines, or restructuring. A fiscal backstop serves as an option from last-resort. Other follow-ups to banks could target their risk governance processes, which fosters further development of efficient and adequate risk management tools. Finally, stress test results are incorporated in the microprudential (e.g. on-site inspections) and macroprudential (e.g. setting capital buffers) supervisory toolkit.

Policy and bank actions lead to behavioral and dynamic feedbacks that could if applicable further aggravate the initial macroeconomic scenario (second-round effects). These macro-financial linkages are modeled in a TD manner (Henry, Kok 2013). A major response channel is through adjustments on the asset side. For example, one way to address a stress test shortfall is to contract credit and tighten credit standards. The corresponding feedback causes limitations in credit supply and thus creates a further shock to output. Another reaction comes from industry deleveraging, which could induce firms to constrain production and spur lay-offs. Finally, higher interest rates contract consumption and may cause additional credit quality deterioration. Through the liability side, bank contagion (Lehman episode) could also surface as a second round effect.

3. Challenges

Evidently MpSTs are here to stay. The topic attracts attention from various sides – academics, bankers, supervisors, policy makers, market analysts and the wide public. The growing body of research provides a far better understanding about the theory and practice of stress-testing but there are still a few questions that remain a challenge.

To publish or not to publish stress test results

As publication of stress test results goes along with costs and benefits, the debate on transparency continues. To publish or not to publish stress test

results? And if so, to what extent? The decision to go public with a stress test firstly corresponds to its mandate, which is usually embedded in the exercise beforehand. Then, it should also consider the underlying goal and approach (EBA 2015). If disclosure is not specifically mandated, authorities use their own discretion to publish results or not when they deem appropriate[17]. Overall there are three ways of approaching transparency – stress test results remain private (no publication), stress test results are disclosed only at aggregate level (partial publication), or stress test results are disclosed at both aggregate and bank-by-bank level (granular publication). Schuermann 2016 suggests that in "peacetime" (supervisory stress test) less detailed results are necessary, while in "wartime" (crisis management) high transparency and individual disclosure is required. Goldstein, Leitner 2015 propose similar disclosure policy. In normal times and banks are perceived as strong, no disclosure is needed, but if banks are viewed as undercapitalized partial disclosure emerges as optimal. In bad times, if results are not known to the bank before the stress test, a distribution of strong and weak banks could be presented. However, if the results are known to the bank beforehand (but not the market), optimal disclosure favors multiple categorization of banks, i.e. publishing individual results.

In practice, the general approach is indeed to publish aggregate TD results as part of the regular monitoring in a financial stability report (central bank stress tests and IMF FSAPs). The system-wide exercises in US, EU, and the UK began with publication of bank data to restore confidence in the banking sector as part of the response to the financial crisis and since then sustained the level of public granularity. Moreover, comprehensive assessments in Europe were finalized with disclosure of individual findings. A distinct feature of public exercises, though, is that they are conducted in a BU manner and the lesson learned is to ensure a credible backstop if individual results are disclosed. Taking into account the academic practical side of the matter, my recommendation is towards appropriate transparency. For sophisticated exercises in terms of scope and governance magnitude, publication of results should suffice market expectations. In that manner, bank information is disclosed together with detailed notes on the overall stress test design. For regular supervisory exercises, on the other hand, individual data may not be published, but there are positives for presenting results of the banking system. Publication of aggregate outcomes enhances the relevance of the risk analysis report and serves as an instrument to draw public awareness on analytical tools for bank assessment. It also increases transparency and is in line with the current IMF FSAP practices. Disclosing of minimum banking system results should be the pursued benchmark by supervi-

[17] In the EU, competent authorities are compelled to run at least annually supervisory stress tests in their national jurisdictions, but there is no standardized requirement for their transparency (CRD IV art. 100).

sors and presented from their discretion in a form deemed most recognizable by the general public.

In the end, the choice of publication and corresponding granularity relies solely on the banking authority taking into account the sophistication of financial intermediation, market development, and environment expectations. Given the considerations of transparency, credibility, and backstop (fiscal) availability, the decision for disclosure is far from being straightforward and still poses a challenge.

Top-down vs bottom-up conundrum

In line with the previous discussion, there are advantages and disadvantages for the TD and BU approach, but the final choice comes down to the decision of disclosing individual results. When the exercise is conducted in a public fashion, it would be unusual to disclose bank data without banks participation. Accordingly, credibility of results must be preserved, which is trade-off with a certain price. Thus, for a public stress test the constrained BU approach is the first step and QA – the second. Due to the resource intensive process, an optimal practice for authorities is to conduct a public BU stress test when deemed necessary[18]. As for the TD approach, there is no bound to the number of in-house simulations, but the common practice is to have at least an annual exercise with disclosure of aggregate outcomes.

One key feature in the TD and BU conundrum is the requirement of resources. In a setting with constrained human capital and compressed time-schedule, an annual BU exercise might seem redundant and overly burdensome to the banking industry. Also, as structural changes appear in longer periods of time, utilization of granular data could be sub-optimal as results do not yield much added-value to the TD approach. If dynamic elements or subsequent contagion effects are sought as important, a TD stress test is currently the way forward. Therefore, when resources are available both approaches could be applied on annual basis, but if not I would suggest the EBA practice of conducting a BU stress test every two years or in a year of implementing new regulatory or accounting frameworks. The recommendation could be applied as well to traditional banking systems, for which the external bank assessment by the market or by other analysts does not depend on regular results from public stress-testing.

[18] The frequency of BU exercises is set annually in the stress test frameworks of the FED (31 participants in 2016) and the BoE (7 in 2016), biennially for the EBA EU-wide stress test (51 in 2016), and ad-hoc for the comprehensive assessments in Ireland 2011 (4), Greece 2013 (4), Spain 2012 (14), Slovenia 2013 (8), Bulgaria 2016 (22), and ECB 2014 (130), 2015 (9), 2016 (4).

Static vs dynamic balance sheet assumption and the inclusion of second round effects

The balance sheet dilemma continues to be a struggle. On one hand, the static assumption ensures comparability especially across diversified sample of banks and is straightforward to implement. The corresponding convenience attracts many practitioners, as well as all EBA stress tests. On the other, including dynamic elements makes the exercise more realistic and its outcome more reliable for capital planning (US CCAR) and recapitalization needs (Greece 2013). The UK approach emphasizes on the banks' financial intermediation function and mandates them to sustain lending under stress (BoE 2015). Demekas 2015 also discusses the dynamic practice as growing worldwide in a TD way. Other country stress tests examples show some degree of dynamism in terms of balance sheet and credit growth (Jobst et al. 2013 and Hristev 2014). On the downside, BCBS 2015 points out to the present need of a satisfactory approach to dynamic modeling given the complex macro-financial linkages that banks are part of. Accordingly, both assumptions are currently applicable.

The dynamic balance sheet is indeed superior in theory. It allows for inclusion of macroprudential aspects in the setup and second round effects such as macroeconomic feedbacks or bank contagion. Nevertheless, in my view due to present significant comparability and modeling challenges, the use of dynamic stress tests for supervisory purposes should correspond to the goal and the approach of the exercise. It is also subject to the necessary resources for ensuring consistency and comparability across banks. Although being more realistic, projecting the balance in a dynamic way brings complexity to a degree that is directly dependent to the sophistication of banking intermediation and business models. Thus, based on the available analytical resources and the level of banks' development as a system, I would advocate for both alternatives of the assumption underlining the dynamic side as the one to move forward.

Application of stress tests for MPIs calibration and systemic risk analyses

There are clear calls for inclusion of stress test results into the calibration process of MPIs (Brazier 2015, Constâncio 2016a, Tarullo 2016) and the analyses of systemic risk (Demekas 2015). At present, few evidences attribute the use of MpSTs for setting capital buffers[19], while for other MPIs information is scarce. A recent relevant step forward to the pure macroprudential aspects of stress-testing is the ECB STAMP€ (Stress Test Analytics for Macroprudential Purposes in the euro area) instrument (Dees et al. 2017). It represents a TD tool

[19] For example, in the US and the UK, stress test results are incorporated in the setting of countercyclical buffer. The practice of the Czech National Bank and the Central Bank of Estonia is to consider simulation findings when determining the level of systemic risk buffer and the buffer for other systemically important institutions.

for supporting design and calibration of macroprudential policy. The analytical framework contributes to the macroprudential purposes of stress-testing by incorporating a dynamic dimension, interaction with the real economy, interconnections between financial institutions, system-wide liquidity assessment, and interaction with non-financial sectors. As a result, the enhanced ECB TD stress-testing allows for impact assessment of macroprudential policy instruments. Contribution of banks to systemic risk assessed via networks analyses or agent-based models is the new frontier in stress-testing. These methods are viewed as better suited for analyzing systemic resilience due to the option of including specific agent dynamics that are not linked with past experiences, assessing their responses to events, and updating all variables in the system accordingly. The computational complexity and technical requirements still pose a practical hurdle to agent-based modeling, but the opportunity to stress test the next "black swan" is an incentive worth for further investigation.

In my view, there is avenue for exploring the accommodation of stress test results in the macroprudential setup. The simulation provides comprehensive analysis about banks' capital and liquidity resilience together with their endurance to contagion. By identifying specific quantitative shortages and general pockets of vulnerability, the MpST could provide additional information to the process of implementing capital instruments. For example, the former may contribute as a supplement indicator to the process of setting a buffer for systemically important institutions, while the latter – for systemic risk buffer. In addition, liquidity rules could be imposed to enrich the current regulatory framework with more conservative stress test assumptions. As MpSTs are hypothetical exercises they could be a valuable source of information for macroprudential policy, but not the solely one. Accordingly, supervisory expertise is important to utilize findings from stress-testing into the application of MPIs.

Marrying the solvency and liquidity stress test frameworks

The global financial crisis of 2007/8 revealed that solvency and liquidity shocks interact together. Accordingly for stress-testing purposes, integrating the solvency and liquidity nexus into one framework gives a more complete view of banks' resilience (BCBS 2015). If taken separately, important interconnections may be missed. To illustrate, signs of capital depletion may prompt funding shocks, fire sales, or run-offs, and vice-versa. Several instances attempt to model solvency and liquidity linkages For example, the Bank of England's Risk Assessment Model for Systemic Institutions (RAMSI) includes shocks to funding liquidity, while Bank of Canada's MacroFinancial Risk Assessment Framework (MFRAF) assumes fire sales and possible runs on top of the macroeconomic scenario[20]. In a similar

[20] For more information about the RAMSI and MFRAF model see Burrows et al. 2012 and Anand et al. 2014, respectively.

fashion, the systemic modules of the ECB's STAMP€ tool bridge the solvency and liquidity link by modeling explicitly the lender of last resort function, fire sales, and funding contagion. The corresponding accumulated losses affect banks' solvency position and consequently trigger higher wholesale funding spreads not only for banks with capital deterioration, but also for those with similar business models. A particular feature of an integrated solvency and liquidity framework is mirrored in its TD feasibility, thus making it not applicable for BU exercises. Moreover, modeling and calibration of a liquidity crisis is challenging due to its "low frequency-high impact" nature. As a result, the majority of regulatory stress tests still assess solvency and liquidity separately[21].

With regards to the importance of the solvency and liquidity nexus, I would suggest that future work on the matter is essential. The link between solvency and liquidity emerges in different channels. One effect reflects higher funding costs as lenders adjust the risk premium for counterparty risk when bank capital depletes. Other channels result in limited market access, fire sale losses, and liquidity run-offs. Finally, the solvency and liquidity issues could turn into interbank market breakdown and cause contagion across the financial system. The solvency and liquidity shocks are mutually amplifying and so an integrated approach for marrying both frameworks is a pressing need.

Limitations, imperfections, and lessons learned

All models are inherently fallible, as Haldane 2009 puts it, because they are not reality. And MpSTs are not an exception. As such they are carried out with certain limitations and imperfections. The overall design of a stress test is coated by expert judgment, which brings a certain degree of subjectivity. Every exercise has predefined assumptions that go beyond the risk modeling and affect the whole process. The stress test is also as robust and credible as the underlying data[22], methodology, and governance. Indeed, significant progress has been made to improve the stress-testing frameworks, data collection, and risk management, but a notch of uncertainty always remains. Further, the stress test happens with some lag, which could have implications on specific risks materialization. The rapid unwinding of financial adversities casts doubt on the scenarios as relevant risks may not be covered (2011 sovereign debt crisis or 2016 UK referendum). Another limitation stems from using historical series and standardized frameworks. As financial intermediation and corresponding risks evolve, stress-testing could miss capturing new risks and create false sense of security. That said, MpSTs are continuously in development to address the

[21] In the US, the annual CCAR goes along with CLAR (comprehensive liquidity analysis and review). Similarly, the 2011 financial measures programme in Ireland had prudential capital assessment review (PCAR) and prudential liquidity assessment review (PLAR) as stand-alone analyses.

[22] Asset quality reviews (AQR), if available, are beneficial for stress test credibility, as they ensure accurate starting data and adequate reporting of risks.

present shortcomings and at the same time attempt to follow Keynes' humble view of being "roughly right than precisely wrong". They are simply an analytical tool that does not substitute a robust regulatory framework, but rather complements the microprudential and macroprudential supervisory arsenal (capital and liquidity instruments, regular risk assessments, financial soundness indicators (FSIs), early warning systems (heat maps), network analyses, systemic risk models, etc.). Accordingly, stress-testing enriches the assessment of banks' health and adds valuable input to the process of implementing policy actions.

Along with the MpSTs' insufficiencies, there are positives related to lessons learned from past experience. One main lesson is that the overall stress test process should follow the set goal. For public exercises such as bank comprehensive assessments or crisis stress tests, credibility must be ensured, which corresponds to adequate severity, sufficient transparency on bank level, and a clear communication. The stress test thus implies a BU approach, a relevant adverse scenario and risk coverage, a QA check, and a fiscal backstop. In order to achieve these prerequisites significant resources should be allocated both in terms of human and financial capital. For the purposes of supervisory monitoring, the stress test is more flexible, which gives an opportunity for including dynamic elements, second-round effects, and solvency-liquidity interactions. The TD approach is currently the way forward to accommodate these complexities and expand the supervisory view of bank reactions under stress. Another lesson follows the use of supervisory data as the central input for MpSTing. It fosters comparability across the sample and is available for both large systemically important and small banks. Regardless of the presented shortcomings, the use of balance-sheet and regulatory capital information is linked to macroeconomic and financial rationale, which promotes the basis for a meaningful supervisory dialogue with banks. If market data is present and the market is well-developed itself, it could be included as an additional source of analysis, but I would recommend keeping supervisory data as the main input for the MpST. A third lesson accounts for the level of severity to the applied stress test framework. The design of scenarios, risk modeling, and assumptions should account for the economic and financial cycle, as well as for the current conservativeness of implemented macroprudential instruments. The final lesson brings the necessity of expert judgment. National expertise utilizes supervisory and economic knowledge that could be reflected in expert add-ons if needed for calibrating the overall severity or managing outliers. The right balance between adverse shocks (tail events) and observed reality (common sense) brings the stress test one step closer to the golden rule of "severe enough, but plausible".

MpSTs are imperfect, yes, they face challenges and have limitations. After all, to quote Charles Goodhart: "fallible as they may be, the conduct of annual

stress tests gives the regulatory authorities their best available chance of dealing with fragile banks while there is still enough time to avert a, potentially contagious, failure."[23] Therefore, as long as the illusion of impeccable precision is avoided, the inherent weaknesses of stress-testing are acknowledged, and the lessons from past experience are accounted for, MpSTing does contribute to a better understanding of potential threats to the banking system and consequently to ensuring financial stability.

Conclusion

In conclusion, the acute combination of appropriate design and expert judgment is essential for the effective conduct of a MpST. The theoretical aspects reveal the art of stress-testing with its benefits and complexities. It indeed fosters resilience, as long as the principles of credibility, relevance, and consistency are accounted for. Additional decisions shape the stress test architecture. First a clear goal has to be set as the foundation of the exercise. Then, the scope of banks, risk coverage, the type of exposures, assumptions, and constraints are defined. A specific attention is drawn on the selected approach considering the stress test purpose and available resources. For instance, the review shows that public exercises should be carried out in a BU manner, while the ones of supervisory nature could employ a BU or TD approach. Also, if the stress test is envisaged to cover dynamic features and second-round effects, the TD option is the leading preference. Next, application of scenarios and risk modeling provides the mechanism channel for translating severe, but plausible events into bank-specific shocks. Quality assurance and validation have a central role in the process in order to ensure robustness of outputs. The analysis brings forward the importance of supervisory data as the main input for quantitative estimates and the necessity of expert judgment as a way for channeling supervisory knowledge and managing consistency. The MpST ends with a report, which presents the results, assesses banks' capacity for absorbing negative shocks and points out the main fragilities. Depending on the mandate or when deemed appropriate, a public disclosure of outcomes may follow. In that manner, accompanying clear communication on the whole stress test process is a must. The simulation proceeds with a follow-up to the results that may prompt a reaction function from supervisors, bank management, or sources of last resort. If applicable, the dynamic responses could aggravate the stress test via second-round effects – macro feedbacks and/or bank contagion. In addition, several issues that remain a challenge are addressed with recommendations as a way forward. Most importantly, the decision on publication of results, approach, and balance-sheet

[23] See Goodhart 2016.

assumptions rely on the banking authority given the stress test goal and design, but supervisors could pursue a minimum disclosure of aggregate results for the banking system and more dynamic elements. Further work on the development of stress tests for calibrating MPIs, analyzing systemic risk, and merging the solvency and liquidity risk is welcomed and would be beneficial for macroprudential purposes.

MpSTing is part of the recipe for financial stability, not the panacea. It is a prominent instrument for assessing potential threats and vulnerabilities to the banking system, but a prudent way is to complement stress tests with other macroprudential tools. Being under the spotlight of growing interest, it is easy to attribute exceptional view to the exercise and form exaggerated expectations. However, stress tests are not prophetic or trying to tell the future, as they remain hypothetical exercises with underlying assumptions and limitations. In the words of Sir Henri Deterding: "It is impossible to see into the future, and it is dangerous not to". Thus, instead of predicting the future, MpSTing contributes to the analytical macroprudential assessment by highlighting areas with vulnerabilities and facilitates preparations for the unexpected.

References

Acharya, V. and S. Steffen, 2014. Falling Short of Expectations? Stress-Testing the European Banking System, CEPS Policy Briefs No. 315, Centre for European Policy Studies.

Acharya, V. and S. Steffen, 2016. Making Sense of the Comprehensive Assessment, in: Anderson, R. (Ed.), Stress Testing and Macroprudential Regulation: A Transatlantic Assessment, London: CEPR Press, pp. 109–19.

Anand, K., G. Bédard-Pagé and V. Traclet, 2014. Stress Testing the Canadian Banking System: A System-Wide Approach, Bank of Canada Financial System Review, June, pp. 61–68.

Basel Committee on Banking Supervision (BCBS), 2013. Liquidity Stress Testing: A Survey of Theory, Empirics and Current Industry and Supervisory Practices, BIS Working Paper No. 24, Bank for International Settlements.

Basel Committee on Banking Supervision (BCBS), 2015. Making Supervisory Stress Tests More Macroprudential: Considering Liquidity and Solvency Interactions and Systemic Risk, BIS Working Paper No. 29, Bank for International Settlements.

Bank of England (BoE), 2013. A Framework for Stress Testing the UK banking system. Available at: http://www.bankofengland.co.uk/financialstability/fsc/Documents/discussionpaper1013.pdf

Bank of England (BoE), 2015. The Bank of England's Approach to Stress Testing the UK Banking System. Available at: http://www.bankofengland.co.uk/financialstability/Documents/stresstesting/2015/approach.pdf

Bologna, P. and A. Segura, 2016. Integrating Stress Tests within the Basel III capital framework: A Macroprudentially Coherent Approach, Bank of Italy Occasional Papers No. 360, Bank of Italy.

Bookstaber, R., J. Cetina, G. Feldberg, M. Flood and P. Glasserman, 2014. Stress Tests to Promote Financial Stability: Assessing Progress and Looking to the Future, Journal of Risk Management in Financial Institutions 7 (1), pp. 16–25.

Borio, C., M. Drehmann and K. Tsatsaronis, 2012. Stress Testing Macro Stress Testing: Does it Live Up to Expectations?, BIS Working Paper No. 369, Bank for International Settlements.

Brazier, A., 2015. The Bank of England's Approach to Stress Testing the UK Banking System, speech at the London School of Economics Systemic Risk Centre, 30 October. Available at: http://www.bankofengland.co.uk/publications/Documents/speeches/2015/speech857.pdf

Brazier, A., 2017. How to: MACROPRU. 5 Principles for Macroprudential Policy, speech at the London School of Economics Financial Regulation Seminar, 13 February. Available at: https://www.bankofengland.co.uk/-/media/boe/files/speech/2017/how-to-macropru-5-principles-for-macroprudential-policy.pdf?la=en&hash=20D7E3AC92616F8B05B642608E7A9812C6D954D8

Breuer, T., M. Jandacka, K. Rheinberger and M. Summer, 2009. How to Find Plausible, Severe, and Useful Stress Scenarios, International Journal of Central Banking 5, pp. 205–24.

Brunnemeier, M. and L. Pedersen, 2009. Market and Funding Liquidity, Review of Financial Studies 22 (6), pp. 2201–38.

Bulgarian National Bank (BNB), 2016a. Report on the AQR and Stress Test of the Bulgarian Banking System. Available at: http://www.bnb.bg/bnbweb/groups/public/documents/bnb_download/bs_aqr_results_a1_en.pdf

Bulgarian National Bank (BNB), 2016b. Macroeconomic Scenarios for the Purposes of the Stress Test of the Bulgarian Banking System: Comparison with the Scenarios for the 2016 EU-Wide Stress Test and Current Macroeconomic Developments in Bulgaria, Economic Review 3, pp. 56–63.

Bulgarian National Bank (BNB), 2016c. 2016 Stress Test Guidelines. Available at: http://www.bnb.bg/bnbweb/groups/public/documents/bnb_download/bs_aqr_st_gl_1_en.pdf

Buncic, D., and M. Melecky, 2012. Macroprudential Stress Testing of Credit Risk: A Practical Approach for Policy Makers, WB Policy Research Working Paper No. 5936, World Bank.

Burrows, O, D. Learmonth and J. McKeown, 2012. RAMSI: a Top-down Stress-testing Model, BoE Financial Stability Paper No. 17, Bank of England.

Candelon, B. and A. Sy, 2015. How Did Markets React to Stress Tests?, IMF Working Paper No. 75, International Monetary Fund.

Čihák, M., 2004. Stress Testing: A Review of Key Concepts, Research and Policy Note No. 2, Czech National Bank.

Čihák, M., 2007. Introduction to Applied Stress Testing, IMF Working Paper No. 59, International Monetary Fund.

Committee on the Global Financial System (CGFS), 2012. Operationalising the Selection and Application of Macroprudential Instruments, CGFS Papers No. 48, Bank for International Settlements.

Constâncio, V., 2016a. The Role of Stress Testing in Supervision and Macroprudential Policy, in: Anderson, R. (Ed.), Stress Testing and Macroprudential Regulation: A Transatlantic Assessment, London: CEPR Press, pp. 51–68.

Constâncio, V., 2016b. Principles of Macroprudential Policy, speech at the ECB-IMF Conference on Macroprudential Policy, 26 April. Available at: https://www.ecb.europa.eu/press/key/date/2016/html/sp160426.en.html

Dees, S., J. Henry and R. Martin, (Eds.) 2017. STAMP€: Stress-Test Analytics for Macroprudential Purposes in the euro area, European Central Bank.

Demekas, D., 2015. Designing Effective Macroprudential Stress Tests: Progress So Far and the Way Forward, IMF Working Paper No. 146, International Monetary Fund.

Dent, K., B. Westwood and M. Segoviano, 2016. Stress Testing of Banks: An Introduction, Bank of England Quarterly Bulletin, pp. 130–43.

Dowd, K., 2015. Central Bank Stress Tests: Mad, Bad, and Dangerous, Cato Journal 35 (3), pp. 507–24.

Eisenberg, L. and T. Noe, 2001. Systemic Risk in Financial Systems, Management Science 47 (2), pp. 236–49.

Elsinger, H., A. Lehar and M. Summer, 2006. Risk Assessment for Banking Systems, Management Science 52 (9), pp. 1301–41.

Enria, A., 2017. Introductory remarks at EBA-IMF Colloquium, New Frontiers on Stress Testing. Available at: https://www.eba.europa.eu/documents/10180/1768574/Andrea+Enria+introductory+remarks+-+EBA-IMF+Colloquium.pdf

Espinosa-Vega, M. and J. Solé, 2010. Cross-Border Financial Surveillance: A Network Perspective, IMF Working Paper No. 105, International Monetary Fund.

European Banking Authority (EBA), 2015. Draft Guidelines on Stress Testing and Supervisory Stress Testing.

European Banking Authority (EBA), 2016. 2016 EU-Wide Stress Test Methodological Note. Available at: http://www.eba.europa.eu/documents/10180/1259315/2016+EU-wide+stress+test-Methodological+note.pdf

European Central Bank (ECB), 2006. Macro Stress Testing Practices, Financial Stability Review, June, pp. 147–154.

European Central Bank (ECB), 2014. Comprehensive Assessment Stress Test Manual. Available at: https://www.ecb.europa.eu/pub/pdf/other/castmanual201408en.pdf

European Systemic Risk Board, 2016. Adverse macro-financial scenario for the EBA 2016 EU-wide bank stress testing exercise, Available at: https://www.esrb.europa.eu/pub/pdf/other/eu-wide_stress_test_adverse_macrofinancial_scenario.pdf?bc317f1a9b60f86fa893ac991de5e52e

Fernandes, M., D. Igan and M. Pinheir, 2015. March Madness in Wall Street: (What) Does the Market Learn from Stress Tests?, IMF Working Paper No. 271. International Monetary Fund.

Financial Times, 2016. Banks Pressed to Step Up Defences Against Cyber Attack, https://www.ft.com/content/2ac58a5c-2422-11e6-9d4d-c11776a5124d, 27 May 2016.

Flannery, M., B. Hirtle and A. Kovner, 2017. Evaluating the Information in the Federal Reserve Stress Tests, Journal of Financial Intermediation 29, pp. 1–18.

Foglia, A., 2009. Stress Testing Credit Risk: A Survey of Authorities' Approaches, International Journal of Central Banking 5 (3), pp. 1–37.

Goldstein, I. and H. Sapra, 2014. Should Banks' Stress Test Results be Disclosed? An Analysis of the Costs and Benefits, Foundations and Trends in Finance 8 (1), pp. 1–54.

Goldstein, I. and Y. Leitner, 2015. Stress Tests and Information Disclosure, FRBP Working Paper No. 15-10, Federal Reserve Bank of Plihadelphia.

Goodhart, C., 2016. In Praise of Stress Tests, in: Anderson, R. (Ed.), Stress Testing and Macroprudential Regulation: A Transatlantic Assessment, London: CEPR Press, pp. 141–53.

Gouriéroux, C., J. Héam and A. Monfort, 2012. Bilateral Exposures and Systemic Solvency Risk, Canada Journal of Economics 45 (4), pp. 1273–1309.

Haldane, A., 2009. Why Banks Failed the Stress Test, speech at the Marcus-Evans conference on stress testing, 13 February. Available at: http://www.bankofengland.co.uk/archive/documents/historicpubs/speeches/2009/speech374.pdf

Henry, J. and C. Kok, 2013. A Macro Stress Testing Framework for Assessing Systemic Risks in the Banking Sector, ECB Occasional Paper Series No. 152, European Central Bank.

Hilbers, P., M. Jones and G. Slack, 2004. Stress Testing Financial Systems: What to Do When the Governor Calls, IMF Working paper No. 127. International Monetary Fund.

Hirtle, B. and A. Lehnert, 2014. Supervisory Stress Tests, FRBNY Staff Report No. 696, Federal Reserve Bank of New York.

Hristev, V., 2014. Bank Stress-Testing: Lessons from Central, Eastern and Southeastern European Countries, Journal of Economic Alternatives 4, pp. 92–109.

Hsiao, KW. And Z. Jiang, 2015. The Pre- and Post-Crisis Stress Testing in the Banking Sector - A Literature Review, Global Credit Review 5, pp. 77–97.

International Monetary Fund (IMF), 2003. External Debt Statistics: Guide for Compilers and Users – Appendix III, Glossary.

International Monetary Fund (IMF), 2012. Macrofinancial Stress Testing – Principles and Practices, IMF Policy Paper.

International Monetary Fund (IMF) and World Bank, 2005. Financial Sector Assessment: A Handbook – Appendix D, Stress Testing.

Jobst, A., L. Ong and C. Schmieder, 2013. A Framework for Macroprudential Bank Solvency Stress Testing: Application to S-25 and Other G-20 Country FSAPs, IMF Working Paper No. 68, International Monetary Fund.

Kapinos, P., C. Martin and O. Mitnik, 2015. Stress Testing Banks: Whence and Whither?, FDIC Center for Financial Research Working Paper No. 07, Federal Deposit Insurance Corporation.

Kunghehian, N., 2013. Breaking the Silos in Stress Testing, in: Stress Testing: European Edition, W. Xoual (Ed.), Moody's Analytics Risk Perspectives, Vol. 1, pp. 14–18.

Licari, J. and J. Suárez-Lledó, 2013. A Macroeconomic View on Stress Testing, in: Stress Testing: European Edition, W. Xoual (Ed.), Moody's Analytics Risk Perspectives, Vol. 1, pp. 56–60.

Melecky, M. and A. Podpiera, 2010. Macroprudential Stress-Testing Practices of Central Banks in Central and South Eastern Europe: An Overview and Challenges Ahead, WB Policy Research Working Paper No. 5434, World Bank.

Moretti, M., Stolz, S. and Swinburne, M., 2008. Stress Testing at the IMF, IMF Working Paper No. 206, International Monetary Fund.

Ong, L. and C. Pazarbasioglu, 2013. Credibility and crisis stress testing, IMF Working Paper No. 178, International Monetary Fund.

Petrella, G. and A. Resti, 2013. Supervisors as Information Producers: Do Stress Tests Reduce Bank Opaqueness, Journal of Banking & Finance 37, pp. 5406–20.

Pierret, D., 2014. Systemic Risk and the Solvency-Liquidity Nexus of Banks, mimeo, NYU Stern School of Business.

Puhr, C. and S. Schmitz, 2013. A View from the Top – The Interaction between Solvency and Liquidity Stress, Journal of Risk Management in Financial Institutions 7 (1), pp. 38–51.

Quagliariello, M., 2009. Macroeconomic Stress-Testing: Definitions and Main Components, in: Stress-Testing the Banking System: Methodologies and Applications, Quagliariello, M. (Ed.), Cambridge University Press, pp. 18–36.

Quagliariello, M., 2017. The Design of a Stress Test in 5 Q&As, Presentation at Joint IMF-EBA Colloquium: New Frontiers on Stress Testing. Available at: https://www.eba.europa.eu/documents/10180/1768556/Session+VI+Mario+Quagliariello.pdf

Schuermann, T., 2016. Stress Testing in Wartime and in Peacetime, in: Anderson, R. (Ed.), Stress Testing and Macroprudential Regulation: A Transatlantic Assessment, London: CEPR Press, pp. 125–39.

Schmieder, C., C. Puhr and M. Hasan, 2011. Next Generation Balance Sheet Stress Testing, IMF Working Paper No. 83, International Monetary Fund.

Sorge, M., 2004. Stress-Testing Financial Systems: An Overview of Current Methodologies, BIS Working Paper No. 165, Bank for International Settlements.

Tarullo, D., 2016. Next Steps in the Evolution of Stress Testing, remarks at Yale University School of Management Leaders Forum, 26 September. Available at: https://www.federalreserve.gov/newsevents/speech/files/tarullo20160926a.pdf

Van den End, J., 2008. Liquidity Stress-Tester: A Macro Model for Stress-Testing Banks' Liquidity Risk, DNB Working Paper No. 175, Netherlands Central Bank.

Woo, V., B. Chelluboina and W. Xoual, 2014. The Effectiveness of the Regulatory Stress Testing Disclosure Process, Financial Markets, Institutions & Instruments 23 (1), pp. 1–70.

DP/109/2018

Constructing an Uncertainty Indicator for Bulgaria

Evgeni Ivanov

Abstract. Given uncertainty's potential role as a driver of business cycles the aim of this paper is to construct a measure of economic uncertainty for the Bulgarian economy. For that purpose, based on monthly qualitative data on Bulgaria from the European Commission Business and Consumer Survey, four different indicators of uncertainty are constructed, all sharing the key assumption that disagreement among respondents can give information on the level of uncertainty in the economy. Since each of these indicators captures a different dimension of the "true" underlying uncertainty, a single composite uncertainty indicator for Bulgaria is also constructed. The results suggest that the uncertainty indicators for Bulgaria are negatively correlated with key macroeconomic variables such as GDP, investment and consumption. Global uncertainty-enhancing shocks appear to be much more significant than domestic ones with the 2008/09 global economic and financial crisis and the 2012 sovereign debt crisis in the EU standing out as one of the clearest examples of periods of heightened uncertainty in Bulgaria. Almost all indicators suggest that uncertainty in Bulgaria in the period after 2013 has fallen to below average levels.

Резюме. Поради ролята на несигурността като потенциален двигател на бизнес цикъла целта на настоящото изследване е да конструира измерител на икономическата несигурност за българската икономика. За тази цел въз основа на месечни данни за България с качествен характер, получени от редовното Проучване на Европейската комисия за нагласите на бизнеса и потребителите, са конструирани четири различни индикатора на несигурността, като всички те се основават на ключовото допускане, че несъгласие сред респондентите може да даде информация относно равнището на несигурност в икономиката. Тъй като всеки от тези индикатори улавя различно измерение на „истинската" ненаблюдавана несигурност, анализът конструира и обединяващ индикатор на несигурността за България. Резултатите показват, че индикаторите на несигурност за България имат отрицателна корелация с основни макроикономически променливи, като БВП, инвестиции и потребление. Глобални шокове, засилващи несигурността, като световната икономическа и финансова криза през 2008–2009 г. и дълговата криза в ЕС през 2012 г. изглежда имат значително по-голямо значение за динамиката на индикатора на несигурност в България спрямо вътрешните за страната шокове. Почти всички индикатори показват, че в периода след 2013 г. несигурността в България е спаднала под средните ретроспективни равнища.

Evgeni Ivanov, Economic Research and Forecasting Directorate, Bulgarian National Bank, Ivanov.e@bnbank.org

Acknowledgment:

I would like to thank the members of the Publications Council of the Bulgarian National Bank and particularly Mariella Nenova, Ass. Prof., Ph.D., and Andrey Vassilev, Ph.D., for their valuable contribution, comments and suggestions.

Contents

1. Introduction .....

2. Overview of the Main Types of Uncertainty Indicators .....

3. Motivation of the Choice of Methodology .....

4. Construction of the Uncertainty Indicators for Bulgaria .....
   4.1. Uncertainty Indicator U1 .....
   4.2. Uncertainty Indicator U2 .....
   4.3. Uncertainty Indicator U3 .....
   4.4. Uncertainty Indicator U4 .....

5. Combining Different Uncertainty Indicators .....

6. Relationship between Uncertainty and Key Macroeconomic Variables .....

7. Conclusion .....

Appendices
   I. Pre-identified Uncertainty-enhancing Events .....
   II. Questions from the BCS Used to Compute U1 .....
   III. Questions from the BCS Used to Compute U2 .....

Bibliography .....

1. Introduction

The reasons behind the sluggish recovery of economic activity in the EU and other advanced economies after the global financial crisis in 2008/09 have been widely discussed in recent years and empirical results suggest that uncertainty has been one of the main factors (ECB, 2016:1; ECB, 2016:2; Barkbu et al., 2015). This is somewhat expected given that in the years after the crisis the EU economy has been constantly subjected to uncertainty-enhancing shocks. Among them worth mentioning are the sovereign debt crisis of 2011/12, the unrests in Ukraine in 2014, the negotiations with Greece on debt restructuring in 2015, the Brexit Referendum in the UK, the failed coup in Turkey and the change in political course of the US after Donald Trump became president, the last three of which happened in 2016. One can also include the intensification of terrorist attacks, as well as the on-going refugee crisis in the EU, which started in 2013 but intensified in 2015. Registering these empirical findings there have been a growing number of publications in recent years that try to answer more precisely the question on what is the role of uncertainty as a driver of macroeconomic fluctuations and how uncertainty can be best measured (EC, 2017).

For Bulgaria the role of uncertainty for macroeconomic developments since the early 2000s is likely to have been significant, since the high openness of the economy would suggest that external shocks are quickly transmitted to the domestic economy. What is more, especially in the period after the global economic and financial crisis in 2008/09 the Bulgarian economy had experienced its own domestic uncertainty-enhancing shocks, mainly related to political uncertainty[1], as well as to the placement of KTB bank under special supervision in 2014.

Economic uncertainty, or also known as "Knightian uncertainty" (Knight, 1921), can be defined as a situation when economic agents lack the knowledge to assess the current and future economic situation with sufficient confidence and/or predict future outcomes (ECB, 2016:2). When deciding on consumption and investment, firms and consumers form expectations on relevant future events on the basis of available data (Manteu and Serra, 2017). These expectations are affected by uncertainty, to the extent that the likelihood of alternative events is unknown or impossible to gauge with precision. In a broad sense, uncertainty can be illustrated as a mean preserving increase in the "tails" of the probability distribution of an event (EC, 2013). An increase in uncertainty makes future outcomes more uncertain in the sense that "tail events", or realisa-

[1] No government after 2009 had managed to complete its full 4-year term, which has led to frequent preliminary parliamentary elections, resulting in four different official governments and three caretaker governments in less than 10 years.

tions at the extreme of the distribution, have a higher probability of occurring (EC, 2013). Wars, terrorist attacks or other unprecedented events are examples where it might be impossible for economic agents to assess the likelihood of the event or its economic impact (ECB, 2016:2).[2]

Economic theory suggests that a high level of uncertainty can depress economic activity through several channels among which investment, employment, consumption and increased risk premium (EC, 2013). The main channel is the so-called "wait-and-see" effect. Faced with high uncertainty economic agents (firms and consumers) might decide to postpone spending decisions in order to avoid making costly mistakes (EC, 2013). A high level of uncertainty gives agents an incentive to delay or cancel decisions involving considerable irreversible costs until uncertainty is resolved and more information is available, thus restraining economic activity (Manteu and Serra, 2017). If firms are uncertain regarding the future demand for their products, they may be forced to cut back on hiring new employees. Precautionary savings might be another channel of transmission of uncertainty shocks. Increased uncertainty regarding economic agents expected income may result in lower current consumption and an increase in current savings in order to smooth future consumption. Other channels through which uncertainty can dampen economic activity include a higher cost of finance, lower asset prices due to increased risk premium (Gilchrist et al., 2010) and reduced willingness of banks to provide financing (Manteu and Serra, 2017).

Since uncertainty is unobservable and one cannot directly measure it, different studies have proposed alternative measures that should ideally reflect the level of uncertainty. However, the results of all these studies are heavily reliant on the choice of methodology for the construction of the uncertainty indicator. The purpose of this paper is to give a short overview of the most commonly-used methods for the construction of an uncertainty indicator with their pros and cons and then to propose and construct a relevant metric for Bulgaria. In terms of contribution to the existing literature this papers is the first to apply these methodologies for Bulgaria and its results can be used for further analysis, particularly in explaining macroeconomic developments, and for nowcasting.

This paper is structured as follows: Section 2 gives a short overview of the most commonly used methods for the construction of uncertainty indicators with

[2] One should clearly distinguish between risk and uncertainty. While uncertainty is related to a situation when economic agents cannot form clear views on the possible future states, risk is a situation where agents are able to form views about the probability distribution of possible future states, based on logic (like when assessing likely outcomes of throwing a dice) or on past experience (if similar shocks – to oil prices, exchange rates, *etc.* – have frequently happened in the past).

their pros and cons, Section 3 motivates the methodology, chosen by the paper, Section 4 is the main section and it contains information on the methodology implementation and results on 4 different uncertainty indicators for Bulgaria. Section 5 proposes ways in which the 4 different indicators, reflecting different dimensions of uncertainty, could be combined in one single indicator. Section 6 then offers some basic descriptive analysis on the relationship between the proposed uncertainty indicators and key macroeconomic variables and, finally, Section 7 concludes the analysis.

2. Overview of the Main Types of Uncertainty Indicators

There are a number of methods that are used in the literature to indirectly measure different aspects of uncertainty and each has its pros and cons.

Uncertainty Indicators Based on Financial Market Volatility:

Financial market uncertainty indicators are typically based on stock market volatility. An advantage is that proxies for uncertainty based on financial market volatility can be calculated in various ways (using data on equity prices, bond yields or exchange rates for example) and at high frequency (typically on daily basis). Low volatility in equity, bond or foreign exchange markets should indicate stable expectations about future outcomes broadly shared across market participants, while heightened volatility should reflect financial market uncertainty about future outcomes (ECB, 2016:2). Probably the most widely-recognised indicator for uncertainty from this family is the VIX index, which was introduced in 1993 by the Chicago Board Options Exchange. The VIX measures the market's expectation of near-term volatility and, hence, reflects the uncertainty that market participants face. It is constructed using the implied volatilities of a wide range of S&P 500 index options, using both calls and puts and, thus, is meant to be forward-looking[3]. However, financial market volatility can change over time even if there is no change in uncertainty about the economic outlook, *i.e.* when changes in risk aversion or sentiment are the main driving factors of market volatility (ECB, 2016:2). In addition, perceptions of uncertainty derived from financial markets might follow different logic from that of business and private households (ECB, 2016:2). Moreover, in countries where the stock market is not well-developed, like in Bulgaria, the volatility of stock market returns is likely to be a poor indicator of uncertainty. However, despite their shortcomings, uncertainty indicators, constructed by using similar

[3] For more information on the calculation of the VIX, please refer to: http://www.cboe.com/micro/vix/vixwhite.pdf

methodology to that of the VIX index, have been widely used in empirical studies on the topic of uncertainty (Bloom, 2009; Moore, 2016; ECB, 2016:1; Meinen and Roehe, 2016; and EC, 2017).

Economic Policy Uncertainty Indicators:

Baker et al. (2015) has developed the so-called "Economic Policy Uncertainty Index" (EPU), which is an index that measures uncertainty, based on frequency counts of newspaper articles containing the words "uncertainty" or "uncertain", "economic" or "economy", as well as other terms, such as "deficit", "regulation", "terrorism", "war", *etc.*[4] Baker et al. (2015) find a strong relationship between the EPU and other measures of economic uncertainty, like implied stock market volatility. They also find that increases in policy uncertainty foreshadow declines in investment, output, and employment in the United States and, in a panel VAR setting, for 12 major economies (Baker et al, 2015). This index is publicly available and regularly updated, which makes it useful for empirical studies that focus on the countries covered by the indicator (Busseti et al., 2015; Meinen and Roehe, 2016; ECB, 2016:2). It covers the US, Europe, as a whole, and key countries from the euro area such as Germany, France, Italy and Spain. Bulgaria and the rest of the countries that joined the EU after 2000 are not covered by the index. However, given the computational complexity of the index, as well as its reliance on publicly available web-based newspaper archives, which vary from country to country, the replication of this type of indices is not yet that widely used. What is more, indicators like the EPU might suffer from a subjective selection of "crisis" words, whose meaning and/or contextual use might moreover change over time, as well as on the choice of newspapers, which might not be representative of the media coverage in their countries (ECB, 2016:2).

Survey-Based Uncertainty Indicators:

Another popular method for calculating indicators for uncertainty is by making use of the qualitative data from the business and consumer surveys that are typically carried out on monthly basis (EC, 2017). The derived uncertainty indicators measure the divergence of business and consumer expectations about the economy and about their finances. The underlying assumption is that the more economic agents disagree in their expectations about future activity, the higher the uncertainty in the economy. These uncertainty indicators became popular with the work of Bachmann et al. (2013) who applied such an approach on US and German data. Since then the methodology has been widely used in the

[4] For more information on the calculation of the EPU, please refer to: http://www.policyuncertainty.com/methodology.html

construction of uncertainty indicators (EC, 2013; Busseti et al., 2015; Meinen and Roehe, 2016; ECB, 2016:2), thanks to its simplicity, coverage, comparability, ease of replication and information richness. One of the main pros of such types of uncertainty indicators is that economic uncertainty is measured based on the replies of individual economic agents, which contrasts with other measures that capture uncertainty indirectly via financial market indicators or occurrences of certain words in the press (EC, 2013). However, the heterogeneity among firms and consumers, as well as differences in the information they possess can cause divergence in opinion across the business cycle irrespective of the subjectively perceived uncertainty (Girardi and Reuter, 2017). As such, the obtained dispersions of expectations could also reflect idiosyncratic issues. This paper will pay special attention to these types of uncertainty indicators in the next two sections of the paper, since their characteristics make them particularly appealing for analysing uncertainty in Bulgaria.

Other Uncertainty Indicators:

Among the other less commonly used uncertainty indicators is the disagreement among professional forecasters, which is obtained as the average standard deviation of the individual probability distributions provided by the professional forecasters (ECB, 2016:2). The idea here is somewhat similar to that of survey-based uncertainty indicators and postulates that increasingly diverse opinions about the economic outlook among forecasters are likely to indicate that the heightened uncertainty makes it more difficult to project future economic developments. A potential downside of measuring dispersion among professional forecasters is that particularly in times of turbulence there might be a tendency of "herding" among forecasters (EC, 2017). Another drawback of these indicators is that the frequency with which they can be calculated may not be as high as that of financial or survey-based uncertainty indicators due to the availability of observations.

Another less commonly-used method to calculate uncertainty indicators is based on obtaining forecast errors from factor models that include large number of variables (ECB, 2016:2). The idea is that increased and more volatile forecast errors would be associated with higher uncertainty. Such an approach has been used in Meinen and Roehe (2016) and the obtained results are comparable to those of other uncertainty metrics analysed in that paper.

3. Motivation of the Choice of Methodology

This analysis has chosen to use measures of uncertainty constructed from survey data, given their advantages over other measures, discussed in Section 2. As already mentioned uncertainty indicators obtained from survey data are simple to calculate, they have high frequency (typically monthly data), they are published timely (usually with a one month lag), they have good coverage (for Bulgaria the starting data of the indicators that can be constructed go back to the early 2000s) and are rich in information, since the sectoral breakdown and the diversity of the questions asked in the surveys can give an idea of the underlying source of increased/decreased uncertainty. Since in this setup economic uncertainty is measured "at source", *i.e.* directly at the level of the economic agents (consumers, firms) who make investment and consumption spending decisions, these uncertainty measures should be more informative about the state of the economy, compared to other uncertainty indicators, constructed in a different way.

As already discussed, survey-based uncertainty measures should be interpreted with caution since they suffer from a non-negligible weakness: notably that their evolvement does not solely reflect changes in underlying uncertainty levels, but also in other forces such heterogeneity among firms and consumers, as well as differences in the information they possess (Girardi and Reuter, 2017). Heterogeneity issues arise from the fact that firms and consumers can have different views on their prospects, as well as current and recent developments, depending on their income, sector of operation, or need of funding. Apart from that, respondents might come up with different answers to the survey questions, because they use different information sets (Girardi and Reuter, 2017). Their assessments of the different concepts inquired in the surveys (*e.g.* the business or income situation) might thus vastly differ and translate into high survey data dispersion, without this necessarily indicating that respondents are uncertain about their assessments. In such a case respondents actually disagree, rather than sense elevated uncertainty (Girardi and Reuter, 2017). However, if the variance of this noise rises, the corresponding increase in the dispersion of survey data could be interpreted as genuine uncertainty.

To obtain the uncertainty indicators for Bulgaria data from the European Commission Business and Consumer Survey (BCS)[5] is used and the paper follows the methodology proposed by Girardi and Reuter (2017) and also used by the ECB (2016:2), by Meinen and Roehe (2016) and by Manteu and Serra (2017). The use of the BCS data has several advantages. The most important one is

[5] For more information see European Commission Business and Consumer Surveys: https://ec.europa.eu/info/business-economy-euro/indicators-statistics/economic-databases/business-and-consumer-surveys_en

that the data collection is based on the same methodology and principals across all countries within the EU, which means that uncertainty indicators derived from the BCS are directly comparable across countries. This is quite useful since it allows one to distinguish between global and country-specific episodes of heightened uncertainty by comparing an uncertainty indicator across different countries. The BCS also offers excellent data coverage for Bulgaria with monthly data, available from the early 2000s, based on a nominal sample size of 1200 respondents from the industrial sector, 1350 respondents from the services sector, 1000 respondents from the consumer sector, 1050 respondents from the retail trade sector and 600 respondents from the construction sector[6].

In the next section the paper will present the methodology used to obtain four different indicators for economic uncertainty for Bulgaria. As previously discussed the core underlying assumption for all four uncertainty indicators is that the more economic agents disagree in their expectations about current and future economic activity and their finances, the higher is the uncertainty in the economy. The four indicators reflect different aspects of economic uncertainty and since they are obtained from the same population distribution this paper will also try to obtain a single indicator that would ideally best explain the "true" underlying uncertainty in the Bulgarian economy (see Section 5).

4. Construction of the Uncertainty Indicators for Bulgaria

Using the EC's BCS data for Bulgaria this paper has constructed 4 different uncertainty indicators (which from now on are called U1, U2, U3 and U4) by following the methodology, proposed by Girardi and Reuter (2017). All indicators are standardised so as to have zero mean and unit standard deviation. This standardisation is essential to make individual indicators comparable in terms of both their mean level and volatility and it also allows me to combine the four indicators into a single one. Apart from that all four uncertainty indicators have been indirectly seasonally adjusted by seasonally adjusting the raw data on the replies of the BCS respondents before carrying out the calculations to obtain the uncertainty indicators.

To check whether these indicators are capable of detecting periods of heightened uncertainty the historical development of each indicator are plotted alongside the starting dates of pre-identified shocks to the global and EU economy that are consider as uncertainty-enhancing (such as the 2008/09 global economic and financial crisis[7] and the sovereign debt crisis in EU), as well as country-

[6] According the EC's "Joint Harmonised EU Programme of Business and Consumer Surveys User Guide", as of March 2016.

[7] Which from here on in paper we also refer to as the "Great Recession"

specific uncertainty-enhancing events in Bulgaria, related mainly to parliamentary elections[8]. The hypothesis that will be considered throughout the analysis is that these pre-identified shocks were periods of heightened uncertainty and the calculated uncertainty indicators should ideally reflect that. Global events are identified in the graphs with a red line, while domestic events are identified with a purple line.

4.1. Uncertainty Indicator U1

The first indicator (U1) measures the disagreement for future activity across all sectors (industry, services, retail trade, consumers, construction) and, thus, should capture uncertainty at the level of the whole economy. This is an expectations-based uncertainty indicator.

Similar to measures based on the dispersion of professional forecasts, but applying the idea to a much larger and more heterogeneous group, the approach rests on the idea that respondents, when replying to questions about their economic expectations, will be more or less concordant in their views, depending on the degree of uncertainty about the future course of the economy (EC, 2017). In times of low uncertainty respondents can be expected to have broadly similar expectations about future developments, while an increasing disagreement between responses indicates increasing uncertainty. To capture this element of dispersion in survey responses, the approach amounts to calculating the cross-sectional standard deviation of individual qualitative survey responses[9] (Bachmann et al., 2013).

This cross-sectional standard deviation is equivalent to:

$$\textbf{Standard deviation on each question}_{\,t} = \sqrt{\boldsymbol{p_t^+} + \boldsymbol{p_t^-} - (\boldsymbol{p_t^+} - \boldsymbol{p_t^-})^2} \qquad (1)$$

where $\boldsymbol{p}^+$ is the fraction of "increase" responses to a survey question at time $\boldsymbol{t}$ and $\boldsymbol{p}^-$ is the fraction of "decrease" responses to a survey question at time $\boldsymbol{t}$.[10]

As to the intuition behind the formula, the first sum under the square root shows that the standard deviation (or dispersion) increases when more respondents give either positive of negative replies, as opposed to neutral replies ("unchanged"). Put differently, the higher the share of "unchanged" replies, the lower is uncertainty among respondents according to this measure, which is based on the assumption that an "unchanged" view on the future is a sign of

[8] For the complete list of the pre-identified uncertainty-enhancing events, please refer to Appendix I.

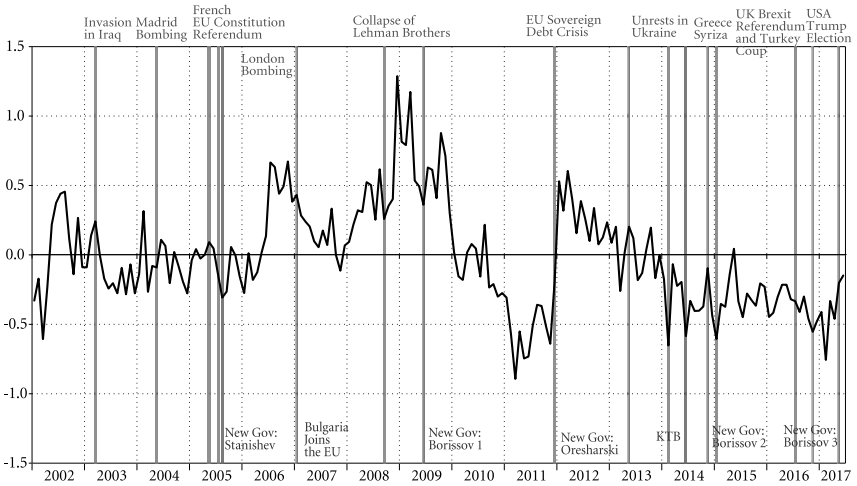
[9] For a detailed list of the questions used to compute U1, please refer to Appendix II.

[10] In the consumer survey there are five possible responses: "very positive", "positive", "very negative", "negative" and "neutral". In this case $\boldsymbol{p}^+$ is computed by summing the fraction of very positive answers with the fraction of positive responses. The same steps are applies when computing $\boldsymbol{p}^-$.

perceived stability. The standard deviation on each question also increases when the second, the quadratic term, gets smaller. The squared difference between positive and negative replies actually measures the dominance of optimistic or pessimistic respondents; if there is no clear dominance, *i.e.* high disagreement, the second term gets small and uncertainty rises; if on the other hand there is strong dominance of either positive or negative views, then these concordant views about the future translate into a large second term, and, thus, a reduction in uncertainty is observed (EC, 2017).

The resulting dispersion on a given answer for a given time period is then standardised as to have zero mean and unit standard deviation. These standardised dispersions from each separate question are then averaged within sectors (such as industry, consumers, construction, *etc.*) to obtain the uncertainty indicators for each sector of the economy. The separate sectoral uncertainty indicators are then averaged[11] across economic sectors in order to obtain the overall uncertainty measure for the whole economy which is our U1 indicator (see Figure 1).

Figure 1. **Uncertainty in Total Economy, Bulgaria** (U1 Indicator)



Note: Since the index is standardised a value of zero would mean that the level of uncertainty in the economy is "average", while values above zero are period with heightened uncertainty and values below zero are periods with "below-average" uncertainty. Movements along the Y-axis are measured in standard deviations due to the standardisation transformation.

Source: EU Commission BCS Survey, own calculations

[11] As an alternative to averaging uncertainty across sectors of the economy, one can assign each sector weights equal to the value added share of that sector in the total value added of the economy under consideration.

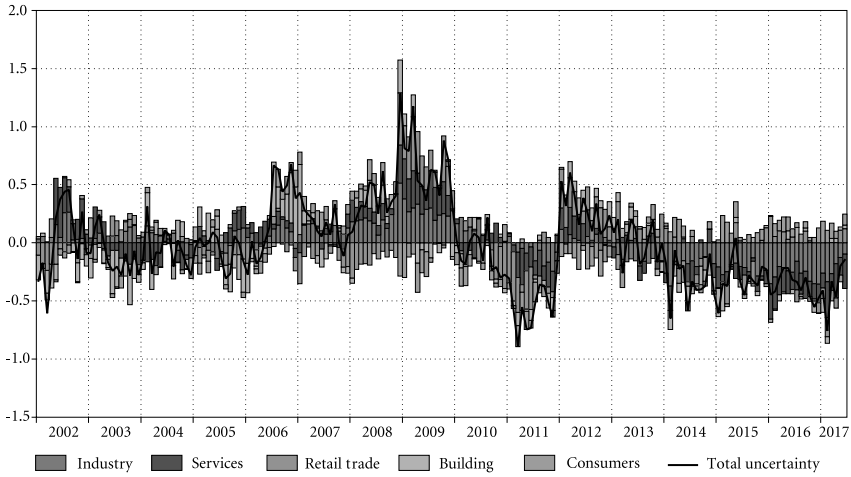
Due to the standardisation, a value of zero on the uncertainty index should be interpreted as "average" uncertainty during the period under consideration, while values above zero should indicate a period of "above average" uncertainty and values below zero – a period with "below average" uncertainty.

The U1 indicator indeed captures well major periods of heightened uncertainty, such as the Great Recession and the EU sovereign debt crisis. However, some of the other global events that have been initially selected and included as uncertainty-enhancing (such as the Brexit Referendum and the unrests in Ukraine) do not result in significant upward spikes of the indicator that would suggest "above average" uncertainty. This could be explained with the fact that these events, despite being uncertainty-enhancing, are not occurring in our main trading partners and, thus, export-oriented sectors in Bulgaria are not faced with uncertainty regarding future orders. One can also notice that country-specific events in Bulgaria (such as the parliamentary elections or the placement of KTB under special supervision) do not generally increase the uncertainty in the country to above-average levels, as measured by U1. The only exception to this is the period before Bulgaria's accession in the EU in 2007, when the indicator increases up to the date of accession and then decreases notably right after the accession. This can be explained with the mixed expectations that people had regarding what would happen to the Bulgarian economy once the country joins the EU.

A useful feature of this type of computation of uncertainty indicators is that one can get the contribution to uncertainty from each sector of the economy given the way the indicator is constructed – the overall uncertainty in the economy is calculated as the average uncertainty across each sector of the economy (see Figure 2). The extent to which uncertainty in different sectors reacts to global and domestic events can vary, depending on the degree of reliance of the particular sector to domestic or foreign markets. In that sense industry and services may depend more on external developments in comparison to consumers and retail trade which will likely depend more on domestic developments.

The sectoral breakdown suggests that the recently observed decrease in uncertainty has been mainly due to industry and to a lesser extent to the services sector. This can probably be explained with the cyclical recovery in the economic activity in our main trading partners from the EU and, respectively, the subsequent increase in external demand for Bulgarian goods and services. Having in mind that firms in the industrial sector in Bulgaria are strongly export-oriented and are highly integrated in global value chains (Ivanova and Ivanov, 2017) an increase in external demand would likely lead to more optimistic future expectations.

Figure 2. **Uncertainty in Total Economy, Bulgaria, Sectoral Breakdown** (U1 Indicator)



Source: EU Commission BCS Survey, own calculations

One can notice that despite the apparent hikes of the uncertainty indicator in 2009 and 2012 U1 is also volatile from one month to the next, which can be explained with the inherent volatility of respondent's answers to the BCS. As already discussed in Section 3 firms' and consumers' assessments regarding the same survey questions are noisy, due to the differences in background characteristics, as well as disagreements due to different information sets.

4.2. Uncertainty Indicator U2

The second indicator (U2) does not only focus on disagreement for future activity but also takes into account respondents' backward-looking assessments for the same questions in subsequent survey rounds (usually three months later for industry, services, retail trade and 12 months later for consumers) across all sectors (industry, services, retail trade, consumers) excluding construction (due to lack of suitable question pairs in the survey)[12]. This indicator is a type of *ex-post* uncertainty measure, based on indirect evaluation of respondents' forecast errors, which combines the uncertainty regarding future activity with respondents' assessment about activity during the past months (Girardi and Reuter, 2017). This means that forward-looking questions such as "How do you expect your production to develop over the next 3 months?" will be combined

[12] For a detailed list of the questions used to compute U1, please refer to Appendix III.

with their corresponding backward-looking questions, such as "How has your production developed over the past 3 months?". One drawback of this type of index is that it can only be constructed with significant time lag because it relies on respondents' retrospective assessment of past developments (Manteu and Serra, 2017). Another drawback is that the indicator rests on the assumption that the economic conditions remain stable between the time of the ex-ante and the ex-post assessments. Should the conditions change, the heterogeneity-induced dispersion of the questions can no longer be reasonably assumed to be similar (Girardi and Reuter, 2017).

The calculation procedure starts in a similar manner to the one used to obtain U1. In the first step the cross-sectional standard deviation from the response on each individual qualitative survey question is calculated.

$$\textbf{Standard deviation on each question}_{\mathbf{t}} = \sqrt{p_t^+ + p_t^- - (p_t^+ - p_t^-)^2} \qquad (2)$$

where p^+ is the fraction of "increase" responses to a survey question at time t and p^+ is the fraction of "decrease" responses to a survey question at time t.[13]

In the second step the obtained standard deviations of the corresponding forward- and backward-looking survey responses by pairs are grouped and for every single pair rescaled dispersions are obtained in the following way:

$$\textbf{Dispersion on each question (Industry, Services, Retail Trade)}_{\mathbf{t\text{-}3}} =$$

$$\ln\left(\frac{\textbf{Standard deviation (question on future)t–3}}{\textbf{Standard deviation (question on past)t}}\right) \qquad (3)$$

$$\textbf{Dispersion on each question (Consumers)}_{\mathbf{t\text{-}12}} =$$

$$\ln\left(\frac{\textbf{Standard deviation (question on future)t–12}}{\textbf{Standard deviation (question on past)t}}\right) \qquad (4)$$

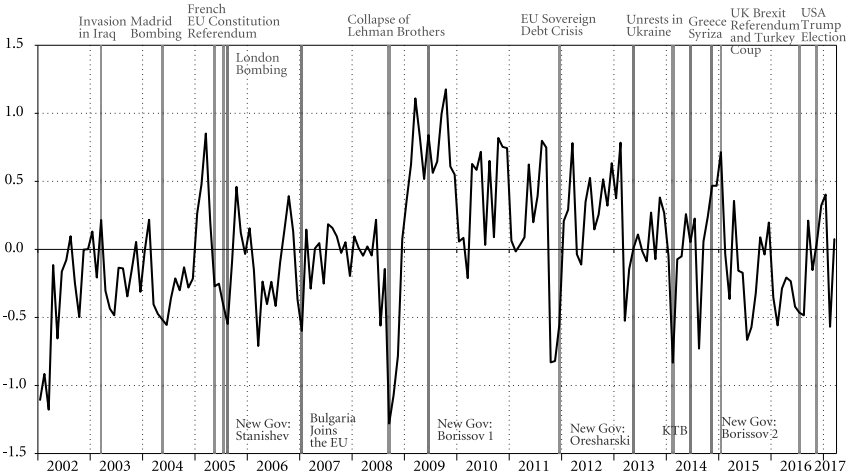
While dispersion in answers to forward-looking questions can be influenced by uncertainty and other factors (namely, heterogeneity and disagreement), dispersion in answers to backward-looking questions should not reflect uncertainty. In practice, the indicator involves scaling the dispersion of answers to the forward-looking questions, as inquired in a given month, by the dispersion of answers to the corresponding backward-looking questions, as inquired some months later, which can be interpreted as a measure of the extent of uncertainty,

[13] In the consumer survey there are five possible responses: "very positive", "positive", "very negative", "negative" and "neutral". In this case p^+ is computed by summing the fraction of very positive answers with the fraction of positive responses. The same steps are applies when computing p^-.

expressed as a share of the "natural" dispersion across the economy (Manteu and Serra, 2017).

The resulting dispersions on each pair of answers for a given time period are then standardised as to have zero mean and unit standard deviation. These standardised "dispersions" from each separate question are then averaged within sectors (such as industry, consumers, *etc.*) to obtain the uncertainty indicators for each sector of the economy. The separate sectoral uncertainty indicators are then averaged in order to obtain the overall uncertainty measure for the whole economy which is the U2 indicator (see Figure 3).

Figure 3. **Uncertainty in Total Economy, Bulgaria** (U2 Indicator)



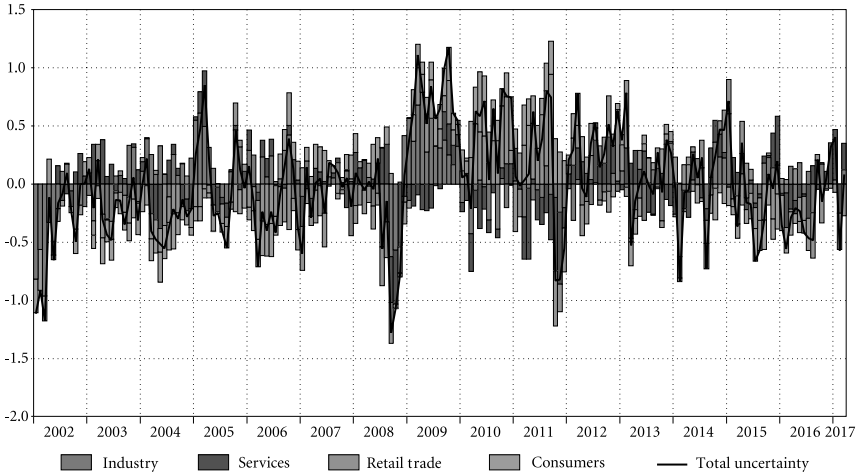
Note: Since the index is standardised a value of zero would mean that the level of uncertainty in the economy is "average", while values above zero are period with heightened uncertainty and values below zero are periods with "below-average" uncertainty. Movements along the Y-axis are measured in standard deviations due to the standardisation transformation.

Source: EU Commission BCS Survey, own calculations

The U2 indicator, like U1, also captures major periods of heightened uncertainty, such as the Great Recession and the EU sovereign debt crisis. In the case of U2 the hikes in uncertainty during these 2 events are relatively weaker as compared to the average variation of the indicator. In general U2 appears to be more volatile than U1, which makes its interpretation and consequently its use for economic analyses more difficult. A possible explanation of why U2 is more volatile than U1 is that the assumption that the economic conditions remain stable between the time of the *ex-ante* and the *ex-post* assessments does not hold, especially in the period after the Great Recession. What stands out is that, unlike

U1, U2 suggests that uncertainty after the 2008/09 global financial crisis has not really returned back to its pre-crisis levels, or to its "average" levels, with events such as the unrests in Ukraine, the negotiations around the Greek debt restructuring, the Brexit referendum and the election of Donald Trump as president of the US contributing for the "above average" uncertainty in recent years.

Figure 4. **Uncertainty in Total Economy, Bulgaria, Sectoral Breakdown** (U2 Indicator)



Source: EU Commission BCS Survey, own calculations

Similar to U1, U2 can also provide the contribution to uncertainty from each sector of the economy because of the way that the indicator is constructed – the overall uncertainty in the economy is calculated as the average uncertainty across each separate sector. The sectoral breakdown, however, gives a somewhat mixed picture of the source of uncertainty, given that sectors frequently switch the direction of their contribution to the overall level of uncertainty, changing between positive and negative contribution over subsequent months (see Figure 4). Services appear to be the most volatile sector, while the industry sector and consumers display more stable results. Despite that U2 manages to capture most of the pre-selected uncertainty enhancing events, the rather muted hike of the indicator during the Great Recession and the large volatility from month to month casts doubt on the reliability of U2 as a good uncertainty proxy, since one would expect that increases/decreases in uncertainty would generally follow a smoother path that reflects the perception of economic agents.

4.3. Uncertainty Indicator U3

The third indicator (U3) is similar to the first one (U1) but it measures the disagreement only related to the future production in the industry sector, rather than the uncertainty in the whole economy (which was the case for U1). While this approach may seem to be too simplistic, there are recognised benefits of using it. The replies on some questions in the BCS (especially those related to expected price developments) may carry ambiguous information and, thus, are likely to introduce additional noise in the uncertainty indicators. What is more, having only one question from which the uncertainty indicator is constructed allows one to analyse the distribution of replies that ultimately determines the dynamics of the uncertainty indicator. These benefits have been widely recognised and this indicator has often been used on its own to measure uncertainty, explaining quite well macroeconomic developments (Bachmann et al., 2013; EC, 2013; Busetti et al., 2015; Meinen and Roehe, 2016). The calculation of U3 starts with obtaining the standard deviation on the responses in the same way as for U1 but this time only from the following question of the BCS survey:

Industry Survey Question 5: How do you expect your production to develop over the next 3 months? It will...

+ increase
= remain unchanged
– decrease

$$\mathbf{U3}_{\mathbf{t}} = \sqrt{\boldsymbol{p}_t^+ + \boldsymbol{p}_t^- - (\boldsymbol{p}_t^+ - \boldsymbol{p}_t^-)^2} \tag{5}$$

where $\boldsymbol{p}^+$ is the fraction of "increase" responses to a survey question at time $\boldsymbol{t}$ and $\boldsymbol{p}^-$ is the fraction of "decrease" responses to a survey question at time $\boldsymbol{t}$.

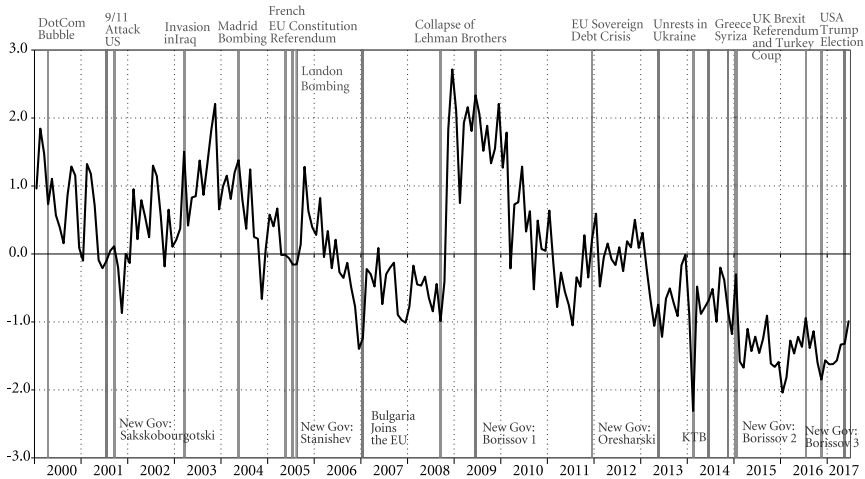
The resulting indicator is then standardised as to have zero mean and unit standard deviation.

The U3 uncertainty indicator captures quite well major periods of global heightened uncertainty, such as the DotCom Bubble, the invasion in Iraq, the Great Recession, the EU sovereign debt crisis and the Brexit Referendum but similar to the previous two metrics (U1 and U2) does not seem to register the pre-selected domestic uncertainty-enhancing events as significant (see Figure 5).

The U3 indicator would imply higher uncertainty in the period around 2003, as compared to U1 and U2. The exceptionally high levels of uncertainty in 2003 can be argued to owe not only to the invasion in Iraq but also to the three largest EU economies (Germany, France, Italy) entering recession. These countries are

the main destinations for exports for the Bulgarian industrial sector and, thus, the worsening of economic activity there increased concerns of Bulgarian firms. This is strongly reflected by U3, since this indicator is constructed only from data from the highly export-oriented industry sector. U3 stands out from U1 and U2 with the much higher amplitude of variation – for example, at its peak after the collapse of Lehman Brothers U3 increases by almost 3 standard deviations, while U1 and U2 at their peaks increased by between 1.0 to 1.5 standard deviations. The U3 indicator also points to a decrease in uncertainty in recent years (similar to U1), but the amplitude of the decrease is much higher than that of U1 (see Figure 5).

Figure 5. **Uncertainty in Industry, Bulgaria** (U3 Indicator)

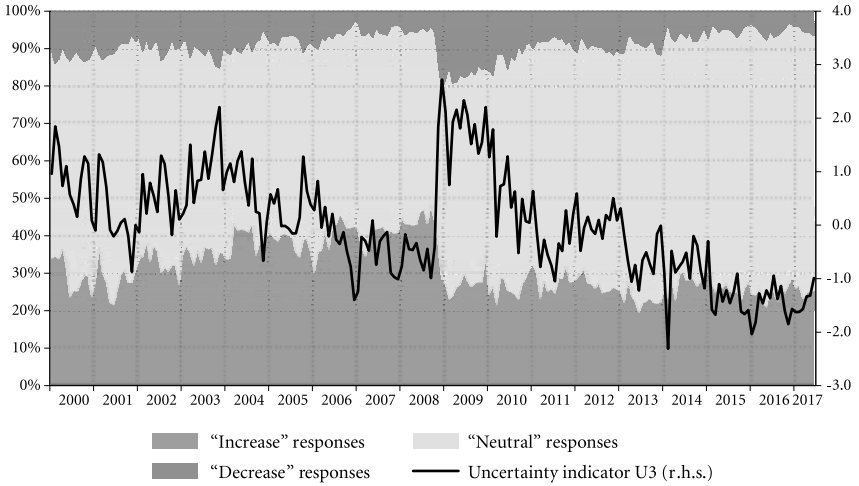


Note: Since the index is standardised a value of zero would mean that the level of uncertainty in the economy is "average", while values above zero are period with heightened uncertainty and values below zero are periods with "below-average" uncertainty. Movements along the Y-axis are measured in standard deviations due to the standardisation transformation.

Source: EU Commission BCS Survey, own calculations

Since this indicator is constructed from the replies on a single question it is easy to plot and visually analyse the distribution of the replies alongside the U3 indicator (see Figure 6).

Figure 6. **Uncertainty in Industry, Bulgaria, Breakdown by Replies** (U3 Indicator)



Source: EU Commission BCS Survey, own calculations

In Figure 6 one can see the underlying volatility of the respondents' answers to the BCS, stemming from the already discussed differences in background characteristics, as well as from disagreements due to different information sets. These monthly volatilities of the survey question replies are inherited by the U3 indicator and cause the indicator to vary from month to month even if no significant event has occurred.

Figure 6 shows that in the period between 2000 and 2008 the fraction of "decrease" replies is quite stable and even falling slightly in the run-up to Bulgaria's accession to the EU. At the same time the fraction of "increase" replies was on an upward path. As a result the fraction of "neutral" responses is decreasing, which, as discussed in the section on U1, leads to an increase in uncertainty because the first terms of the U3 indicator $(p^{+}+p^{-})$ becomes larger. However, if there is clear dominance of one of the replies, like in the period between 2004 and 2007 when the "increase" responses clearly outweigh the "decrease" responses, there is less disagreement among respondents which is a clear sign of lower uncertainty. To account for that the second quadratic term of the U3 indicator $(p^{+}+p^{-})^{2}$ acts as a corrective mechanism and reduces uncertainty (because of the minus sign in front of this second term in the equation). The "above average" uncertainty, during the 2008/09 crisis is characterised by low fraction of "neutral" responses and similar fractions of "increase" and "decrease" responses, which indicates disagreement among respondents and

no clear dominance of either answer. In recent years uncertainty has decreased because the fraction of "increase" responses has dominated the fraction of "decrease" responses, and also the fraction of "neutral" responses has been higher than its historical average.

4.4. Uncertainty Indicator U4

The fourth indicator (**U4**) is constructed in a slightly different way, compared to the previous three indicators. Rather than using the detailed breakdown of "increase" and "decrease" responses from the BCS, this indicator uses dispersion of changes in the overall balance of replies between consecutive rounds of the survey (Girardi and Reuter, 2017). The idea is that a high degree of uncertainty may not be only reflected in disagreement between respondents to a given question but may also be reflected in the resulting balance scores diverging across different questions, *i.e.* increased dispersion across questions rather than within questions (Manteu and Serra, 2017). When uncertainty is low and the economy is growing steadily, respondents' assessment of different questions asked in the survey should be more or less commonly shared and the balances on their replies will not change that much from month to month. The opposite should be true in times of uncertainty. For example, when the economy is exiting from a trough, the dispersion of balance scores is likely to increase as expectations on leading indicators (such as for expected orders) turn positive, while expectations for other lagging indicators, such as employment, stay unchanged or continue to worsen (ECB, 2016:2).

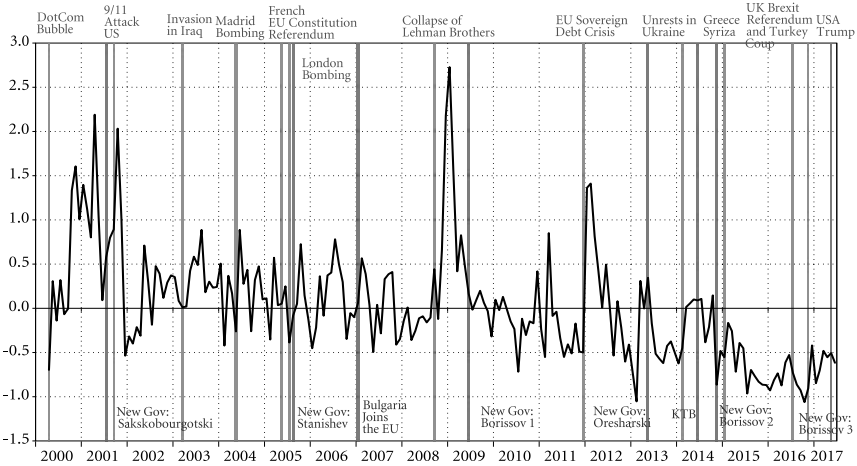
The first step of the U4 indicator computation is to obtain from all questions across all sectors of the survey (5 sectors and a total of 35 questions on past/present/future assessments) the change in each balance score compared to its balance score three months ago:

$$\textbf{Change}_{i,t} = \textbf{Balance Score}_{i,t} - \textbf{Balance Score}_{i,t-3} \qquad (6)$$

where i is each individual question in the survey.

The second step is to standardise the obtained changes on each balance as to have zero mean and unit standard deviation, so that they can be directly comparable. Then the standard deviation across questions' balances changes is calculated within each sector of the economy, as to obtain 5 standard deviations, one for each sector of the economy (industry, services, retail trade, construction and consumers). The sectoral standard deviations are then averaged across sectors in order to obtain the final measure of uncertainty U4 (see Figure 7).

Figure 7. **Uncertainty in Total Economy, Bulgaria** (U4 Indicator)



Note: Since the index is standardised a value of zero would mean that the level of uncertainty in the economy is "average", while values above zero are period with heightened uncertainty and values below zero are periods with "below-average" uncertainty. Movements along the Y-axis are measured in standard deviations due to the standardisation transformation.

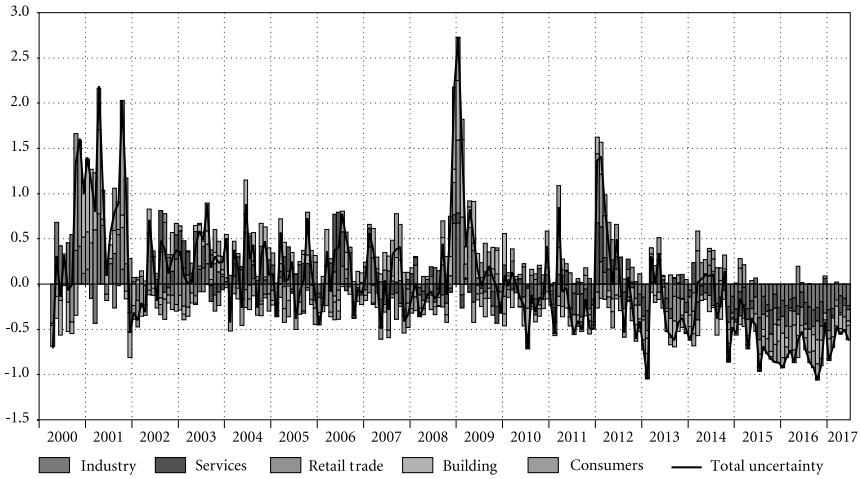
Source: EU Commission BCS Survey, own calculations

The U4 uncertainty indicator performs adequately in terms of capturing major periods of heightened uncertainty, such as the DotCom Bubble, the 9/11 attacks in the US, the invasion in Iraq, the Great Recession, the EU sovereign debt crisis and the unrests in Ukraine. However, similar to the previous uncertainty indicators, U4 does not seem to register domestic uncertainty-enhancing events as significant. Another similarity between U4, U1 and U3 is that they point to a decrease in uncertainty to below average levels in recent years (see Figure 7).

The U4 indicator, similar to U1, U2 and U3 cannot be treated as an unbiased uncertainty indicator but should rather be treated as a proxy for uncertainty. This is because the cross-question perspective captures elements of uncertainty but is also influenced by other forces, mainly heterogeneity. Here, unlike for U1, U2 and U3 the heterogeneity does not refer to the characteristics of the individual respondents, but of the questions (Girardi and Reuter, 2017). According to Girardi and Reuter (2017) the concepts inquired by the different survey questions are heterogeneous in terms of the degree to which they typically lead or lag the business cycle (*e.g.* employment plans, structurally, tend to lag production expectations). As a result, normal business cycle fluctuations can produce increases in the cross-question dispersion of month-on-month changes, which are not strictly related to uncertainty.

The sectoral breakdown of the U4 indicator (see Figure 8) suggests that the decrease in uncertainty observed in the last three years is quite broad-based and is not concentrated in a single sector of the economy.

Figure 8. **Uncertainty in Total Economy, Bulgaria, Sectoral Breakdown** (U4 Indicator)



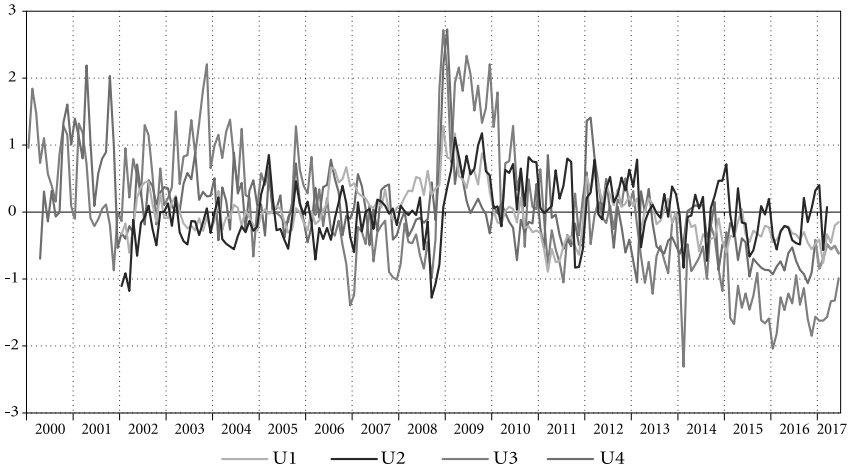
Source: EU Commission BCS Survey, own calculations

5. Combining Different Uncertainty Indicators

Since each of the four calculated indicators (U1, U2, U3, U4) captures a different aspect of the "true" unobserved uncertainty in the economy their dynamics can at times diverge (see Figure 9). While all four indicators spike during the 2008/09 global financial crisis and the 2012 sovereign debt crisis in Europe, suggesting "above average" uncertainty, their dynamics are quite diverse in the period between 2002 and 2005. Such a divergence is also observed in the period between 2015 and 2017, when most indicators suggest that uncertainty has gone down to "below normal" levels, but the magnitude of this decreased uncertainty varies notably among the different indicators. The U3 measure, which measures the uncertainty in the industry sector in Bulgaria, appears to react stronger to uncertainty increasing/decreasing shocks, compared to the rest of the indicators which reflect uncertainty in the total economy and not just the industrial sector. As already commented in Section 4.3 this can be a result of the dependence of the industrial sector on external developments due to its high export-orientation and its high integration in global value chains. On the

other hand, the U2 indicator's readings would not imply a significant decrease in uncertainty in recent years.

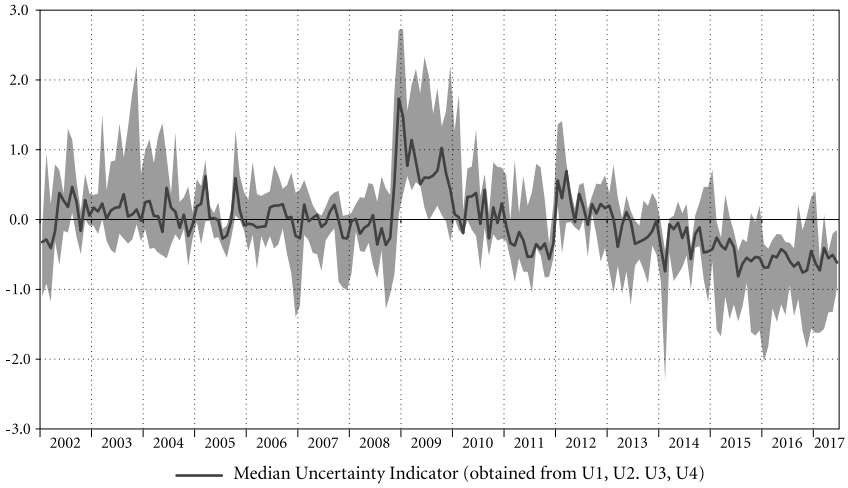
Figure 9. **Comparison of Uncertainty Indicators, Bulgaria** (U1, U2, U3 and U4)



Source: EU Commission BCS Survey, own calculations

Given that no indicator is a perfect proxy for uncertainty, it might be preferable to compile a composite measure of uncertainty which captures and consolidates the information content of the different uncertainty proxies (ECB, 2016:2) and also smooths away the noise inherent to any particular measure. This has been widely used in practice with the method of aggregation varying from simple average of several uncertainty metrics, or weighted average of several uncertainty metrics, to taking the median of several uncertainty metrics, and finally to using a principal components model on a set of several uncertainty metrics and taking out the first principal component as the composite index (Busetti, 2015; ECB, 2016:2; Moore, 2016; Manteu and Serra, 2017). This paper has chosen to use the median aggregation approach, similar to the ECB (2016:2), in order to obtain a composite uncertainty indicator (see Figure 10).

Figure 10. **Composite Uncertainty Indicator, Bulgaria** (Median of U1, U2, U3, U4)



Note: The shaded area around the uncertainty indicator represents the area between the maximum and the minimum standard deviation of all four uncertainty indicators U1, U2, U3 and U4 that are used in order to construct the composite uncertainty indicator.

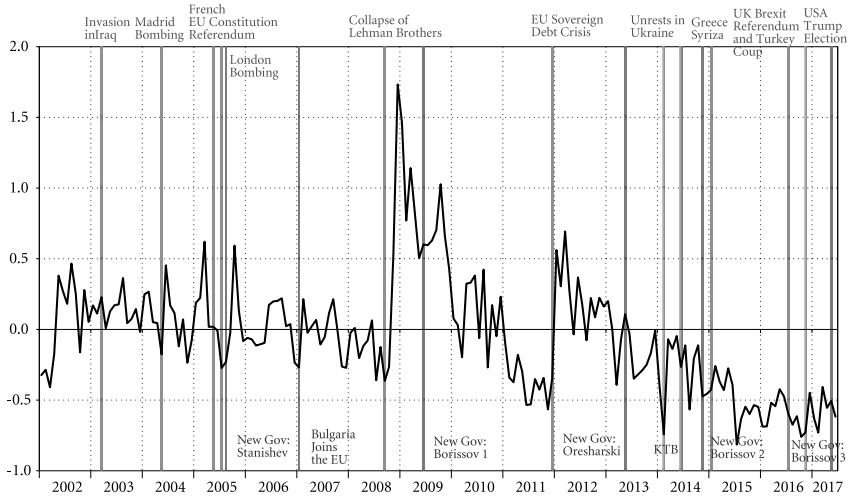
Source: EU Commission BCS Survey, own calculations

The ECB (2016:2) suggests that it is better to use the median rather than the mean when constructing a composite indicator for uncertainty given that the typically found characteristics of uncertainty indicators are not associated with those of a normal distribution. Namely, such indicators display a wide range of observations, which at their extremes can vary up to three standard deviations from the mean. Another observation that the ECB (2016:2) made and which motivates the choice of median over mean is that uncertainty indicators tend to show a positive skewness, which implies that the mass of the distribution of observed uncertainty levels is concentrated at below average levels. In other words, the indicator suggests that there are more frequent and longer-lasting periods of low uncertainty than of high uncertainty. The third characteristic that makes the use of the median more preferable than the use of the mean is that uncertainty indicators typically display high kurtosis, which would imply that the distribution has tails that asymptotically approach zero more slowly than a Gaussian distribution (ECB, 2016:2).

The obtained composite uncertainty indicator for Bulgaria captures the joint development of all underlying uncertainty proxies and it performs adequately in capturing periods that were pre-selected as uncertainty-enhancing, especially

those related to global events, such as the invasion in Iraq, the Great Recession and the sovereign debt crisis in Europe (see Figure 11).

Figure 11. **Composite Uncertainty Indicator, Bulgaria** (Median of U1, U2, U3, U4)



Source: EU Commission BCS Survey, own calculations

The correlations between the constructed uncertainty indicators (presented in Table 1) is generally quite high and statistically significant with the exception of the U2 indicator, which appears to be less correlated with the other uncertainty indicators. This could be due to the more volatile readings of the U2 indicator, compared to the other indicators and also because U2 is the only indicator that does not point to a reduction in uncertainty to "below average" levels in recent years.

Table 1. **Correlation between Uncertainty Indicators**

	U	U1	U2	U3	U4
U	1				
U1	0.8***	1			
U2	0 44***	0.27**	1		
U3	0.84***	0.49***	0.23*	1	
U4	0.82***	0.63***	0.12	0.67***	1

Source: EU Commission BCS Survey, own calculations

An overview of the constructed indicators would suggest that, although similar in nature, the uncertainty proxies can point to different conclusions for the level of uncertainty in the economy, especially for specific time periods. Given that uncertainty is unobserved and there is no formal approach to testing which indicator best reflects the underlying uncertainty in the economy, the composite indicator would be preferred to the separate uncertainty proxies. This is because it captures the joint development of all underlying uncertainty proxies, it smooths out the noise inherent to any particular measure and it also draws information from all sectors of the economy.

However, the constructed uncertainty indicators may be differently related to macroeconomic developments, depending on the aspects of the "true" unobserved uncertainty in the economy that they reflect. For example, the U3 measure, which focuses on future production expectations in the industrial sector, may be more closely related to investment activity in the country as opposed to U1, U2 and U4, which also cover several questions from the consumer survey and would thus be more appropriate when analysing uncertainty at the level of consumers. The next section of the paper examines the relationship between the uncertainty indicators and key macroeconomic variables which has important implications for further work with these indicators and their potential use in modelling activities.

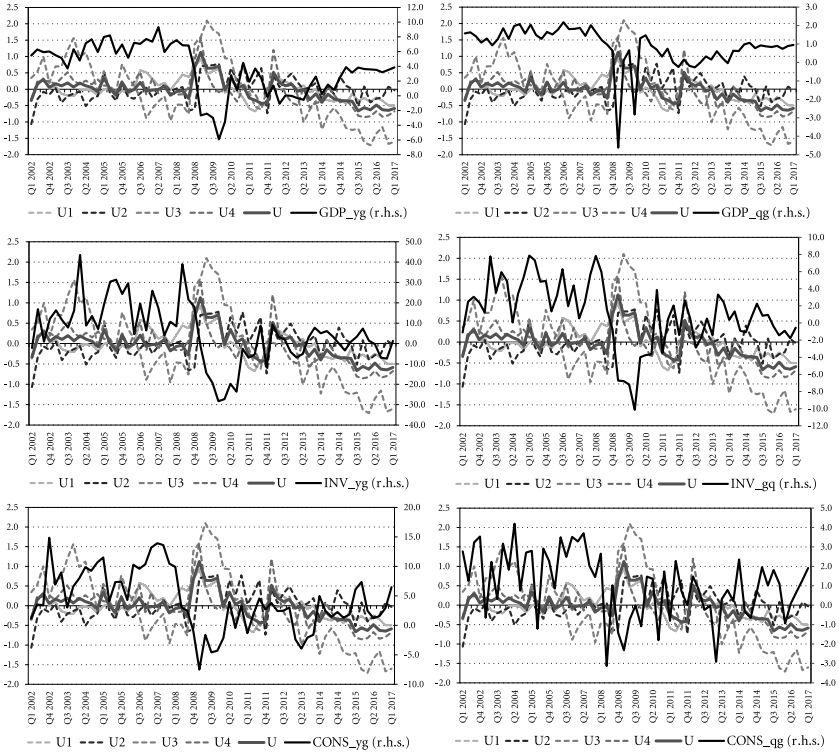
6. Relationship between Uncertainty and Key Macroeconomic Variables

Uncertainty indicators should have negative correlation with key macroeconomic variables, as one would expect an adverse contemporaneous or lagged impact of uncertainty on activity (ECB, 2016:2). In other words, proxies for uncertainty should be high in periods of recession and low during periods of resilient growth. One reason is that negative news shocks (such as terrorist attacks or wars) that can cause recessions also cause higher uncertainty at the same time. Another reason for heightened uncertainty during recessions is that recessions might themselves increase uncertainty (ECB, 2016:2). Figure 12 plots the developments of real economic variables (quarter-on-quarter and year-on-year growth rates in real gross domestic product, real gross fixed capital formation and real private consumption) alongside the constructed uncertainty indicators (U1, U2, U3, U4 and the Composite Uncertainty Indicator U).

A visual inspection of Figure 12 suggests that the constructed uncertainty indicators generally display the desired negative relationship with the developments of GDP, investment and private consumption. However, the strength of the

relationship appears to vary from indicator to indicator, as well as across time. All uncertainty indicators appear to display stronger negative relation with GDP than with investment and consumption, both when quarter-on-quarter and year-on-year growth rates are considered. U3 and U4 are the indicators that display weaker negative relationship with investment and consumption in the period between 2002 and 2005. Also in recent years both indicators suggest that uncertainty has gone down to levels below those in the pre-crisis period, which does not correspond to the more muted developments of the considered macroeconomic variables, which are still lower than their pre-crisis rates. In general, the combined uncertainty indicator U, as well as U1 and U2 appear to capture better the dynamics of GDP, investment and private consumption.

Figure 12. **Uncertainty and Real Economic Developments, Bulgaria**



Note: "qg" stands for quarter-on-quarter growth rate, while "yg" stands for year-on-year growth rate.

Source: NSI, EU Commission BCS Survey, own calculations

However, further analyses are needed in order to identify which uncertainty indicator is most useful in explaining real economic developments. A first step would be to statistically examine the correlations between the constructed uncertainty indicators and key macroeconomic variables.

Section 1 discussed several channels, identified in the economic theory, through which increased uncertainty can affect the economic activity – namely, investment, employment, consumption and increased risk premium. Table 2 presents the correlation of the obtained uncertainty indicators (U1, U2, U3, U4 and the Composite Indicator U) with real gross domestic product (GDP), real gross fixed capital formation (INV), real private consumption (CONS), unemployment rate (UNEMP) and the 5-year credit default swap rate (CDS 5-y; as a proxy for Bulgaria's risk premium) in order to get some intuition about these relationships for Bulgaria. The correlation between the calculated uncertainty indicators and the first-order lags of these macroeconomic variables are also included to partially account for possible delayed effects of uncertainty. All series used in this table have been transformed to have quarterly frequency and the correlations are evaluated for the period between the first quarter of 2002 and the first quarter of 2017.

Table 2 shows that all uncertainty indicators for Bulgaria display negative correlation with the contemporaneous GDP growth rates (both quarter-on-quarter and year-on-year) and in most cases this correlation is statistically significant, which is exactly what is expected by theory. The correlations between the uncertainty indicators and the lagged growth rates of GDP appear to be smaller and less significant but are still negative in almost all cases. One can observe a strong and significant positive correlation between the uncertainty indicators and the risk premium proxy (CDS rates), which also holds when the contemporaneous CDS rate is substituted with its first-order lag. This supports initial expectations that increased economic uncertainty could lead to higher risk premium. The correlation coefficients between the uncertainty indicators and the contemporaneous and lagged real growth rates of investment and private consumption appear to have the expected sign in most cases (with the exception of U4) although the correlation significance is not that strong. Unemployment rate and its first-order lag do not seem to be significantly related to any of the uncertainty indicators, apart from U1, where the correlation coefficient has the wrong sign.

Table 2. **Correlation between Uncertainty Indicators and Key Macroeconomic Variables**

	U1	U2	U3	U4	U
GDP_qg	-0.27**	-0 41***	-0.16	-0.18	-0.33**
GDP_yg	-0.11	-0 44***	-0.2	0.02	-0.22*
INV_qg	-0.12	-0.32**	-0.2	0.08	-0.18
INV_yg	0.11	-0.35***	-0.06	0.23*	-0.01
CONS_qg	-0.17	-0.28**	-0.07	0.03	-0.16
CONS_yg	-0.08	-0.28**	-0.16	0.03	-0.15
UNEMP	-0.24*	-0.19	0.2	0.05	0.01
CDS 5-y	0.24*	0.14	0 41***	0.34***	0.38***
GDP_qg(-1)	-0.01	-0.32**	-0.2	-0.03	-0.12
GDP_yg(-1)	0.04	-0.52***	-0.15	0.14	-0.12
INV_qg(-1)	-0.01	-0.51***	-0.16	0.1	-0.15
INV_yg(-1)	0.16	-0.55***	0.06	0.26**	0.04
CONS_qg(-1)	0.01	-0.17	-0.03	0.02	-0.02
CONS_yg(-1)	0.04	-0.35***	-0.13	0.1	-0.07
UNEMP(-1)	-0.24*	-0.15	0.13	0.06	-0.01
CDS 5-y(-1)	0.24*	0 42***	0 46***	0.3**	0 44***

*Notes: "**qg**" stands for quarter-on-quarter growth rate, while "**yg**" stands for year-on-year growth rate; Statistical significance of the correlation in the table is coded in the following way: *** statistically significant at the 1-percent level, ** statistically significant at the 5-percent level, * statistically significant at the 10-percent level; First-order lags of the macroeconomic variables are denoted by (-1) after the name of the variable in the table.*

Source: EU Commission BCS Survey, NSI, Bloomberg, own calculations

In general, the results from Table 2 suggest that the constructed uncertainty indicators are more strongly related to the year-on-year growth rates of the selected macroeconomic variables, rather than their quarter-on-quarter rates. The U2 indicator appears to have statistically the most significant negative correlation with GDP, investment and consumption, followed by the composite uncertainty indicator U, and U1.

Simple correlations are not an ideal approach to examine the usefulness of the constructed uncertainty indicators for explaining economic developments. The connection between the uncertainty indicators presented and economic activity can be best described with models that explore the mutual interdependence between these variables, without imposing a priori a causal relationship

(Manteu and Serra, 2017). In recent literature Vector Autoregression (VAR) models have been a commonly used tool for this purpose (Meinen and Roehe, 2016; Manteu and Serra, 2017; Girardi and Reuter, 2017), in particular when estimated using Bayesian techniques that reduce the overfitting problems of traditional VAR models. These models could not only quantify the effects of uncertainty on macroeconomic developments but also provide a framework that can be used for forecasting purposes.

However, apart from common factors such as uncertainty, each macroeconomic indicator (GDP, investment, consumption, *etc.*) is also likely to be affected by specific factors. Therefore, it is more meaningful to construct a separate model for each macroeconomic indicator. Such analyses are beyond the scope of this paper but Ivanov et al. (2018) combines the obtained uncertainty indicators with Bayesian VAR estimation techniques to study in details the relationship between uncertainty and investment in Bulgaria. The results from the historical decomposition point to the conclusion that uncertainty has been one of the key drivers of business investment dynamics in the country, while the impulse responses suggest that an increase in uncertainty can depress investment for up to two years. Using similar approaches to analyse the relationship between uncertainty and key macroeconomic indicators such as GDP, private consumption could be a possible area for future research.

7. Conclusion

Uncertainty has been identified as one of the key factors for the sluggish recovery of economic activity in the EU and other advanced economies. While uncertainty is an unobservable and not quantifiable variable, the economic literature offers different ways to measure the change in uncertainty, such as by constructing indicators based on financial market volatility, on survey response disagreements and on word-counts in newspaper articles. Each of these approaches has its pros and cons.

Given uncertainty's potential role as one of the drivers of business cycles, it is important to construct and monitor such indicators of uncertainty (ECB, 2016:2). An assessment of the level of uncertainty is essential to be able to understand and interpret past and current macroeconomic developments. This paper follows the methodology of Girardi and Reuter (2017) in order to construct four uncertainty indicators for Bulgaria, using qualitative data from the European Commission Business and Consumer Survey. Since each of these four indicators captures a slightly different dimension of the "true" underlying uncertainty, the paper also tries to combine them into a single composite uncertainty indicator.

The paper has chosen to use uncertainty indicators, obtained from survey data, because they are easy to calculate, they can be constructed on a monthly basis with data publication delay of only one month and because they have good coverage for Bulgaria (starting from the early 2000s). Another advantage is that due to the harmonisation of the data gathering procedure and methodology in the BCS, the uncertainty indicators that are constructed for Bulgaria are directly comparable to those for other countries from the EU.

The results suggest that the uncertainty indicators for Bulgaria are more sensitive to global uncertainty-enhancing shocks, compared to domestic shocks. The 2008/09 global economic and financial crisis and the 2012 sovereign debt crisis in the EU stand out as one of the clearest examples of periods of heightened uncertainty that all four indicators identify for Bulgaria. Almost all indicators suggest that uncertainty in Bulgaria in the period after 2013 has gone down to below average levels, mainly on the back of the industry sector.

Descriptive analysis shows that the uncertainty indicators for Bulgaria exhibit the desired countercyclical behaviour and are negatively correlated with key macroeconomic variables, such as GDP, investment and consumption and their first-order lags. The strongest relation is between the uncertainty indicators and GDP. There also appears to be a significant co-movement between the uncertainty indicators and the indicator for the risk premium of Bulgaria (proxied by the 5-year credit default swap rate).

One possible area for future research would be to analyse the relationship between these uncertainty indicators and macroeconomic variables in a more detailed manner, for example by incorporating them in structural VAR models and obtaining impulse responses and historical decomposition of the shocks from the model.

Appendices

Appendix I. **Pre-identified Uncertainty-enhancing Events**

Table 2. **Global Uncertainty-enhancing Events**

Event	Occurrence
DotCom bubble burst	April 2000
9/11 terrorist attacks in the US	September 2001
Invasion of US in Iraq	March 2003
Madrid bombings terrorist attack	May 2004
French referendum on ratifying the EU Constitution	May 2005
London bombings terrorist attack	July 2005
Collapse of Lehman Brothers in the US	September 2008
Beginning of the sovereign debt crisis in Europe	December 2011
Unrests in Ukraine after Yanukovych and his government lost power	February 2014
Syriza wins the governmental elections in Greece	January 2015
UK Brexit referendum; Failed military coup against government in Turkey	July 2016
Donald Trump wins US Presidential elections	November 2016

Table 3. **Domestic Uncertainty-enhancing Events**

Event	Occurrence
Election of new government: Sakskoburgotski	July 2001
Election of new government: Stanishev	August 2005
Bulgaria joins the EU	January 2007
Election of new government: Borissov 1	June 2009
Election of new government: Oresharski	May 2013
Placement of KTB bank under special supervision	June 2014
Election of new government: Borissov 2	November 2014
Election of new government: Borissov 3	May 2017

Appendix II. **Questions from the BCS Used to Compute U1**

The U1 uncertainty indicator uses the following questions from the BCS survey for Bulgaria where the three possible answers that respondents can give are "increase", "decrease", or "remain the same" (in the case of consumers the "increase" option consists of two components – "strongly increase" and "increase moderately", and the "decrease" option consists of two components – "strongly decrease" and "decrease moderately"):

Industry survey:

Question 5: How do you expect your production to develop over the next 3 months? It will...

Question 6: How do you expect your selling prices to change over the next 3 months? They will...

Question 7: How do you expect your firm's total employment to change over the next 3 months? It will...

Question 12: How do you expect your export orders to develop over the next 3 months? They will...

Services survey:

Question 3: How do you expect the demand (turnover) for your company's services to change over the next 3 months? It will…

Question 5: How do you expect your firm's total employment to change over the next 3 months? It will...

Question 6: How do you expect the prices you charge to change over the next 3 months? They will...

Consumer survey:

Question 2: How do you expect the financial position of your household to change over the next 12 months? It will...

Question 4: How do you expect the general economic situation in this country to develop over the next 12 months? It will...

Question 6: By comparison with the past 12 months, how do you expect that consumer prices will develop in the next 12 months? They will…

Question 7: How do you expect the number of people unemployed in this country to change over the next 12 months? The number will...

Question 9: Compared to the past 12 months, do you expect to spend more or less money on major purchases (furniture, electrical/electronic devices, *etc.*) over the next 12 months? I will spend…

Retail trade survey:

Question 3: How do you expect your orders placed with suppliers to change over the next 3 months? They will...

Question 4: How do you expect your business activity (sales) to change over the next 3 months? It (They) will...

Question 5: How do you expect your firm's total employment to change over the next 3 months? It will...

Construction survey:

Question 4: How do you expect your firm's total employment to change over the next 3 months? It will...

Question 5: How do you expect the prices you charge to change over the next 3 months? They will...

Appendix III. **Questions from the BCS Used to Compute U2**

The U2 uncertainty indicator uses the following questions from the BCS survey for Bulgaria where the three possible answers that respondents can give are "increase", "decrease", or "remain the same" (in the case of consumers the "increase" option consists of two components – "strongly increase" and "increase moderately", and the "decrease" option consists of two components – "strongly decrease" and "decrease moderately"):

Industry survey:

Question 1: How has your production developed over the past 3 months? It has...

Question 5: How do you expect your production to develop over the next 3 months? It will...

Question 11: How have your orders developed over the past 3 months? They have...

Question 12: How do you expect your export orders to develop over the next 3 months? They will...

Services survey:

Question 2: How has demand (turnover) for your company's services changed over the past 3 months? It has...

Question 3: How do you expect the demand (turnover) for your company's services to change over the next 3 months? It will…

Question 4: How has your firm's total employment changed over the past 3 months? It has...

Question 5: How do you expect your firm's total employment to change over the next 3 months? It will...

Consumer survey:

Question 1: How has the financial situation of your household changed over the last 12 months? It has...

Question 2: How do you expect the financial position of your household to change over the next 12 months? It will...

Question 3: How do you think the general economic situation in the country has changed over the past 12 months? It has...

Question 4: How do you expect the general economic situation in this country to develop over the next 12 months? It will...

Question 5: How do you think that consumer prices have developed over the last 12 months? They have…

Question 6: By comparison with the past 12 months, how do you expect that consumer prices will develop in the next 12 months? They will…

Retail trade survey:

Question 1: How has (have) your business activity (sales) developed over the past 3 months? It has… (They have...)

Question 4: How do you expect your business activity (sales) to change over the next 3 months? It (They) will...

Bibliography

Bachmann, R., S. Elstner, E. Simis (2013). "Uncertainty and economic activity: Evidence from business survey data", American Economic Journal: Macroeconomics 5, 217–249.

Baker, S., N. Bloom, S. Davis (2015). "Measuring Economic Policy Uncertainty", NBER Working Paper Series, Working Paper 21633.

Barkbu, B., S. Berkmen, P. Lukyanstau, S. Saksonovs, H. Schoelermann (2015). "Investment in the Euro Area: Why Has It Been Weak?", IMF Working Paper, WP/15/32.

Bloom, N. (2009). "The impact of uncertainty shocks", Econometrica 77(3), 623–685.

Busetti, F., C. Giordano, G. Zevi (2015). "Main drivers of the recent decline in Italy's non-construction investment", Occasional Papers N. 276, Bank of Italy.

EC (2013). "Assessing the Impact of Uncertainty on Consumption and Investment", Special focus in "Quarterly Report on the Euro Area", Volume 12, N.2 (2013).

EC (2017). "Measuring uncertainty using survey data – what do we measure?", Special topic in "European Business Cycle Indicators 4th Quarter 2016", Technical Paper 013.

ECB (2016:1). "Business investment developments in the euro area since the crisis", Article in ECB Economic Bulletin, Issue 7 2016.

ECB (2016:2). "The impact of uncertainty on activity in the euro area", Article in ECB Economic Bulletin, Issue 8 2016.

Gilchrist, S., J. Sim, E. Zakrajsek (2010). "Uncertainty, Financial Frictions, and Investment Dynamics", 2010 Meeting Papers 1285, Society for Economic Dynamics.

Girardi, A., A. Reuter (2017). "New uncertainty measures for the euro area using survey data", Oxford Economic Papers, Volume 69, Issue 1: 278–300.

Ivanov, E., N. Ivanova, D. Kasabov, K. Karagyozova (2018). "Determinants of Investment in Bulgaria", Discussion Papers DP/110/2018, Bulgarian National Bank.

Ivanova, N., E. Ivanov (2017). "The Role of Bulgaria in Global Value Chains", Discussion Papers DP/105/2017, Bulgarian National Bank.

Knight, F. (1921). "Risk, Uncertainty, and Profit", Boston: Houghton Mifflin.

Manteu, C., S. Serra (2017). "Impact of uncertainty measures on the Portuguese economy", Working Papers 9 (2017), Banco de Portugal.

Meinen, P., O. Roehe (2016). "On measuring uncertainty and its impact on investment: cross-country evidence from the euro area", Discussion Paper No 48/2016, Deutsche Bundesbank.

Moore, A. (2016). "Measuring Economic Uncertainty and Its Effects", Research Discussion Paper RDP 2016-01, Reserve Bank of Australia.

DP/110/2018

Determinants of Investment in Bulgaria

Evgeni Ivanov, Neli Ivanova, Daniel Kasabov,
Kristina Karagyozova

Abstract. Investment activity in Bulgaria, as in many EU countries, contracted substantially during the Great Recession and despite the recently witnessed recovery of business investment remains below pre-crisis levels. The main purpose of this study is to empirically examine the impact of key drivers of business investment in Bulgaria. We estimate a Bayesian VAR model to evaluate the importance of cyclical factors, such as demand, capacity utilisation, economic uncertainty, availability of financing (external and internal) and cost of financing for investment dynamics over the period 2000Q1 – 2017Q1. We establish external demand as the main driver of business investment in the country. Investment dynamics in Bulgaria are also strongly dependent on the levels of economic uncertainty and corporate profits, while foreign financing and lending rates appear to have lower impact. Our empirical results suggest that the observed slowdown of business investment in the post-crisis period can be explained by the reduced external demand and the increased economic uncertainty in the immediate aftermath of the crisis. After 2013 the negative effects from decreased external demand and heightened uncertainty fade away and these two factors contribute positively to investment growth. However, low corporate profits and tightening of financial conditions in the aftermath of the sovereign debt crisis prevented a more notable recovery. After the second half of 2015, our results point to a recovery in corporate profits, as well as favourable financing conditions, which on top of the already observed growth in external demand and decreasing economic uncertainty support the recovery of investment activity in the country. A large part of the fluctuations in business investment in most recent years and especially in 2015 and 2016 that could not be explained by the VAR analysis of demand factors, uncertainty and financing resources, seems to be driven by the dynamics of EU co-financed investment.

Evgeni Ivanov, Economic Research and Forecasting Directorate, Bulgarian National Bank, Ivanov.E@bnbank.org

Neli Ivanova, Economic Research and Forecasting Directorate, Bulgarian National Bank, Ivanova.Nl@bnbank.org

Daniel Kasabov, Economic Research and Forecasting Directorate, Bulgarian National Bank, Kasabov.D@bnbank.org

Kristina Karagyozova, Economic Research and Forecasting Directorate, Bulgarian National Bank, Karagyozova.K@bnbank.org

Acknowledgment:

We would like to thank the members of the Publications Council of the Bulgarian National Bank and particularly Mariella Nenova, Ass. Prof., Ph.D., and Andrey Vassilev, Ph.D., for their valuable contribution, comments and suggestions.

Резюме. Инвестиционната активност в България, както и в редица държави от ЕС, съществено намаля по време на Голямата рецесия и въпреки наблюдаваното напоследък възстановяване на бизнес инвестициите, остава на предкризисни равнища. Основната цел на настоящото изследване е да изучи емпирично влиянието на основните движещи сили на бизнес инвестициите в България. Ние оценяваме Байесов VAR модел за оценка на важността на циклични фактори, като търсене, натовареност на производствените мощности, икономическа несигурност, достъпност на финансовия ресурс (външен и вътрешен) и цена на финансиране за динамиката на инвестициите през периода между първите тримесечия на 2000 г. и 2017 г. Констатираме, че основен двигател на бизнес инвестициите в страната е външното търсене. Освен това динамиката на инвестициите в България силно зависи също от равнището на икономическа несигурност и от корпоративните печалби, докато външното финансиране и лихвените проценти по кредитите имат изглежда по-малко влияние. Получените емпирични резултати показват, че наблюдаваният спад на бизнес инвестициите в следкризисния период може да се обясни с намаляването на външното търсене и със засилилата се икономическка несигурност непосредствено след кризата. След 2013 г. негативните ефекти от намаляването на външното търсене и от високата несигурност отшумяха и тези два фактора се отразиха положително върху растежа на инвестициите. Ниските корпоративни печалби и затягането на финансовите условия след кризата на държавните дългове предотвратиха по-значително възстановяване. След втората половина на 2015 г. нашите резултати показват възстановяване на корпоративните печалби, както и благоприятни условия на финансиране, които непосредствено след наблюдаваното вече нарастване на външното търсене и намаляването на икономичческата несигурност подпомогнаха инвестиционната активност в страната. Значителна чанст от колебанията в бизнес инвестициите в последните години и особено през 2015 г. и 2016 г., които не може да се обяснят чрез VAR анализа на факторите на търсенето, несигурността и финансовите ресурси, изглежда са продукт на динамиката на съвместното финансиране с ЕС.

Contents

1. Introduction .....

2. Literature Review .....

3. Investment Dynamics in Bulgaria .....
   3.1. General Overview of Investment in Bulgaria .....
   3.2. Main Factors behind Investment .....

4. Data and Methodology .....
   4.1. Methodology .....
   4.2. Data .....

5. Results .....

6. Conclusion .....

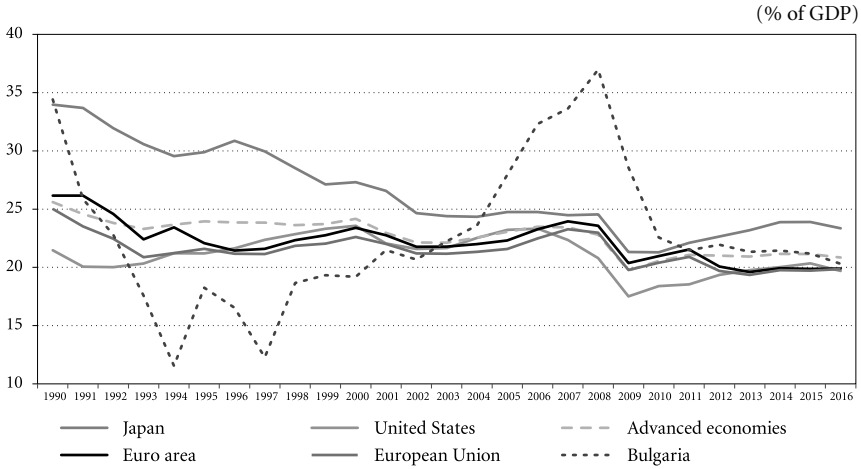
Appendices
   1. Investment Data .....
   2. Survey Results on the Purpose of Investment and Factors Driving Investment Activity .....
   3. Measures of Uncertainty .....
   4. Variables Used .....
   5. EU Funds and Their Role for Government and Corporate Investment .....

Bibliography .....

1. Introduction

Investment plays a key role not only in determining short-run growth prospects through cyclical swings but also in determining the long-run potential growth of the country through capital stock accumulation. Without sufficient investment, the capital stock cannot be renewed regularly, which impedes technological progress and prevents capital deepening from increasing labour productivity (ECB, 2016:1).

Figure 1. **Total Investment as Share of GDP**



Source: IMF WEO April 2017

Investment across countries from the European Union (EU), including Bulgaria, has been hit hard since the onset of the Great Recession in 2009 and the subsequent recovery was slower than in other advanced economies, such as the US and Japan (see Figure 1). While initially after the Great Recession investment in the EU started to recover in line with investment in other advanced economies, the sovereign debt crisis in Europe in 2012 led to a new weakening of investment in the region (EIB, 2016), including in Bulgaria. The fall in EU's total investment witnessed during the Great Recession and the decline in investment in 2011–2012 were a result not only of lower non-construction investment[1], but also a strong downward correction in overheating housing markets and persisting budgetary constraints that resulted in lower construction and public investment in some countries (ECB, 2016:1). As of 2016 total investment

[1] Non-construction investment covers investment in (1) machinery, equipment and weapons systems, (2) intellectual property products and (3) agricultural products.

as a share of GDP remained below its pre-crisis level in the EU even though a slow recovery has been observed in most Member States since early 2013, driven mainly by non-construction investment.

Similar dynamics were observed in Bulgaria. Following a period of strong growth in the period 2002–2008, total investment in Bulgaria decreased sharply in 2009 with both non-construction and construction investment contributing to this decrease. In the following years investment in Bulgaria remained weak, with non-construction investment starting to recover more notably after 2013. At the same time construction investment remained depressed after the real estate boom and bust and is unlikely to regain its pre-crisis levels soon[2].

Recent studies on the topic (Barkbu et al, 2015; ECB, 2016:1; Kose et al., 2017) have identified that the factors that determine the dynamics of non-construction investment are different from those determining the dynamics of construction investment. Moreover, these studies suggest focusing on the so-called "business investment", which typically encompasses certain components of the non-construction investment, and is particularly relevant in boosting an economy's future productive capacity. Relevant studies have identified several factors which have played a key role for the observed business investment weakness in the EU after the 2009 crisis. Among them are weak demand, depressed profit margins, impaired financial sectors and economic and political uncertainty[3].

The aim of this paper is to empirically assess the impact of such cyclical factors on business investment[4] dynamics in Bulgaria and to quantify their contributions, which will shed more light on the determinants behind the business investment slowdown observed in the country after the Great Recession. The estimation is done using a Bayesian Vector Autoregression model (BVAR), based on quarterly data for Bulgaria for the period 2000Q1 – 2017Q1. The paper also attempts to examine in more details the role of EU-related investments for the country's investment dynamics, given that in recent years public investment, co-financed by the EU under different operational programs, has played an important role in Bulgaria, as was the case in most Central and Eastern European countries (EIB, 2016). To the best of our knowledge this paper is the first attempt to disentangle and quantify the drivers behind business investment dynamics for the case of Bulgaria.

[2] See Section 3 for more details.

[3] See Section 2 for a more detailed overview of the main findings in related studies.

[4] We use real investment in machinery, equipment and weapons systems as a proxy for business investment in Bulgaria. For more information on the choice of investment proxy, please refer to Appendix 1 A.

The paper is structured as follows: Section 2 gives an overview of the results of similar studies on the causes and factors behind the investment slowdown, focusing mainly on countries from the EU. Section 3 outlines the historical developments in investment in Bulgaria, offering more descriptive details about its composition and dynamics, which reinforces the choice of model specification and facilitates the interpretation of the empirical results of the paper. Section 4 presents the technical details on the methodology and data used for our analysis and the motivation behind choosing them. Section 5 presents and discusses the results from our analysis. Finally, in Section 6 we conclude by summarising our results.

2. Literature Review

Given the importance of investment for the overall economic growth, the topic of what the drivers of investment are and why the recovery in investment has been weak in most advanced economies in the years after the Great Recession has become one of the central research topics in many international institutions, among which are the IMF (2014), the ECB (2016:1), the World Bank (Vashakmadze et al., 2017) and the EIB (2015). The preferred method used in empirical studies on the determinants of investment typically is structural VAR models, which is also the approach that we have adopted in this paper. Evidence suggests that financial crises have long-lasting effects on investment with a decline of 3 to 3.5 percentage points in the investment-to-GDP ratio three years after the crisis (Barkbu et al., 2015). The recovery in the aftermath of the 2008–2009 financial crisis was particularly slow across most EU countries, including Bulgaria, and the factors behind it have been thoroughly analysed. Some of the identified factors in the literature include weak demand, economic uncertainty, financing conditions, profit expectations and capacity utilisation.

Weak demand and the prospects for sluggish future economic growth have been established as one of the core reasons for the weak post-crisis recovery of investment across Europe (EIB, 2015). The European Commission (2017) has repeatedly identified insufficient demand as one of the most important factors limiting firms' production in its industrial surveys, which is likely to have a negative impact on firm's investment decisions. Furthermore, according to analysis conducted by the IMF (2014), output changes can explain trends in investments activity in most euro area (EA) member states. This was further supported by an empirical study on the case of Spain which empirically proves, using an accelerator model, that output almost entirely explains the dynamics of investment in the country (Barkbu et al, 2015). Busetti et al. (2015) examine the causes of exceptionally low investment in Italy since 2007 and find that demand

conditions, along with cost of capital, uncertainty and credit supply have a statistically significant impact on firms' capital spending. External demand, in particular, has been found to play a dominant role in explaining investment dynamics in small open economies (CNB, 2017). Kose et al. (2017) explore the evolution of investment growth in emerging market and developing economies (EMDE), using a structural VAR model. They conclude that the investment slowdown in EMDE could be attributed to a large extent to the weak economic activity in major economies.

Another important factor for investment dynamics that has been established in the existing literature is economic uncertainty. In general increased uncertainty is considered to lead to a postponement of investment (Busetti et al., 2015). As Bloom et al. (2007) present, heightened uncertainty, apart from affecting directly investment, also reduces the responsiveness of investment to a demand increase, which reduces the effects of policy stimulus. There are many other studies that examine the potentially adverse impact of uncertainty on investment. A recent study by Meinen and Rohe (2016) develops five types of uncertainty measures to investigate the role of uncertainty for investment decisions with a focus on Germany, France, Italy and Spain. Based on both descriptive analysis and VAR models, the authors conclude that investment responds negatively to uncertainty shocks and that uncertainty can account for a significant portion of the decrease in investment in the course of the Great Recession. A report published by the EIB (2013) proves that uncertainty, together with demand, have been the main drivers of the investment slowdown in the euro area since 2010. A study by the European Commission (2013) concludes that the increase in uncertainty has a significant negative effect both on investment and on consumption. Busetti et al. (2015) confirm that uncertainty had a negative impact on investment during the Great Recession and establish it as one of the factors explaining the delayed recovery of the economy.

Credit constraints can also have a sizeable negative impact on investment since they limit the availability of resources that firms could use to expand their activities. However, studies for the euro area find that financing conditions have been a constraining factor only in certain countries. Barkbu et al. (2015) establish a negative correlation between financial constraints and investment for Italy, Portugal and Spain. In addition, applying an extended version of an accelerator model, Busetti et al. (2015) estimate that for the case of Italy constraints imposed by tight credit conditions accounted for about one-third of the fall in investment in the periods 2008–2009 and 2012–2013. The ECB (2016:1) uses a VAR model and finds that credit constraints have been one of the main factors supressing the euro area's business investment in the years after the sovereign debt crisis.

Earned profit is another factor that has been found to play a role in firms' decisions to invest. Sustaining profits means that firms can invest without having to borrow funds from credit markets and, thus, could easily and more quickly expand their production if they assess that there is enough demand for their products. Profitability is also a positive signal which may lead to further investments in case returns are high enough to compensate for the risk and the cost of capital (EIB, 2016). The ECB (2016:1) finds that the current business investment recovery in the euro area is being driven mostly by profit growth, improving demand and falling interest rates. Mizen and Vermeulen (2005) confirm that investment in Germany and the UK is sensitive to sales growth and operating profits.

Capacity utilisation could also be considered as a factor that can determine investment activity (ECB, 2016:1, Busetti et al., 2015), since it reflects firms' choices on how to meet current and expected demand. Capacity utilisation can give an indication of the cyclical position of the economy. For example, in the aftermath of the Great Recession demand was greatly reduced and as a result capacity utilisation decreased, which in turn made firms' new investments in production extension unjustified. On the other hand, in times of economic expansion when demand is booming and capacity utilisation is rising, firms are under pressure to invest more in order to extend or improve their production capabilities.

In addition to traditional borrowed and own financial resources, transfers from the EU have also played an important role in determining investment dynamics, especially in new member state countries (EC, 2016:2). EU co-financed investment is unevenly spread across the years and as such affects significantly not only the level, but also the growth rate of economy wide investment (Eesti Pank, 2016). In some member states, EU transfers have played a countercyclical role in the aftermath of the crisis, offsetting declines in other investment (EC, 2016:2 and Eesti Pank, 2016). For other member states, positive contributions from EU co-financed investment were concentrated mostly in 2014–2015 and apart from their impact on government investment they have also played a prominent role in explaining the dynamics of investment by non-financial corporations (CNB, 2016). In 2016 EU funds absorption declined at the switchover between the programing periods, having an adverse impact on investment growth in all Central and Eastern European countries (NBP, 2017).

Having already identified the main cyclical factors proposed in the economic literature as key drivers of investment dynamics, it is worth discussing in more details how investment in Bulgaria has evolved over the years and which of the outlined factors have been in play.

3. Investment Dynamics in Bulgaria

This section analyses the investment dynamics in Bulgaria over the period 2000–2016 and aims at identifying the main drivers of investment in Bulgaria before, during and after the global financial crisis, outlining the possible factors behind the slow recovery of investment after 2009. The main findings are used as an input for section 5 where we quantify the impact of the identified factors on investment activity through the lens of a Bayesian VAR model.

3.1. General Overview of Investment in Bulgaria

Gross fixed capital formation in Bulgaria accounts for more than 20% of GDP. As such it plays a key role not only in determining short-run growth prospects through cyclical swings but also in determining the long-run potential growth and the productive capacity of the economy.

Following the introduction of the currency board arrangement in July 1997 and the achieved macroeconomic stabilisation, Bulgaria had significant investment needs. This reflected the low initial level of the capital stock (relative to the size of the economy) compared with other new member state countries (NMS)[5] and the high rate of depreciation of the existing capital stock (see Table 1). In addition to this, a mix of cyclical factors that we will discuss here has also contributed to developments of investment in the country.

Table 1. **Indicators Related to Investment Activity in Bulgaria**

Indicators	2000–2003	2004–2008	2009–2011	2012–2016
Investment (% of GDP)	18.8%	27.2%	23.7%	20.7%
Nomianl interest rates on new business, non-fnancial corporations	10.9%	8.5%	8.7%	7.0%
Real (*ex-post*) interest rates on new business, non-fnancial corporations	6.4%	1.7%	5.6%	7.8%
Capital stock-to-GDP ratio	1.9	1.8	2.1	2.2
Capital stock-to-GDP ratio (new member states)*	*2.0*	*2.0*	*2.2*	*2.3*
Capital intensity (net capital stock at 2010 prices per person employed)	30.04	32.98	42.78	51.19
Consumption of fixed capital (% of previous period net capital stock)	8.1%	9.1%	7.5%	6.7%
Consumption of fixed capital (% of GDP)	14.4%	15.1%	14.8%	14.4%
Marginal efficiency of capital (change in GDP at constant market prices of year T *per* unit of gross fixed capital formation at constant prices of year T-5)	0.28	0.26	0.01	0.09
Net returns on net capital stock (2010 = 100)	104.68	127.59	103.57	84.87

**New member state countries (NMS) are Czech Republic, Estonia, Croatia, Latvia, Lithuania, Hungary, Poland, Romania, Slovenia and Slovakia.*

Source: European Commission (AMECO database), BNB

[5] Here we include the countries that joined the EU after 2004 and that have similar economic and historical development to Bulgaria. These countries are Czech Republic, Estonia, Croatia, Latvia, Lithuania, Hungary, Poland, Romania, Slovenia and Slovakia.

Based on its dynamics, investment activity in Bulgaria could be divided into three main periods. The first period spans from 2000 to 2008 and is characterised by an investment boom. The share of investment during that period increased substantially from a low of 16.8% of GDP in 2000 to a peak of 33.0% in 2008. This period was characterised by favourable cyclical developments in Bulgaria and in our main trading partners from the EU, the construction boom in Bulgaria over the period 2004–2008 and a decline in investors' risk perceptions. The second period covers the post-crisis years (2009–2011) and is characterised by a decline in investment. The cumulative decline in real gross fixed capital formation for that period was around 35%. The observed weak recovery of investment activity in 2012–2016 can be distinguished as the third period. During this period the growth rate of real gross fixed capital formation in Bulgaria averaged 0.8% per year for a total increase of 4.1%, with government investment playing an important role. The EU funds programming cycle has had a notable effect on investment in this third period, with EU funds supporting investment activity in the period 2012–2015 and leading to a temporary decline in 2016.

We can generally distinguish between two broad categories of investment – construction investment and non-construction investment. Examining the breakdown of gross fixed capital formation by asset type could shed more light on how the identified factors have affected investment activity over the post-crisis period and which components have driven the observed investment recovery in recent years. Data suggests that the decline in investment after the global financial crisis was the consequence both of a strong downward adjustment in the construction sector and of lower non-construction investment activity (see Figures 2 and 4). Eight years after the crisis gross fixed capital formation in construction, both residential and non-residential, is still around 40% below its pre-crisis peak level. Although gross fixed capital formation in assets other than construction has recovered somewhat recently, it is still more than 20% below the peak reached in 2008 (see Figure 4).

Figure 2. **Breakdown of Real Total Investment**

(annual percentage changes, percentage points)

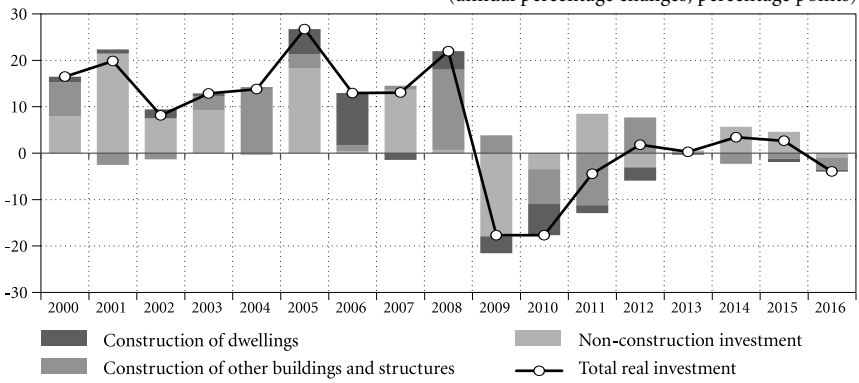
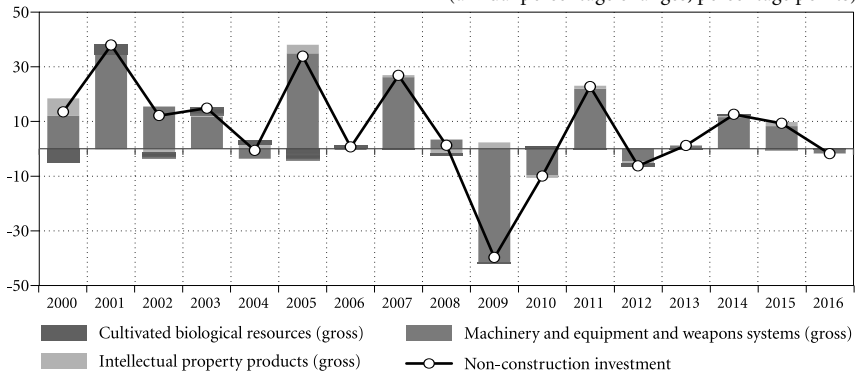


Figure 3. **Non-construction Investment Growth Rate and Contribution by Major Asset Type**

(annual percentage changes, percentage points)



Source: Eurostat

The decline in construction investment after the global financial crisis reflected mainly a prolonged period of housing market adjustment to the changed economic environment in Bulgaria. The heightened economic uncertainty, negative labour market developments, increased saving rate, negligible FDI inflow to the residential sector and tightened financial conditions suppressed household demand for dwellings, resulting in a prolonged period of decline in house prices. Starting from the second half of 2008, house prices tended to decline strongly and by the end of 2011 they were already more than 30% lower than the pre-crisis levels. This dampened investment in the construction sector (BNB, 2017).

Apart from housing, the depressed economic activity reduced the need for new non-residential buildings and also forced the public sector to cut back on infrastructure expenditures, especially in the period 2010–2011.

Figure 4. **Real Investment Compared with the Pre-crisis Peak Levels**

(index, 2008 = 100)

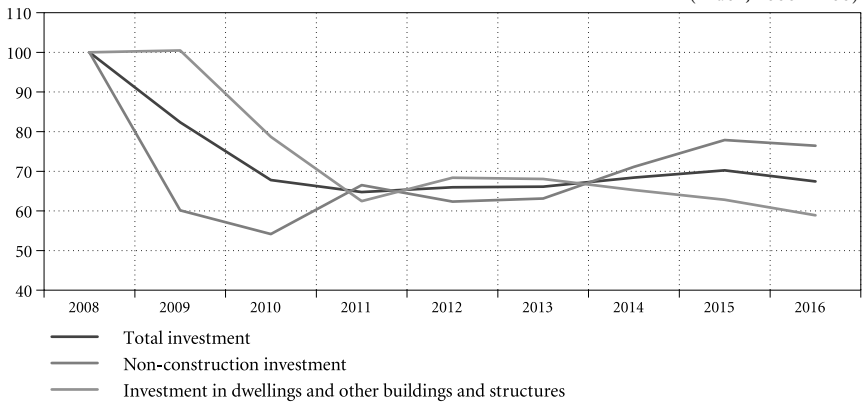
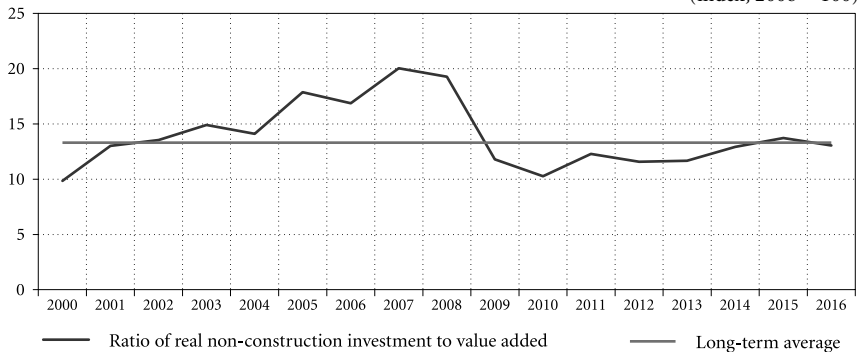


Figure 5. **Ratio of Real Non-construction Investment to Value Added**

(index, 2008 = 100)



Note: Long-term average is calculated for the period 1998–2016.

Source: Eurostat

Non-construction investment declined markedly in period 2009–2010 and started to recover afterwards, though at much slower growth rates compared with the pre-crisis period. Currently, the level of non-construction investment is significantly below the pre-crisis peak observed in 2008 (see Figure 4). Figure 5 indicates that in the run-up to the global financial crisis there was some "over-investment" in Bulgaria, measured by the ratio of real non-construction invest-

ment to value added[6], which possibly reflected favourable cyclical conditions, high marginal return on capital and positive firms' expectations about future demand, lower cost of finance and large FDI flows. In the aftermath of the global financial crisis both the level and growth rate of non-construction investment fell by much more than the value added. Following the cyclical recovery of the economy, however, the ratio of real non-construction investment to value added started to rise and now it is close to its long-run average[7]. In the years after the 2008/09 crisis, non-construction investment, which in Bulgaria consists almost entirely of investment in "machinery, equipment and weapon systems", has become the main driver of gross fixed capital formation (see Figures 3 and 4).

Having briefly discussed how investment in Bulgaria has evolved over the years we now turn to a more detailed discussion on the likely significance of key factors that drive investment decisions in Bulgaria.

3.2. Main Factors behind Investment

As already discussed in Section 2, according to the economic literature the level of firms' desired capital stock is determined mostly by demand conditions, expectations of return on capital, the cost of financing, availability of funding and economic uncertainty. In the case of Bulgaria, EU funds are also likely to play a role for investment dynamics. The detailed elaboration of the potential effects of the identified factors on investment dynamics in Bulgaria would be useful in order to interpret the results from our VAR model in Section 5.

Demand Conditions

Demand for investment in Bulgaria was strong in the post-1997 period taking into account the low initial level of the capital stock and the high rate of depreciation of the existing capital stock (see Table 1). Since mid-1997 the Bulgarian economy began to open up, with the exports of goods and services-to-GDP ratio reaching 64% in 2016, up from 36% in 2000. The EU became Bulgaria's most important trading partner and the trade and financial linkages between them deepened further as a result of Bulgaria's accession in the EU. Apart from trade, this process was also supported by the observed at the time large FDI inflows from EU countries in the financial sector and in the export-oriented manufac-

[6] The ratio of real non-construction investment to value added gives a measure of the size of investment in the overall economy. As this ratio tends to be pro-cyclical over time, comparing it with the long-term value gives an indication of the cyclical behaviour of investment and the presence of over/under-investment (ECB, 2016:1).

[7] Long-run average is calculated for the period 1998–2016. In addition to the business cycle, the ratio of real business investment to gross value added reflects factors such as the capital intensity of the economy (see Table 1) and structural changes over time.

turing sector, which helped facilitate the integration of Bulgaria in Global Value Chains. As a result, investment activity in Bulgaria became more and more closely related to the cyclical developments in the EU, and, respectively, to the external demand for Bulgarian goods and services (see Figure 6).

Figure 6. **Non-construction Investment and External Demand for Bulgarian Exports (at constant prices)**

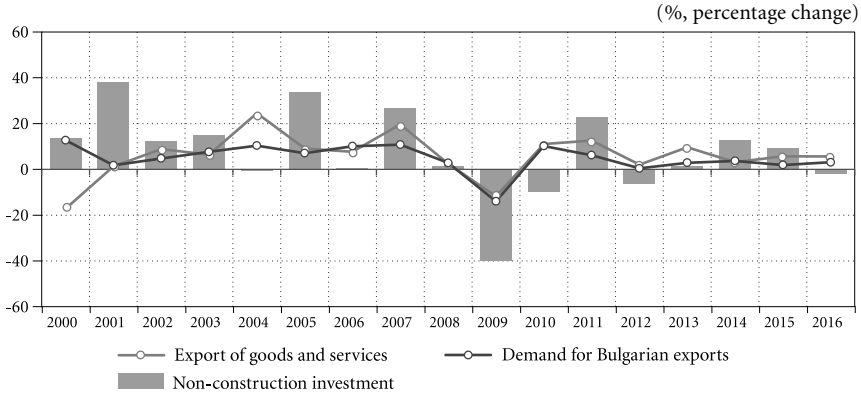
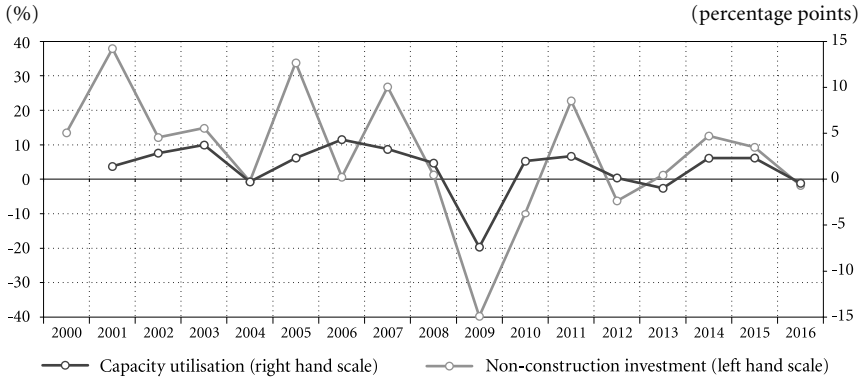


Figure 7. **Change in Capacity Utilisation in the Manufacturing Sector and Growth Rate of Real Non-construction Investment**



Note: Non-construction investment in Figures 6 and 7 is defined as overall investment excluding construction and is used as a proxy for investment in manufacturing.

Source: ECB, NSI

Over the period 2000–2008 demand for investment was stimulated by favourable cyclical conditions in both Bulgaria and our main trading partners from the EU. On the one hand, firms were faced with strong external and domestic

demand at a time of high capacity utilisation, which pushed up non-construction investment (see Figure 7). On the other hand, the rising demand for new dwellings and non-residential buildings, stimulated by the positive labour markets developments, the easing financial conditions and FDI inflows in real estate activities, supported investment in the construction sector (Kotseva and Yanchev, 2017).

Survey data reveals that during the pre-crisis period more than 40% of firms in the manufacturing sector tended to report demand as the most important positive driver of investment[8]. As a result, the share of investment serving extension purpose increased from 32% in 2002 to 50% in 2008. The increase in capacity utilisation observed at that time in the manufacturing sector was another factor which necessitated further expansions of firms' capital stock (see Figure 7).

During the post-crisis period (2009–2011) investment in Bulgaria was pushed down by cyclical conditions as the economy (especially the construction sector) underwent large downward economic adjustments, accompanied by a gradual deleveraging of the private sector. Due to a prolonged depressed domestic demand, investments were aimed mostly at rationalisation rather than expansion of productive capacity. At the same time, given the importance of exports to the EU for the total growth of the Bulgarian economy, weak external demand over the period 2009–2011 amid the sovereign debt crisis in Europe amplified negative domestic developments. The deterioration of demand conditions created significant production capacity under-utilisation during the post-crisis period which limited the need for further increases in gross fixed capital formation. With the strengthening of economic activity in Bulgaria and the EU in recent years, investment activity has started to recover and since 2014 the share of firms that undertake investment in order to expand productive capacity has been increasing. Data on the capacity utilisation rate in the manufacturing sector also suggests increasing needs for further investment.

Economic Uncertainty

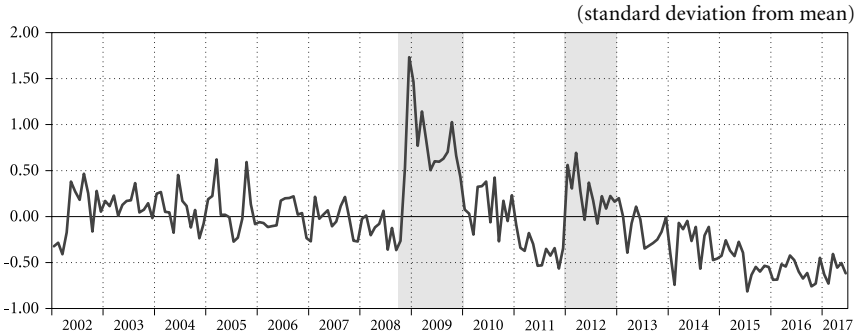
Economic theory suggests that uncertainty has a detrimental effect on economic activity by giving economic agents the incentive to postpone investment, consumption and employment decisions until uncertainty is resolved, and by pushing up the cost of capital through increased risk premia (EC, 2013).

In this paper we use five uncertainty indicators, constructed by Ivanov (2018), following the methodology of Girardi and Reuter (2017). These indicators are based on monthly qualitative survey data, taken from the European Commis-

[8] See Appendix 2 "Survey results on the purpose of investment and factors, driving investment activity" for more information.

sion's Business and Consumer Survey. The core underlying assumption is that the more economic agents disagree in their assessment and expectations about current and future economic activity, the higher is the uncertainty in the economy[9]. In this section we present only the composite uncertainty indicator for Bulgaria, which consolidates the information of the different uncertainty indicators (see Figure 8) but for the purposes of the VAR model we use all five uncertainty indicators before selecting the most appropriate one for the empirical estimation.

Figure 8. **Composite Uncertainty Indicator, Bulgaria**



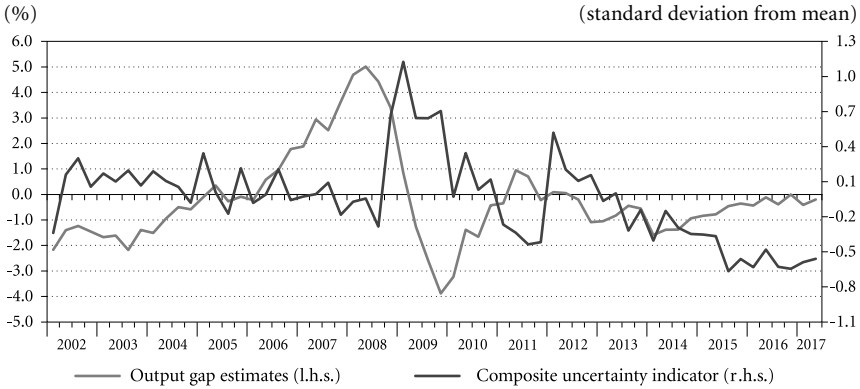
Note: Since the index is standardised a value of zero would mean that the level of uncertainty in the economy is "average", while values above zero indicate periods with heightened uncertainty and values below zero are associated with periods with "below-average" uncertainty. Movements along the Y-axis are measured in standard deviations due to the standardisation transformation.

Source: EU Commission BCS Survey, own calculations

Based on this indicator we can see that the global financial crisis and sovereign debt crisis triggered an increase in uncertainty in Bulgaria (indicated by the shaded area in Figure 8), likely depressing economic activity and investment. In recent years (2013–2017) uncertainty has declined in line with the strengthening of economic activity in Bulgaria and our trading partners, which is likely to have had a positive effect on investment dynamics.

[9] For more information on the construction and specificities of these indicators please refer to Appendix 3.

Figure 9. **Uncertainty and the Business Cycle in Bulgaria**



Notes: "Output gap estimates" refers to our assessment of the output gap in Bulgaria, estimated using a production function approach on quarterly data. The composite uncertainty indicator is also transformed to have quarterly frequency in order to be comparable to the output gap estimates.

Source: NSI, EU Commission BCS Survey, own calculations

As expected, the uncertainty index in Bulgaria moves in the opposite direction of the business cycle in the country (see Figure 9). One reason is that negative news shocks (such as terrorist attacks or wars) that can trigger recessions also cause higher uncertainty. Another reason for heightened uncertainty during recessions is that recessions might themselves increase uncertainty (ECB, 2016:2).

Cost of Financing and Availability of Funding

The investment boom in Bulgaria over the period 2000–2008 was sustained by large FDI inflows (see Figure 10). FDI inflows followed an upward trend from 5.6% of GDP in 2002 to a peak of 27.9% in 2007 and 18.1% in 2008. Most of the FDI inflows were absorbed by the construction, real estate, financial, manufacturing, trade and transport sectors. On average, around half of the FDI inflows in Bulgaria were used to finance corporate investment (mostly manufacturing and development of domestic trade infrastructure), one fifth were allocated to the banking sector (stimulating domestic credit), while the remainder went to the corporate real estate sector – largely financing the building of vacation real estate aimed at foreigners (Mitra, 2011). Rapid credit growth in the pre-crisis period and declining interest rates further supported accumulation of capital (see Table 1 and Figure 11). This reflected sound macroeconomic fundamentals (successful macroeconomic stabilisation, restoration of confidence in the bank-

ing sector and strong GDP growth), capital inflows and increased competition in the banking sector following the privatisation of state banks.

Figure 10. **FDI Inflows by Sector of Economic Activity**

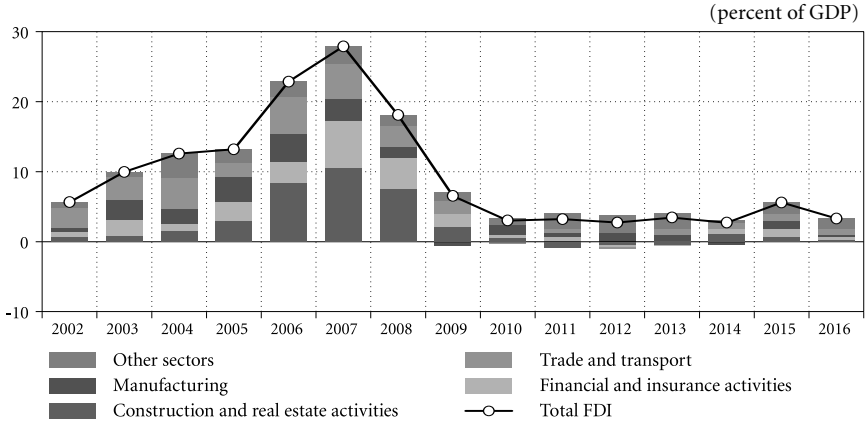
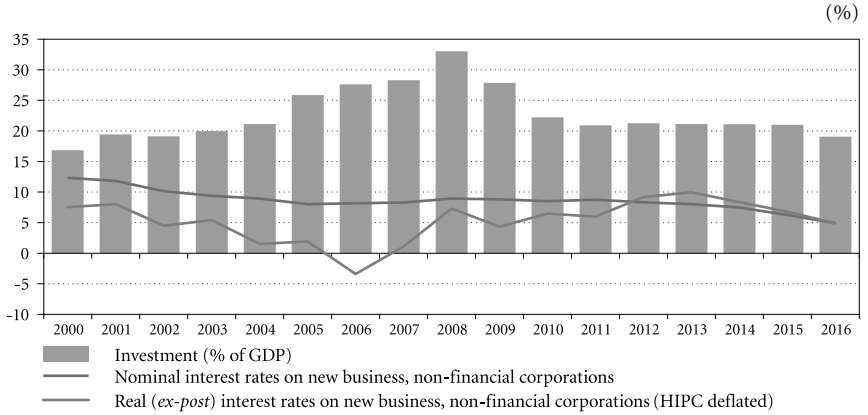


Figure 11. **Investment and Cost of Financing**



Source: European Commission (AMECO database), BNB

The global financial crisis and sovereign debt crisis increased economic uncertainty across Europe and FDI inflows decelerated markedly due to investors' increased risk-aversion. As a result one of the financing channels that Bulgarian firms had at their disposal in the pre-crisis period was substantially reduced. At the same time, real financing costs for firms (proxied by real (ex-post) interest rates) increased in the post-crisis period (see Table 1) as spreads rose and banks experienced problems with deteriorating portfolios. Thus, bank lending

conditions deteriorated and credit-supply shocks limited investment activity. According to the Bulgarian National Bank's quarterly bank lending survey, banks tightened credit standards markedly over the period from the third quarter of 2008 to the first quarter of 2010 (Karamisheva, 2016). Since the beginning of 2010 an easing of credit standards has been observed. In recent years nominal and real costs of external financing have trended downwards and have become increasingly supportive of business investment due mostly to the expansionary monetary measures in the euro area. With the recovery of bank lending, more firms have been able to take advantage of low interest rates to invest more.

Corporate Profits and Expectations of Returns on Capital

Other factors that determine investment decisions are the dynamics of corporate profits and expectations of returns on invested capital. The levels of corporate profits and retained earnings influence firms' capacity to finance investment with internal means. During the pre-crisis period gross investment rate tended to increase in line with the upward trend of gross profit share[10] (see Figure 12 or net returns on capital stock in Table 1). This profitability-type indicator corresponds to the share of the value added created during the production process remunerating capital and could be viewed as an approximation of operating profit before depreciation. The increase in gross profit share for the overall economy reflected the low initial level of the capital stock and thus the high marginal efficiency of capital (see Table 1). Declining risk perceptions and expectations of higher future returns based on the EU membership and prospects of euro adoption (IMF, 2007) further stimulated investment activity.

After the global financial crisis firms' profitability, measured by gross profit share, deteriorated, likely having a negative effect on gross fixed capital formation. Gross profit share dynamics was a consequence of successive years of lacklustre GDP growth rate and a prolonged period of housing market adjustment. In addition to the business cycle, structural factors also contributed to the weakening of investment activity. The post-crisis period was marked by a substantial slowdown in TFP growth and potential output growth, which were not expected prior to the crisis (see Figure 13). As a result economic agents revised corporate profit growth expectations downwards compared with the pre-crisis period, thereby depressing domestic and foreign investment.

[10] Gross profit share is a measure of profitability and is calculated as the ratio of gross operating surplus to value added.

Figure 12. **Gross Investment Rate and Profit Share (Total Economy)**

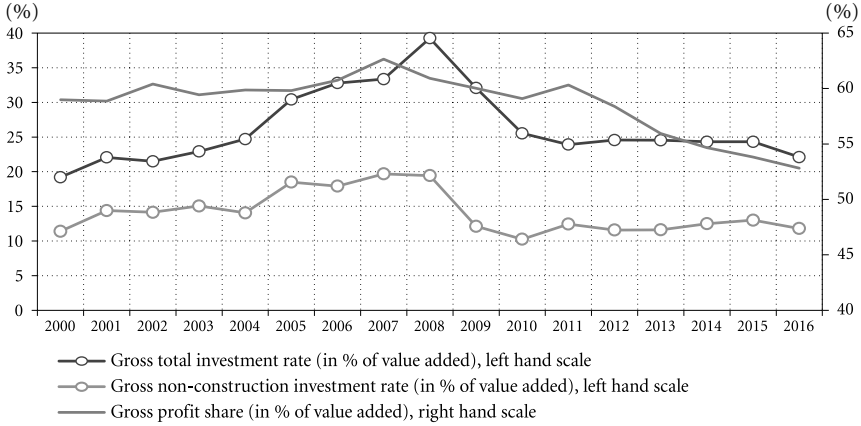
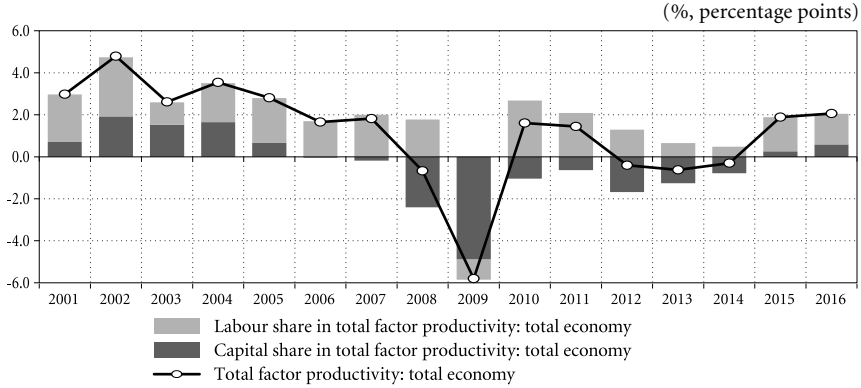


Figure 13. **Breakdown of Total Factor Productivity (TFP) Growth Rate**



Notes: (i) Non-construction investment in Figure 12 is defined as overall investment excluding construction; (ii) TFP decomposition in Figure 13 is based on estimates of the European Commission. According to the methodology used, potential output is estimated by using a Cobb-Douglas production function, and unlike standard approaches it explicitly accounts for the level of efficiency with which inputs (capital and labour) are being used in the production process. As a result, TFP is explicitly modelled as a function of the level of efficiency of factor inputs (Havik, K. et al., 2014).

Source: NSI, European Commission (AMECO database)

Impact of EU Funds on Public and Private Investment

The programming cycle of EU funds is another factor that influences investment activity since the country joined the EU in 2007[11]. Similar to other NMS countries, the absorption of EU funds in Bulgaria has probably played an important role in explaining recent developments in business investment that could not be explained by demand factors, uncertainty or financing resources.

The largest part of the EU structural and investment funds in Bulgaria under both 2007–2013 and 2014–2020 financial frameworks have been allocated in support of government investment in transport and environmental projects[12]. Some of the infrastructure projects have also been undertaken by state owned companies that are classified outside the general government sector (a notable example is the expansion of the underground railway in Sofia with financial support from the EU and the national budget amounting to BGN 1 billion or 8% of the 2007–2013 ERDF and Cohesion Fund allocations)[13]. EU transfers to private companies have remained more limited, even though their role has increased in the 2014–2020 programming period as compared to the 2007–2013 financial framework. As an alternative to grant funding, the EU-wide priorities for the current programming period have focussed more on various financial instruments, which mobilise banks' financing and private capital to support investment projects. The role of these instruments for stimulating productive private investment has been strengthened by the additional mechanisms, available under the Investment plan for Europe, launched in 2014[14].

Given the financing priorities of the EU cohesion policy, one could expect three main channels through which EU co-financed investment could stimulate business investment: 1) providing direct financial support for individual investment projects in non-financial corporations, 2) creating favourable conditions for business investment through better infrastructure, technologies and enhanced human capital as a result of higher government investment and 3) lowering the cost of bank lending for business investment projects as a result of either explicit guarantees through various financial instruments in place or lower risk profile of EU co-financed investment projects.

There are, however, concerns that the stimulating impact of EU funds might remain limited if EU co-financed investment does not add to but rather replaces

[11] Before joining the EU in 2007 Bulgaria had access to pre-accession EU funds. For further details of the pre-accession EU funds please refer to Paliova and Lybek (2014).

[12] Details on the allocations and financing priorities in the 2007–2013 and 2014–2020 programming periods can be found in MF (2016).

[13] Source: Unified management system for the EU structural instruments in Bulgaria, www.eufunds.bg.

[14] See Appendix 5 on detailed information for EU funds allocations and financial priorities.

market-based investment. Moreover, if national government spending on co-financing EU projects increases considerably and this translates into tax or government debt increases, real interest rates may also increase, resulting in private investment crowding-out. Last but not least, if EU grants are not spent efficiently, their impact on business investment might be rather limited and short-lived.

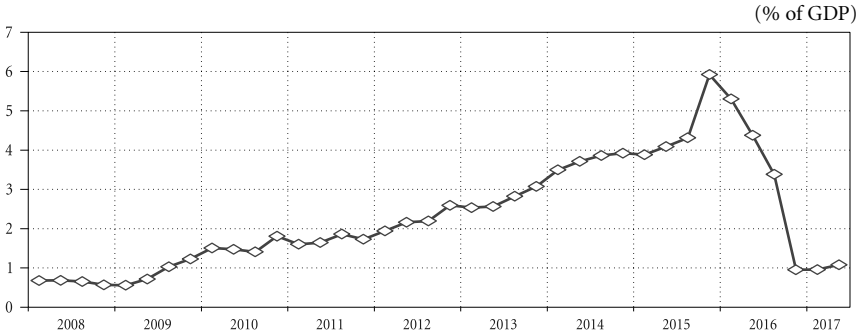
Given the large scope of EU co-financed extensions and improvements in existing infrastructure in Bulgaria[15], there is a strong argument in support of a positive overall impact of EU funds on business investment. While substitution effects may not be completely ruled out, the access to cheaper and more certain EU-related financing is likely to have supported the implementation of projects that may not have been undertaken otherwise, especially in the aftermath of the global financing crisis, when investment activity was impeded by growing uncertainty and scarce external and domestic sources of financing.

The available data on EU funds absorption rates, EU co-financed government investment and capital transfers from EU budget to non-financial corporations (see Appendix 1B on available data sources) indicates that EU related flows to both the government and non-financial corporations have played an important role in explaining the dynamics of business investment, especially in most recent years. The impact of EU funds-related flows was relatively weak until the end of 2011 as the absorption rate (as defined by the rate of payments relative to the overall funding available[16]) was less than a quarter of the allocated Cohesion policy budget (23.8%). The absorption rate strongly increased in 2013–2015 and as a result EU co-financed government investment had a notable positive contribution for economy-wide investment in 2012–2014 and was the main factor behind investment growth in 2015 (see Figure 15). Capital transfers from government to non-financial corporations, related to the implementation of EU co-financed projects, have also increased substantially after 2012 and have peaked in 2015.

[15] The implementation of infrastructural investment projects under the 2007–2013 programming period resulted in the completion of 175 km. of new roads and the reconstruction of respectively 1040 km. and 234 km. of roads and railroads (data as of 2014, EC, 2016:1)

[16] MoF data on Cohesion policy payments (both EU and nationally co-financed). The total budget has been augmented in the course of the implementation of the programming periods to reflect shift in financing and deductions.

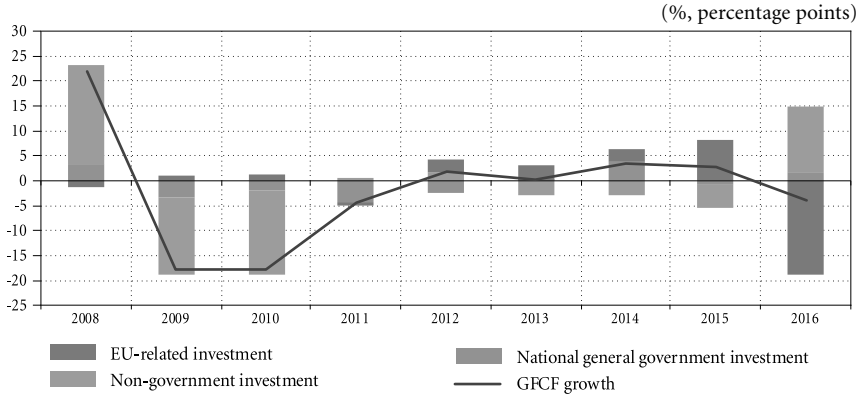
Figure 14. **Capital Spending on EU Programmes**



Note: Four quarters rolling data, includes capital spending in government units and transfers to non-government units on programmes financed by the pre-accession funds, Structural funds, the Cohesion fund, European Agriculture Fund for Rural Development and European Fisheries Fund, as well as other projects and programs, financed with EU funds. Includes co-financing from the national budget.

Source: MoF monthly reports on the implementation of the Consolidated fiscal programmes

Figure 15. **Contribution of EU Co-financed Government Investment to the Growth of Economy Wide Investment**



Note: EU co-financed government investment is proxied by quarterly data on capital transfers received by the government and an additional adjustment is made to account for co-financing from the national budget.

Source: Eurostat, own estimates

The switchover of the 2007–2013 and the 2014–2020 financial frameworks resulted in delays in the implementation of projects under the new program-

ming period[17]. As a result there was a sharp decline in EU co-financed investment in 2016 that caused a decline of economy wide investment, despite favourable non-government investment and national government investment dynamics (see Figure 15). EU flows to direct SME support and other capital transfers to non-financial corporations were also lower as compared to the peak, recorded in 2015.

In summary of this section, it can be argued that mostly cyclical factors supported the slow investment recovery in recent years. The investment recovery in Bulgaria reflected mainly increased non-construction investment, particularly business investment, and is likely to have been driven by improving demand, lower economic uncertainty and easing financing conditions, with additional support from EU-fund related investments.

4. Data and Methodology

Having reviewed the related literature on the topic and having identified the most relevant drivers behind investment dynamics we can now focus on building the econometric framework which will be used to analyse business investment in Bulgaria. The descriptive analysis above reveals that investment activity in Bulgaria over the period 2000Q1–2017Q1 was driven mostly by the cyclical developments in factors, such as external demand, uncertainty and financing conditions. The quantification of the relative importance of each of these factors could be estimated through the lens of a Bayesian VAR model.

4.1. Methodology

From a methodological perspective, we use Bayesian VAR models to estimate the impact of the identified factors on business investment activity. Vector autoregression models (VARs) are very flexible time series models that can capture complex dynamic interrelationships between macroeconomic variables. The flexibility and ability to fit the data comes from the rich parameterization of VAR models. However, that comes at the risk of overfitting the data, of having imprecise inference and large uncertainty about the future paths projected by the model. The relatively short time span of most macroeconomic time series for countries like Bulgaria further aggravates the problem of dense parametrization. A solution to this issue is to employ Bayesian techniques to estimate the VAR model. In Bayesian estimation we use informative priors in order to shrink the

[17] The delay was common across all NMS and was explained largely by common factors: complex legislative and administrative procedures, which require: 1) adoption of relevant EU regulations at the EU level, and 2) negotiations between member states and the EU for finalising the operational programmes and partnership agreements.

richly parameterised unrestricted model towards a parsimonious naïve benchmark model, hence reducing estimation uncertainty (Giannone et al., 2012).

A general VAR model with n endogenous variables, p lags and m exogenous variables can be written as:

$$y_t = A_1 y_{t-1} + A_2 y_{t-2} + \ldots + A_p y_{t-p} + Cx_t + \varepsilon_t, \text{ where } t = 1, 2 \ldots, T \quad (1)$$

$$E(\varepsilon_t) = 0 \quad (2)$$

$$E(\varepsilon_t \varepsilon'_s) = \Sigma \text{ if } t = s \quad (3)$$

$$E(\varepsilon_t \varepsilon'_s) = 0 \text{ if } t \neq s \quad (4)$$

$y_t = (y_{1,t}, y_{1,t}, ..., y_{n,t})$ is a $n \times 1$ vector of endogenous variables, $A_1, A_2, \ldots, A_p$ are p matrices of dimension n×n, C is a $n \times m$ matrix, and xt is a $m \times 1$ vector of exogenous regressors which can be constant terms, time trends, or exogenous data series. $\varepsilon_t = (\varepsilon_{1,t}, \varepsilon_{1,t}, ..., \varepsilon_{n,t})$ is a vector of residuals following a multivariate normal distribution, *i.e.* $\varepsilon_t \sim N(0, \Sigma)$

In more compact notation, the VAR model can be written as:

$$Y = XB + \varepsilon \quad (5)$$

Where Y is a $T \times N$ matrix, where T denotes the total number of observations and N stands for the number of variables (endogenous and exogenous), while B is a matrix of suitable dimension containing the model's unknown parameters.

In the VAR framework the parameters of interest are the coefficients of the model, gathered in a matrix B in (5), along with the residual covariance matrix Σ. Unlike the traditional frequentist approach, in Bayesian econometrics every parameter of interest is treated as a random variable, characterised by some underlying probability distribution. The ultimate goal in Bayesian econometrics is to identify these distributions in order to produce estimates and carry inference on the model. This is done by combining the prior information we may have about the distribution of these parameters (the prior distribution) with the information contained in the data (the likelihood function) to obtain an updated distribution accounting for both these sources of information. This updated distribution is commonly known as the posterior distribution.

For the estimation of the VAR model we use a Matlab package, called *Bayesian Estimation, Analysis and Regression (BEAR) Toolbox*, developed by Dieppe et al. (2016). The model parameters are estimated using an Independent Normal-

Inverse-Wishart prior distribution[18], similar to Meinen and Rohe (2016). This prior has an advantage over the standard Minnesota and the Normal-Inverse-Wishart priors because it allows for more estimation flexibility. Unlike the Normal-Inverse-Wishart, the Independent Normal-Inverse-Wishart prior does not impose dependence between the variance of the residual term and the variance of the VAR coefficients (Dieppe et al., 2016). As the name suggests, it involves setting the prior for the VAR coefficients and the error covariance independently. Under this prior, analytical expressions for the marginal posterior distributions are not available, but numerical methods, such as the Gibbs sampling algorithm, could be used for posterior inferences (Blake and Mumtaz, 2012).

Since the number of parameters in the VAR model is large and presenting all of them is cumbersome, in Section 5 we adhere to the commonly accepted practice to report functions of the VAR coefficients, which summarise information better. In particular, we report impulse response functions and historical decompositions. Impulse response functions describe the response of variable y_i following a shock in variable y_j, everything else being constant. In turn, historical decomposition allows us to decompose the variance of each endogenous variable into its constituent components, whereby each component is due to one structural shock in the model. Structural shocks are identified using a recursive ordering (Cholesky decomposition).

4.2. Data

In this section we describe the choice of variables that we have used for the empirical estimation in the analysis. For the purpose of our VAR specifications we use quarterly data on a number of macroeconomic series, most of which cover the period from 2000Q1 up to 2017Q1. For a detailed list of the sources and individual series coverage please refer to Appendix 4.

The series that we use as a proxy for business investment in Bulgaria is real investment in machinery, equipment and weapons systems published by EUROSTAT. For more information on the choice of a measure of investment and the issues related to this choice please refer to Appendix 1 A.

External demand in the model is proxied by the weighted real import of goods and services of Bulgaria's main trading partners – a series compiled by the ECB. As indicated in the literature review section, weak demand and sluggish growth outlook have been empirically identified as some of the core reasons

[18] Specifically, we set hyperparameter $\lambda_1 = 0.8$, $\lambda_2 = 0.5$, $\lambda_3 = 1$, and $\lambda_4 = 100$ which are the recommended values in the BEAR Toolbox for this particular prior. We also set the number of iterations to 10 000 with the number of burn-in iterations equal to 9 000.

for the investment slowdown, especially for small open economies. This is why we expect higher external demand to lead to higher investment and in the case of Bulgaria this effect is likely to be quite significant as the country is export-oriented.

To account for economic uncertainty in Bulgaria we use 5 different indicators that have been calculated from the EU Commission's Business and Consumer Survey (BCS), following a methodology by Giradi and Reuter (2017). For more information on the calculation and the interpretation of these uncertainty indicators please refer to Appendix 3. Although each of the five uncertainty measures has been separately used in our estimations, in the final model we present the results of the uncertainty indicator that gives us the best fit and economic interpretation. We expect that heightened uncertainty will have a negative impact on investment in the country as suggested by both theoretical and empirical work.

Capacity utilisation in the manufacturing sector is included in the specification in order to account for the economy's supply side motives for investment. We expect that higher levels of capacity utilisation will result in higher investments.

The effect of profits on investment activity in our model is represented by including in the VAR the share of gross operating surplus in gross value added for the manufacturing sector. Data is taken from the National Statistical Institute. In line with relevant literature we expect that an increase in profits would lead to higher investment.

To account for availability and importance of external funding for Bulgaria we have included FDI flows, the data for which we have obtained from the BNB statistics. An increase in FDI flows in the economy is expected to increase investment activity through the external financing of investment channel.

To account for possible credit constraints and cost of financing we include the weighted interest rate on new loans to the corporate sector. This data is taken from the BNB's interest rate statistics. We expect that the lower the interest rates to the corporate sector are, the easier the access to loan financing is, and, thus, the higher will be the increase in investment.

As suggested by recent studies certain variables enter the VAR model in (log) levels (CNB, 2017; Meinen and Rohe, 2016; Manteu and Serra, 2017). Investment and external demand series are transformed by taking their natural logarithm and FDI flows are taken as a share of GDP.

Using the outlined variables, we have estimated a VAR model with six lags of each endogenous variable for the case of Bulgaria over the 2000:Q1 – 2017:Q1 period. We have chosen the following recursive ordering of variables, which would allow us to obtain the structural shocks of our VAR model:

External demand
Uncertainty
Capacity Utilisation
Gross Profit Share
FDI Flow
Real NFC Lending Rates
Business Investment

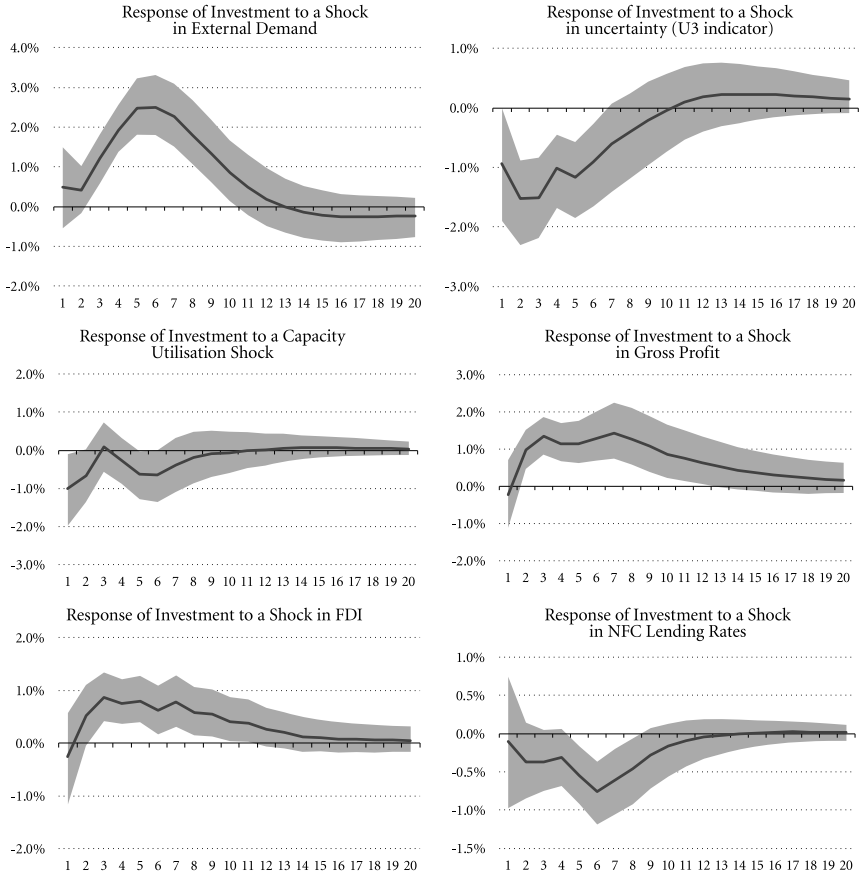
This Choleski ordering is similar to imposing short-run restrictions, such that we impose a particular causal chain on the data. In this setup the variable ordered first (in our case external demand) does not contemporaneously react to variables ordered below it, but will affect all the ones ordered below it contemporaneously. The variable ordered second (uncertainty in Bulgaria) responds contemporaneously only to the variable above it in the ordering. Finally, the variable ordered last (business investment) responds contemporaneously to shocks in all variables ordered above it but shocks in investment do not affect the variables above it in the current period.

A possible shortcoming of our specification is that we do not explicitly include a proxy for future demand expectations, such as the Economic Sentiment Indicator (ESI), published by the EC. However, expectations indicators such as the ESI focus mainly on the near-term future, typically 3 months for firms and 12 months for consumers. As investment decisions tend to have long-term consequences for firms, they are planned carefully and medium to long-term demand expectations are likely to be more important. In the short-run firms can face changes in demand by altering their capacity utilisation. Similar to Barkbu et al. (2015), ECB (2016:1) and Meinen and Rohe (2016) the setup of our model does not include an expectations proxy, which would suggest that economic agents are backward-looking and future demand prospects are assessed primarily based on current demand developments. However, short-term future expectations of economic agents are partly accounted for by the uncertainty indicator, which is based on economic agents' disagreement on future production dynamics.

5. Results

In order to determine the direction in which each of the factors above affects business investment and to establish their significance we look at the impulse response functions (IRFs) from the estimated VAR Model (see Figure 16).

Figure 16. **Impulse Response Functions**



Notes:
The line denotes the median response of investment and the blue area around it denotes the 90% confidence bands.
On the X-axis are plotted the 20 time periods (which in our case are quarters) for which we examine the effect of shocks in our endogenous variables. On the Y-axis we have the response of business investment (in percent from baseline path) to a positive one standard-deviation structural shock in each variable of interest.

The results from the impulse response functions suggest that in Bulgaria one of the factors that has the biggest and most persistent effect on business investment is external demand. This is consistent with findings in the literature on investment drivers in small open economies (CNB, 2017) and supports the initial hypothesis based on the descriptive analysis in Section 3. The maximum impact on business investment from one standard deviation increase in external demand is observed in the second year after the shock.

The effects of uncertainty shocks to investment in Bulgaria are provided in the top right panel. Here, we present the results of the model when the U3 uncertainty indicator is used, as it appears to be the most significant and the one that fits best in the model. This is intuitive given that U3 is an indicator that was specifically designed to account for the economic uncertainty in the industrial sector, which is the main investor in gross fixed assets. When we substitute the U3 uncertainty indicator in our model with another of our uncertainty measures we obtain impulse responses that are correct in terms of the expected direction of the effect but there is a certain loss of statistical significance. The statistically significant response of investment to uncertainty shocks proves that uncertainty has a crucial and long-lasting negative effect on economic agents' investment decisions in Bulgaria. As expected, higher uncertainty constrains investment and leads to firms postponing their investment decisions. Under higher uncertainty agents in Bulgaria tend to adopt a "wait-and-see" approach, depressing investment activity for up to two years.

Judging from the impulse response functions, increases in capacity utilisation do not have a significant impact on business investment in Bulgaria. One possible explanation is that high capacity utilisation on its own does not necessarily transfer into an immediate investment increase. If firms are uncertain about how sustainable demand prospects are, they may choose to respond to the increased demand by pushing up further their production capacity, thereby deferring new investments. This would imply that positive effects on investment from increased capacity utilisation would tend to materialise in the medium term conditional on continuous improvement in demand conditions and firms' sentiments.

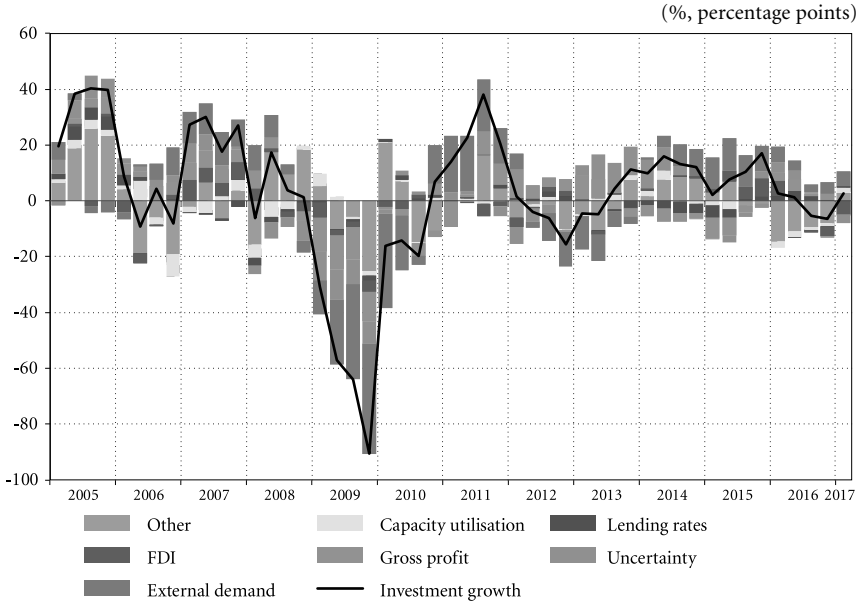
Increases in profit (proxied by the share of gross operating surplus in gross value added for the manufacturing sector) have a positive impact on business investment, reflecting firms' increased capacity to finance investment with internal means. The positive reaction of business investment is very persistent and is maximised in the end of the first year after the initial shock in profit. This result is consistent with empirical research that establishes a positive link between increasing profits and investment decisions (ECB, 2016:1).

Moving to external sources of financing, the reaction of business investment to increases in foreign direct investment is positive and significant (after the second quarter), confirming our findings in Section 3 that FDI has been an important investment financing channel in Bulgaria.

The results from the applied VAR suggest that the reaction of business investment to changes in the cost of financing is relatively less significant. The negative response of investment activity to an increase in the non-financial corporations' lending rates materialises from the first quarter after the initial shock in interest rates and the maximum (negative) impact from an increase in lending rates is observed in the sixth quarter after the shock.

The Choleski ordering in our VAR model also allows us to construct a historical decomposition of business investment dynamics and in this way to quantify the effects of different factors that have driven business investment in different periods of time. Figure 17 shows that results from the VAR broadly confirm our initial expectations about the main drivers of business investment that were stated in Section 3.

Figure 17. **Business Investment Growth – Historical Decomposition**



Note: "Other" here includes the combined effects from the lagged investment and from the exogenous shocks.

Based on the historical decompositions of the VAR model it can be noted that business investment in Bulgaria in the years right before the crisis was largely driven by the observed increases of external demand associated with the cyclical upturn in our main trading partners and optimism, related to Bulgaria's accession in the EU. Increasing corporate profits and low levels of uncertainty further supported the dynamics of business investment in the years prior to the crisis. Furthermore, we can conclude that financing conditions also had quite a notable contribution to the investment boom in the run-up to the crisis. In 2005 and 2006 we observe a positive contribution from favourable domestic lending rates, which in 2007 and 2008 is overtaken by the positive contribution of foreign financing in the form of FDI inflows. Firms' capacity to finance investment with internal means, proxied by the profit share, was also a particularly important factor for business investment after 2005.

In the 2009 – 2010 period lower external demand and the negative impact it had on the overall economy appear to have been the biggest drag on business investment. The negative effect of the lower external demand was further amplified by the heightened uncertainty triggered by the global financial crisis. Both the internal and the external channels of financing of investment were impaired, which exacerbated the decline in investment. Firms' profits and ability to finance investment using internal sources were negatively affected, which is suggested by the observed negative contribution of the share of gross operating surplus in GVA to business investment dynamics. Moreover, the higher risk-aversion of investors as a result of the crisis was reflected in lower FDI flows which also limited business investment in the period. The historical decomposition of the VAR model also indicates that the increase in real non-financial corporations' lending rates up until the end of 2009 had a negative impact on business investment.

With the recovery of external demand and the exhaustion of the negative impact of uncertainty, business investment in Bulgaria also started to recover in the beginning of 2011. However, the 2012 sovereign debt crisis in Europe lead to a new increase in uncertainty, followed by a supressed external demand, which resulted in another weakening of business investment in Bulgaria that lasted up to the second half of 2013.

As evidenced by the historical decomposition, the decrease in uncertainty supported the recovery of business investment in Bulgaria after the second half of 2013. Higher external demand also contributed to investment growth especially after the second half of 2014, offsetting a temporary increase in uncertainty around the end of 2014 and the first half of 2015 triggered by domestic and international developments (see Appendix 3 for more information). In 2014 and

2015 the increasing external demand, the recovery of FDI flows and the increasing absorption of EU funds (assumed to be reflected in "other factors") were the main factors that helped investment return to positive growth rates despite the negative contribution from low corporate profits and tight financial conditions. The sharp decline of EU co-financed investment observed throughout 2016 could explain the decreasing business investment dynamics in spite of the positive impact of demand factors, lower uncertainty and recovering profitability. In the beginning of 2017 business investment recorded positive growth rates on account of continued recovery of external demand.

Overall, the results from the VAR model suggest that external demand has been the most significant determinant of business investment. In addition, the effect of uncertainty on investment activity has also been quite substantial. Availability of financing (both domestic and external) has had a more pronounced effect on investment in the pre-crisis years as compared to the post-crisis period. Constraints regarding availability of funding have been mainly a drag on investment in the aftermath of the crisis. The absorption of EU funds could also be identified as an important factor for business investment in Bulgaria, judging by the contribution of the "other factors" component throughout the years. Together with demand factors it could explain most of the observed investment growth in the beginning of 2017.

6. Conclusion

Gross fixed capital formation is a key driver of the business cycle and plays a prominent role in determining the potential growth of a country. As such, the observed slowdown in investment activity across advanced economies, and particularly countries in the EU, including Bulgaria, observed after the Great Recession, has raised questions on the drivers behind this slowdown.

This paper looks into the drivers behind business investment in Bulgaria and quantifies their effect in order to explain the observed slowdown of investment in the aftermath of the Great Recession and the followed weak recovery in more recent years. This is done using a Bayesian VAR estimation with which we have evaluated the relative importance of factors, such as demand, capacity utilisation, economic uncertainty, firms' profitability, foreign direct investments and lending rates for business investment dynamics during the period 2000Q1 – 2017Q1. Separately from the VAR analysis, we have also examined the role of EU-related funds for explaining business investment dynamics in Bulgaria, given the considerable amount that Bulgaria managed to absorb in recent years under EU programmes.

The results from our VAR model suggest that external demand is the main driver of business investment in the country, which is consistent with the findings of similar studies on small open economies. Apart from external demand, we find that the level of economic uncertainty and corporate profits also play a crucial role in determining the level of business investment in the economy. At the same time the importance of foreign direct investment and lending rates for investment activity appears to be less significant.

From the historical decomposition, obtained from our VAR model, we can conclude that almost all of the factors considered in the analysis contributed positively for the business investment growth in the run-up to the Great Recession. External demand increases, associated with the cyclical upturn in our main trading partners, low levels of uncertainty and increasing corporate profits were the main factors that stimulated investment growth in Bulgaria in that period. Lending rates (in 2005–2006) and FDI inflows (in 2007–2008) also contributed to the business investment growth. Our estimations suggest that the observed slowdown of business investment in the post-crisis period can be explained by the reduced external demand and the increased economic uncertainty in the immediate aftermath of the crisis, which lasted up to mid–2013. After the second half of 2015 our results point to a recovery in corporate profits, as well as favourable financing conditions, which on top of the already observed growth in external demand and decreasing economic uncertainty support the recovery of investment activity in the country. Additional evidence suggest that the dynamics of EU co-financed investment has also been a notable factor in explaining business investment dynamics, especially in 2015 and 2016.

Appendices

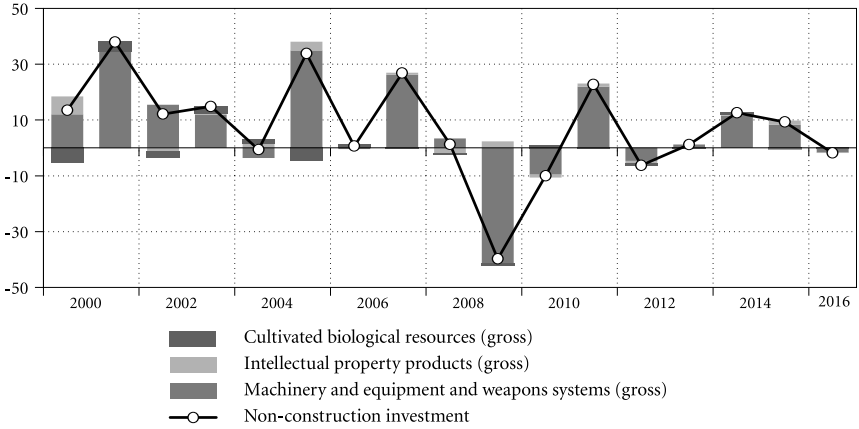
Appendix 1. **Investment Data**

A. Business Investment Data

In this paper we follow the example of recent studies on the topic (Barkbu et al, 2015; ECB, 2016:1; Kose et al., 2017) and for the purpose of our VAR analysis we separate construction investment from total investment in order to extract the part of investment that is particularly relevant in boosting an economy's future productive capacity. In order to obtain our investment proxy we use the breakdown of Gross Fixed Capital Formation by main asset type, which is available from Eurostat for all EU countries. As can be seen in Figure 18, non-construction investment in Bulgaria is driven almost entirely by investment in "machinery, equipment and weapons systems" (encompassing transport investment, information and communication technology (ICT) equipment, other machinery and equipment, and weapons systems).

Figure 18. **Non-construction Investment Growth Rate and Contributions by Major Asset Type**

(annual percentage changes, percentage points)



Source: Eurostat

As such we have decided to use investment in machinery, equipment and weapons systems as our proxy for "business" investment, similar to Meinen and Rohe (2016). Our choice is further supported by the fact that machinery and equipment investment arises mainly in the corporate sector and, thus, its dynamics should reflect firms' investment decisions (ECB, 2016:1). We should note that

in the way that we have defined business investment it is not only private sector investment but also contains some part of public investment. Unfortunately, due to data limitations, we cannot separate out this public investment component which for Bulgaria and other NMS could be relevant in particular periods, especially in the years with high absorption rates of EU-related funds, related to the programming cycle of EU funds. For more advanced countries such as the "core" euro area economies this public investment component is likely to be relatively smaller, with the ECB (2016:1) estimating it at around 10% of euro area investment, with some variation across countries.

B. EU Funds-related Data

Regarding the impact of EU funds on government investment we can use two sources of information: 1) annual data on EU transfers to government and non-government units (source: ESCB Government Finance Statistics, EU budget transactions) and 2) quarterly data on capital transfers received by the government (source: Eurostat, quarterly non-financial accounts of the general government sector). The latter is preferred in this analysis due to its higher frequency and the inclusion of capital transfers from pre-accession funds, which had a significant share in government GFCF in the 2007–2009 period[19].

Assessing the impact of EU funds on non-government investment is not straightforward. Data on capital transfers from the EU budget to non-government units from the ESCB GFS on EU budget transactions is not available for Bulgaria, while data from the same source for NMS[20] indicates that the share of capital transfers to non-government in all capital transfers from the EU budget has varied a lot across years and across countries. On average for these countries, capital transfers to non-government units represented between 34-42% of all capital transfers in the period 2007–2015, but have increased notably in 2016 to 48.1% of all capital transfers.

As alternative sources of data that provide indications for the relative importance of EU funds in financing private investment we use 1) Consolidated Fiscal Programme national methodology cash-based data on capital transfers to non-financial corporations, related to the implementation of EU co-financed projects from all five EU funds and 2) payments on projects financed by Operational Programme "Competitiveness" excluding payments related to financial

[19] While Table GFS Table 1b also includes information on transfers from pre-accession funds, these flows are not split between capital and current transfers and their impact on government investment cannot be assessed.

[20] Data is available for Czech Republic, Estonia, Hungary, Latvia, Lithuania, Poland, Slovakia and Slovenia.

instruments. Both sources provide information only on projects' costs, financed by the EU budget and the national budget, while the amount provided by final beneficiaries could also be quite substantial (up to 50% of the total budget). These sources are clearly second best alternatives as they are both cash-based and their coverage does not correspond to the sectoral coverage of the national accounts. In particular, the first source includes EU financed and national budget financing to non-financial corporations that have been classified in the general government sector (notably the National Railway Infrastructure Company). The second source has a more limited coverage and includes data on the one operational programme aimed to support projects in SME. It does not include support for investment by agricultural producers from the EAFRD. The beneficiaries of OP "Competitiveness" are typically private companies and both financial and no-financial corporations are eligible for support from this programme.

Appendix 2. **Survey Results on the Purpose of Investment and Factors Driving Investment Activity**

Two times per year, in spring and autumn, the European Commission conducts a survey among the companies in the manufacturing sector, collecting information on realised and planned investment, purpose of investment and main factors influencing investment. This appendix summarises the results of this survey for Bulgaria over the period 2002–2016.

Survey results suggest that over the period 2002–2003 around 60% of overall investment was dedicated to the replacement of worn-out equipment and/or rationalisation. In the following years, however, this pattern changed and firms tended to report more investments as serving extension purposes (see Figure 19). The gradual shift in the focus of investment projects could be attributed to the significant FDI inflows over the period 2004–2008, the integration of domestic firms in global value chains as well as to the business cycle (Ivanova and Ivanov, 2017). During periods of economic upswings it could be expected that more investments are dedicated to the expansion of production capacity, compared with downturns when investments are likely to aim mainly at replacing worn-out plant or equipment. Similar to findings for other EU countries, the relative share of investments dedicated to the extension of production sights in Bulgaria is positively correlated with the growth rate in gross fixed capital formation (EC, 2017). Contrary to the pre-crisis period, in the aftermath of the global financial crisis (2010–2014) investment in Bulgaria has mainly served rationalisation purposes as firms aimed at improving their efficiency and competitiveness. In 2013 the share of investment for rationalisation stood at its highest level since the start of the survey 2002. With the cyclical recovery of the economy in 2015–2017, firms started again to report investments that are more aimed at the extension of production capacity.

In terms of the factors that stimulate or limit investment activity, the survey distinguishes between demand, financial, technical (*e.g.* technological factors and the availability of labour) and other factors (*e.g.* taxation). Demand and financial conditions are by far the most important factors behind investment decisions in Bulgaria, regardless of the phase of the business cycle (see Figure 20). In the pre-crisis period demand and financial factors exerted a stimulating effect on investment. This reflected the robust domestic demand growth over the period 2000–2008 combined with favourable external environment. At the same time, the downward trend of interest rates and the availability of credits contributed to easing in financial conditions. In 2009 all factors declined while investment activity contracted severely following the global financial crisis. In 2011–2013 firms in the manufacturing sector that participated in the survey

reported an increase in realised investment, with demand and financial factors being the two main contributors. This was in line with the increase in capacity utilisation over the same period and the gradual recovery of the economy. In addition to demand and financial conditions, increasingly more managers started to report technical factors as relevant for investment decisions, with the share of those managers outpacing pre-crisis levels. According to survey results, investment has declined over the last two years but this possibly could be attributed to the end of the EU programming period.

Figure 19. **Purpose of Investing**

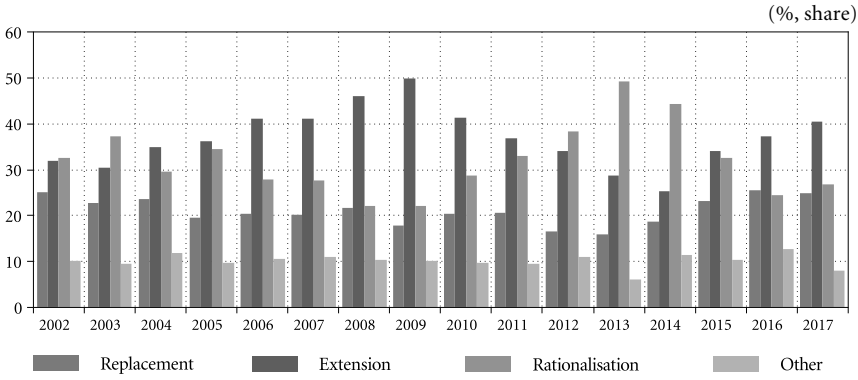
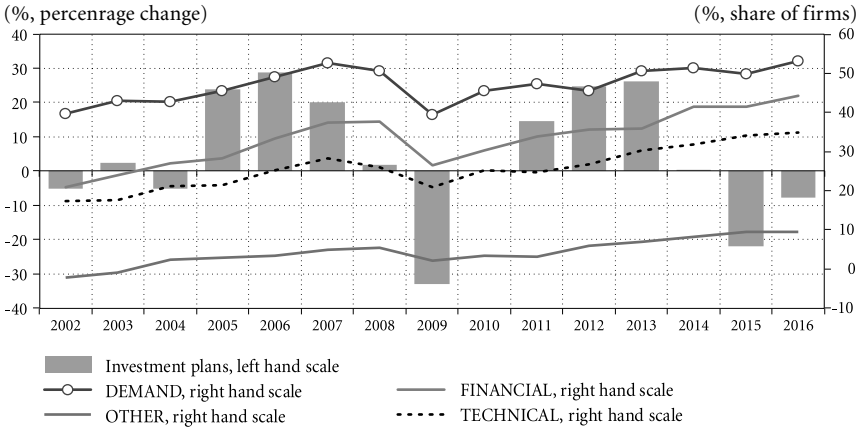


Figure 20. **Factors Driving Investment and Realised Investment in the Manufacturing Sector**



Source: European Commission (AMECO database)

Appendix 3. **Measures of Uncertainty**

There are a number of methods (EC, 2017) that are used to indirectly measure uncertainty, such as financial market volatility indicators, disagreement among professional forecasters, survey-derived indicators and Economic Policy Uncertainty indicators. All of these approaches have their pros and cons.

For our analysis we have chosen to use measures for uncertainty constructed from survey data, since economic uncertainty is measured "at source", i.e. directly at the level of the economic agents who make investment and consumption spending decisions (EC, 2013). To that aim we use data from the European Commission Business and Consumer Survey (BCS)[21] and by following the methodology of Girardi and Reuter (2017) we obtain 4 different uncertainty indicators and a composite uncertainty indicator for Bulgaria, similar to Ivanov (2018)[22]. The derived indicators measure the divergence of business and consumer expectations about the economy and about their finances. The core underlying assumption is that the more economic agents disagree in their assessment or expectations about economic activity, the higher is the uncertainty in the economy. Uncertainty can, thus, be illustrated as a mean preserving increase in the "tails" of the probability distribution of an event (EC, 2013). This distinguishes it from risk, where economic agents are able to form views about the probability distribution of possible future states, based on logic or on past experience (ECB, 2016:2).

The first indicator (U1) measures the disagreement for future activity across all sectors (industry, services, retail trade, consumers, construction) and, thus, should capture uncertainty at the level of the whole economy. This is an expectations-based uncertainty indicator.

The second indicator (U2) does not only focus on disagreement for future activity but also takes into account respondents' backward-looking assessments for the same variables in subsequent survey rounds across all sectors (Girardi and Reuter, 2017). This indicator is a type of ex-post uncertainty measure, based on indirect evaluation of respondents' forecast errors.

The third indicator (U3) is similar to the first indicator (U1) but it measures the disagreement only related to the future production in the industry sector (single question in the survey), rather than the uncertainty in the whole economy

[21] For more information see European Commission Business and Consumer Surveys: https://ec.europa.eu/info/business-economy-euro/indicators-statistics/economic-databases/business-and-consumer-surveys_en

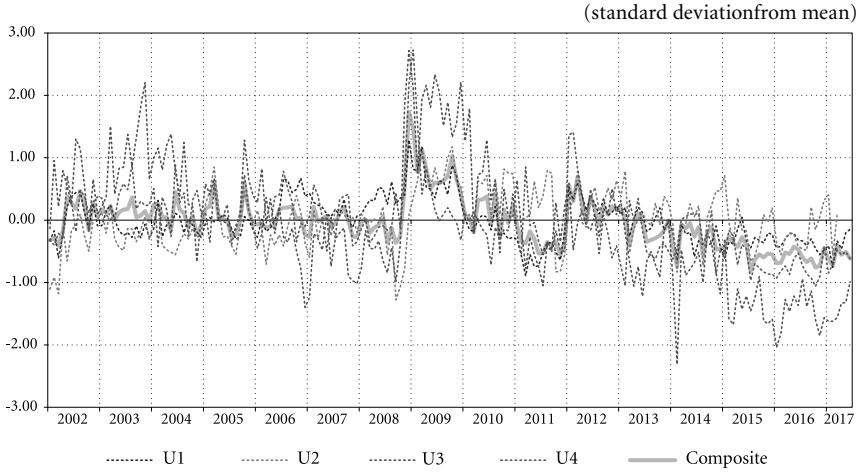
[22] In this section we give a short overview of the methodology and finding of Ivanov (2018). For more information and a detailed discussion on how the uncertainty measures for Bulgaria were derived, please refer to that paper.

(which was the case for U1). This indicator is particularly relevant in explaining investment dynamics in Bulgaria, since the industrial sector is the main investor in gross fixed assets in the country.

The fourth indicator (U4) is constructed in a slightly different way, compared to the previous three indicators. Rather than using the detailed breakdown of "increase" and "decrease" responses from the BCS, this indicator uses dispersion of changes in the overall balance of replies between consecutive rounds of the survey (Girardi and Reuter, 2017). The idea is that a high degree of uncertainty may not be only reflected in disagreement between respondents to a given question but may also be reflected in the resulting balance scores diverging across different questions, i.e. increased dispersion across questions rather than within questions (Manteu and Serra, 2017).

Given that no indicator is a perfect proxy for uncertainty, it might be preferable to compile a composite measure of uncertainty which captures and consolidates the information content of the different uncertainty proxies (ECB, 2016:2). We have chosen to use the median aggregation approach, similar to the ECB (2016:2), in order to obtain a composite uncertainty indicator. The four uncertainty indicators and the composite indicator are presented in Figure 21.

Figure 21. **Uncertainty Indicators, Bulgaria**



Source: EU Commission BCS Survey, own calculations

Appendix 4. **Variables Used**

The table in this section contains detailed information on the variables used in the empirical part of our analysis.

Name	Variable	Source	Coverage	Transformation
INV_M	**Business Investment** – real investment in machinery, equipment and weapon systems	Eurostat, seasonally adjusted	2000Q1-2017Q2	Natural logarithm transformation
WTV	**External Demand for Bulgarian Goods and Services** – weighted index of the real imports of Bulgaria's main trading partners	ECB, seasonally adjusted	2000Q1-2017Q2	Natural logarithm transformation
U1	**Uncertainty Indicator 1** – related to the future expectations in the economy	BCS, own calculations	2000Q1-2017Q1	Standardised index
U2	**Uncertainty Indicator 2** – related to backward-looking assessments	BCS, own calculations	2000Q1-2017Q1	Standardised index
U3	**Uncertainty Indicator 3** – related to the future production in the industry sector	BCS, own calculations	2000Q1-2017Q1	Standardised index
U4	**Uncertainty Indicator 4** – obtained from balance scores	BCS, own calculations	2000Q1-2017Q1	Standardised index
U_C	**Composite Uncertainty Indicator**	BCS, own calculations	2000Q1-2017Q1	Standardised index
GOS_P	**Share of Gross Operating Surplus in GVA for the manufacturing sector**	National Accounts, four quarters rolling data	2000Q1–2017Q1	Natural logarithm transformation
CU	**Capacity Utilisation in the Manufacturing Sector**	Business Surveys, BNB seasonally adjusted	2000Q1 – 2017 Q1	N/A
IR	**Real Non-Financial Corporations (NFC) Lending Rates (GVA deflated)**	BNB Interest Rates Statistics	2000Q1–2017Q1	N/A
FDIFLOW_GDP	**FDI Inflow as a Share of GDP**	BNB Balance of Payment Statistics (BPM5 and BPM6)	2000Q1–2017Q1	N/A

Appendix 5. **EU Funds and Their Role for Government and Corporate Investment**

As described in Appendix 1B two alternative sets of data are used to quantify the impact of EU funds on government investment and investment of non-financial corporations.

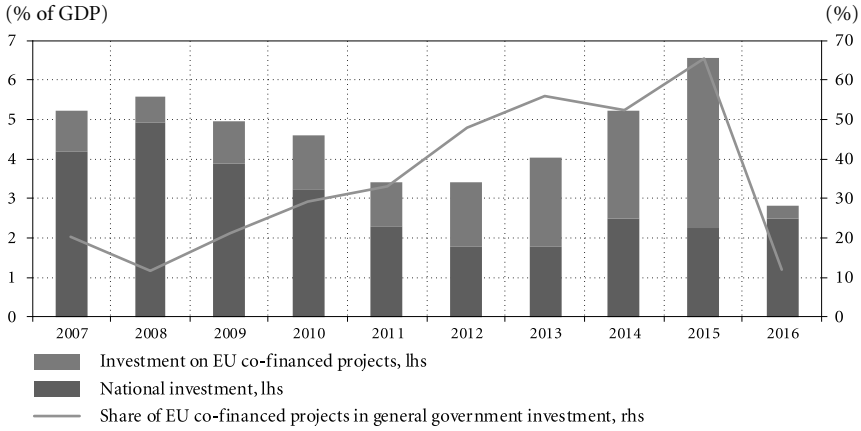
As already noted, EU funds have been one of the major factor for the dynamics of government investment in recent years. The share of EU co-financed projects in total government investment has increased from just 11.7% in 2008 to the substantial 65% in 2015, when EU co-financed investment peaked at 4.3% of GDP (see Figure 22).

Capital transfers from government to non-financial corporations, related to the implementation of EU co-financed projects, have also increased substantially after 2012 and have peaked in 2015, when they accounted for 1.2% of GDP, or 8.8% of total investment by non-financial corporations (see Figure 23). It has to be notes that these transfers also included co-financing on projects, implemented by state owned companies, while direct support to small and medium enterprises was more limited. Under the 2007–2013 framework, the total budget of the Operational programme "Competitiveness of the Bulgarian economy" amounted to EUR 1.2 billion or 1.9% of private investment (0.4% of GDP) and under the 2014–2020 framework, there are two operational programmes (OPs "Innovations and Competitiveness" and "SME Initiative") that finance projects related to enterprise support and innovation with a total budget of EUR 1.4 billion, or 2.4% of private investment (0.4% of GDP).

In addition to grants, EU funded financial instruments are also meant to play an important role in supporting private investment. Around 30% of the 2007–2013 Operational programme "Competitiveness of the Bulgarian economy" budget was allocated to financial instruments for enterprises (JEREMIE initiative). The funds allocated to financial instruments for enterprises under the 2014–2020 financial framework amount to EUR 337 million and include 100% of the budget of the "SME Initiative" (EUR 102 million) and the contracted funds of Fund Manager of Financial Instruments in Bulgaria EAD on OP "Innovations and Competitiveness" (EUR 235 million)[23].

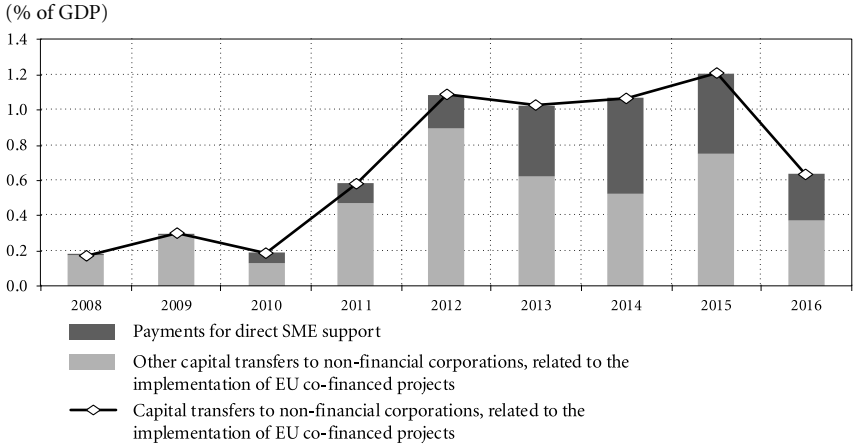
[23] As of April 2017 the Fund Manager of Financial Instruments in Bulgaria EAD has contracted financial resources from four operational programmes totalling BGN 1.2 billion. In the previous programing period, the majority of the funds allocated to financial instruments were the funds under the JEREMIE programme (finances as part of OP "Competitiveness), while financial engineering instruments for urban development, energy efficiency and renewable energies (Jessica initiative) were of marginal important (EUR 33 million).

Figure 22. **EU Co-financed Investment as a Share of Total Government Investment**



Source: Eurostat, own estimates

Figure 23. **Capital Transfers from Government to Non-financial Corporations, Related to the Implementation of EU Co-financed Projects**



Note: Payments to direct SME support represent yearly payments under OP "Competitiveness" excluding payments related to financial instruments for enterprises.

Source: MoF, Unified management system for the EU structural instruments in Bulgaria, own estimates

With the launch of the Investment plan for Europe ("Juncker Plan") in 2014, the scope for interventions to stimulate investment in EU countries has been broadened further. The Juncker Plan's European Fund for Strategic Investments (EFSI) aims to support strategic investments in key areas such as infrastructure, education, research and innovation, as well as risk finance for small businesses. Its main goal is to overcome bottlenecks in current investment environment and to mobilise private investment in times when public resources are scarce. As of May 2017 total financing in Bulgaria under the Juncker Plan's for SME support and support of infrastructure and innovation projects amounts to of EUR 206.9 million, with expected overall impact on investment of nearly EUR 1.124 billion (after financial intermediaries' and private companies' capital has been mobilised).[24]

[24] According to latest information published on MoF website as of July 2017.

Bibliography

Barkbu, B., S.P. Berkmen, P. Lukyantsau, S. Saksonovs and H. Schoelermann (2015). "Investment in the Euro Area: Why Has It Been Weak?", IMF Working Paper.

Blake, A., H. Mumtaz (2012). "Technical Handbook – No. 4 Applied Bayesian econometrics for central bankers", Centre for Central Banking Studies, Bank of England, p. 38.

Bloom, N., S. Bond and J. Van Reenen (2007). "Uncertainty and Investment Dynamics", Review of Economic Studies, 74(2), pp. 391-415.

BNB (2017). "Dynamics of House Prices in Bulgaria between 2000 and 2016", Economic Review Issue 1/2017.

Busetti, F., C. Giordano, G. Zevi (2015). "Main Drivers of the Recent Decline in Italy's non-construction investment" Banca D'Italia. Occasional Papers 276.

Czech National Bank (2016). "The impact of drawdown of EU funds on investment by non-financial corporations", Box in the CNB Inflation Report, Issue IV/2016.

Czech National Bank (2017). "Private investment – an analysis of individual data on non-financial corporations", Box in the CNB Inflation Report, Issue I/2017.

Dieppe, A., R. Legrand and B. Roye (2016). "The BEAR toolbox", ECB Working Paper Series № 1934/July 2016.

EC (2013). "Assessing the Impact of Uncertainty on Consumption and Investment", Special focus in "Quarterly Report on the Euro Area", Volume 12, N.2 (2013).

EC (2016:1). "Ex post evaluation of Cohesion Policy programmes 2007–2013, focusing on the European Regional Development Fund (ERDF) and the Cohesion Fund (CF): Task 3 Country Report Bulgaria", Directorate-General for Regional and Urban Policy European Commission B-1049 Brussels.

EC (2016:2). "Public finances in EMU–2016", institutional paper 045/2016

EC (2017). "European Business Cycle Indicators: 4th Quarter 2016", Technical Paper 013.

ECB (2016:1). "Business investment developments in the euro area since the crisis", Article in ECB's Economic Bulletin, Issue 7/2016.

ECB (2016:2). "The impact of uncertainty on activity in the euro area", Article in ECB Economic Bulletin, Issue 8/2016.

Eesti Pank (2016). "Transfers from the European Union", Box in the Estonian Economy and Monetary Policy quarterly bulletin, Issue 1/2016.

EIB (2013). "Factors influencing investment during the financial crisis and deep economic recession: the European experience since 2008" Investment and Investment Finance in Europe.

EIB (2015). "Investment in Tangible Capital" Investment and Investment Finance in Europe.

EIB (2016). "Investment and Investment Finance in Europe: Financing productivity growth", Annual Report 2016.

Giannone, D., M. Lenza and G. Primiceri (2012). "Prior selection for Vector Autoregressions", ECB Working Paper Series №1494/November 2012.

Girardi, A., A. Reuter (2017). "New uncertainty measures for the euro area using survey data", Oxford Economic Papers, Volume 69, Issue 1: 278–300.

Havik, K., K. Morrow, F. Orlandi, C. Planas, R. Raciborski, W. Röger, A. Rossi, A. Thum-Thysen and V. Vandermeulen (2014). "The Production Function Methodology for Calculating Potential Growth Rates & Output Gaps", European Commission, Economic Papers 535.

IMF (2007). "Bulgaria: Selected Issues", IMF Country Report No. 07/390.

IMF (2014). "Article IV Consultations: Euro Area Policies". Country Report 14/199, Washington, D.C.: IMF.

Ivanov, E. (2018). "Constructing an Uncertainty Indicator for Bulgaria", BNB, DP/109/2018.

Ivanova, N. and E. Ivanov (2017). "The Role of Bulgaria in Global Value Chains", BNB, DP/105/2017.

Karamisheva, T. (2016). "Analysis of the Bank Lending Survey Results for Bulgaria (for the 2003–2014 Period)", Bulgarian National Bank DP/102/2016.

Kose, M., F. Ohnsorge, L. Ye, E. Islamaj (2017). "Weakness in Investment Growth. Causes, Implications and Policy Responses", World Bank Policy Research Working Paper 7990.

Kotseva, P. and M. Yanchev (2017). "Analysis of the Housing Market Developments and the Underlying Macroeconomic Fundamentals in Bulgaria", BNB, DP/103/2017.

Manteu, C., S. Serra (2017). "Impact of uncertainty measures on the Portuguese economy", Working Papers 9 (2017), Banco de Portugal.

Meinen, P. and O. Rohe (2016) "On Measuring Uncertainty and its Impact on Investment: Cross-country Evidence from the Euro Area", Deutsche Bundesbank Discussion Paper No 48/2016.

Ministry of Finance (2016). "EU funds in Bulgaria: assessment of the macroeconomic effects from the implementation of EU co-financed programmes ", available only in Bulgarian, Sofia 2016.

Mitra, P. (2011). "Capital Flows to EU New Member States: Does Sector Destination Matter?", IMF Working Paper WP/11/67.

Mizen, P. and P. Vermeulen (2005). "Corporate Investment and Cash-Flow Sensitivity: What Drives the Relationship", ECB Working Paper Series, No. 485 / May 2005.

Narodowy Bank Polski (2017). "The absorption of EU structural funds in 2016", Box in the NBP inflation report, Issue March 2017.

Paliova, I. and T. Lybek (2014). "Bulgaria's EU Funds Absorption: Maximizing the Potential!", IMF Working Paper 14/21, International Monetary Fund, Washington, DC.

Vashakmadze, E., G. Kambou, D. Chen, B. Nandwa, Y. Okawa, D. Vorisek (2017). "Regional Dimensions of Recent Weakness in Investment. Drivers, Investment Needs and Policy Responses" World Bank Policy Research Working Paper WPS 7991.