CONTENTS

Neutral Interest Rates in CEEMEA — Moving in Tandem with Global Factors
   Clemens Grafe, Sara Grut, Lorenzo Rigon

Impact of Banking Supervision on Banking System Structure: Conclusion from Agent-Based Modelling
   Alexey Ponomarenko, Andrey Sinyakov

Hidden Negative Capital in Banking Before and After the Senior Management Change at the Bank of Russia
   Mikhail Mamonov

The Folk Theorem of Decreasing Effectiveness of Monetary Policy: What Do the Data Say?
   Ugo Panizza, Charles Wyplosz

Reflections on Central Banking, Protectionism and Globalization
   Jacob A.Frenkel

Reflections on Reflections from the Russia’s Mirror (Commentary to Dr. Jacob Frenkel’s Article)
   Alexander Morozov
# Russian Journal of MONEY AND FINANCE

**V. 77 №1, MARCH 2018**

## CONTENTS

### Foreword

*Ksenia Yudaeva*........................................................................................................... 3

### Interest Rates

Neutral Interest Rates in CEEMEA — Moving in Tandem with Global Factors

*Clemens Grafe, Sara Grut, Lorenzo Rigon* .......................................................... 6

### Banking Supervision: Russia

Impact of Banking Supervision on Banking System Structure: Conclusion from Agent-Based Modelling

*Alexey Ponomarenko, Andrey Sinyakov* ........................................................... 26

Banks’ Hidden Negative Capital Before and After the Senior Management Change at the Bank of Russia

*Mikhail Mamonov* ............................................................................................ 51

### Innovations in Monetary Policy

The Folk Theorem of Decreasing Effectiveness of Monetary Policy: What Do the Data Say?

*Ugo Panizza, Charles Wyplosz* ........................................................................... 71

### Policy Discussion

Reflections on Central Banking, Protectionism and Globalization

*Jacob A. Frenkel* ................................................................................................ 108

Reflections on Reflections from the Russia’s Mirror (commentary to Dr. Jacob Frenkel’s Article)

*Alexander Morozov* .......................................................................................... 124

---

All the articles are freely available at: [www.cbr.ru/money-and-finance](http://www.cbr.ru/money-and-finance)
The Russian Journal of Money and Finance is a scientific peer-reviewed journal published by the Bank of Russia. Since 2018 it is issued quarterly both in Russian and in English.

The Journal was established in 1938 (from the middle of 1941 to the end of 1945 was not published).

The journal is included in the list of leading scientific journals of the Higher Attestation Commission and in the Web of Science — RSCI.

Submission guideline: www.cbr.ru/money-and-finance/information-for-authors
www.cbr.ru/money-and-finance
Letter from the Editor

Ksenia Yudaeva,
First Deputy Governor of the Bank of Russia

Dear colleagues,

I am happy to introduce the first 2018 issue of the Russian Journal of Money and Finance — not just this year’s first but first in its new form. In changing the journal’s concept we aim to enhance its scholarly component and to improve the quality of publications by giving preference to articles reporting the results of original research. This has meant introducing a considerably more thorough article selection and reviewing process.

We have set the goal of publishing only high-level research that is of interest to all experts in macroeconomics and finance. This academic journal seeks to promote research in macroeconomics and finance in emerging markets and to further economic science in Russia by acquainting Russian scholars with the latest developments in the world economics, particularly in the areas directly related to central banks’ mandate. The journal is now also published in English because we wish to give the international audience an idea of the Russian researchers’ results and development of their research.

To what extent we have achieved this in the first issue of this renovated journal is certainly for our readers to judge. But I am pleased to note that this issue features high-quality studies by Russian and foreign researchers on contemporary subjects for monetary policy in Russia and throughout the world, as well as papers analyzing the situation in the Russian banking sector.

Now a few words about the papers themselves and implications of their findings for economic policy.

The paper by Clemens Grafe, Sara Grut and Lorenzo Rigan estimates the neutral interest rate for Central and Eastern European countries, including Russia. The neutral interest rate is a central bank’s interest rate in a state of economic equilibrium, i.e., when GDP growth rate of a given country is equal to potential growth rate, with inflation on the target level. From a methodological perspective, this study presents an interesting case of using the currently widespread technique of estimating neutral, or equilibrium, interest rates as applied to
emerging markets. In large economies with reserve currencies, the neutral rate is mainly determined by internal factors. However, in smaller and emerging economies, interest rates are largely driven by external factors. In Grafe et al.'s paper the results of econometric analysis suggest that whereas in Eastern European countries the key driver of neutral rate movements is the US neutral rate, in Russia, it is terms of trade, primarily for oil, that play a more important role. I can add that as the Russian economy’s dependence on oil prices lessens due to, among other factors, changeover to inflation targeting and adoption of the budget rule (under which revenues from commodities exports in excess of the preset level go to reserves rather than to the government budget), one could expect a decline in the neutral rates’ dependence on oil prices.

Two papers appearing in this issue analyze the effectiveness of monetary policy innovations in countries with reserve currencies which have found themselves in a situation of near-zero interest rates. Ugo Panizza and Charles Wyplosz test the hypothesis that the effectiveness of central banks’ policy to buy up assets weakens with time. Their results are not entirely definitive, though. If the effectiveness of policy is analyzed by estimating shadow interest rates (i.e., rates that could produce the same effect as an asset-buying policy), then numerous tests confirm the hypothesis that the effectiveness of this policy weakens over time. But tests relying on central bank announcements about quantitative easing for analysis do not support this hypothesis. I would draw the following conclusion from these results: the effectiveness of quantitative easing does weaken over time, and thus monetary policy is incapable of dealing with all the problems in an economy. But as far as instruments are concerned, communication policy is in itself a fairly effective one, thanks above all to its influence on financial markets’ expectations. I would also like to draw Russian researchers’ attention to how thoroughly the authors study the problem and how many tests for the robustness of results they perform. Unfortunately, studies by many Russian researchers so far show a lack of this thoroughness.

The key points of Dr. Jacob Frenkel’s paper, which is published in the section covering economic policy, is closely related to the conclusions of the above-mentioned article. Jacob Frenkel, a well-known macroeconomist and the Bank of Israel’s former Governor, has been long setting forth the idea that the so-called unconventional policy measures should be resorted to on a short-term basis, under extraordinary circumstances in the markets and should not be the only measures used under such circumstances. Data on the global economic situation leads Dr. Frenkel to conclude that the world economy is ready for these measures to be scaled down. The trends and approaches covered in his article — with regard to both monetary policy and
globalization issues — are directly related to both Russia’s economic situation and its economic policy. Alexander Morozov, Director of the Bank of Russia’s Research and Forecasting Department, covers this angle in his brief comment.

The remaining two papers published in this issue analyze the situation in the Russian banking sector. The study by Alexey Ponomarenko and Andrey Sinyakov analyzes the effect of banking supervision enhancement (entailing license withdrawal from weak players) on competition in the banking sector. This is a theoretical study with an analysis of potential implications relying on an agent-based model. I would like to note that this is now an extensively developing model class which is fairly often used to analyze financial stability issues. It helps depart from the representative agent concept and examine the results of interaction between different agent types. This study suggests that the impact of supervision toughening on various banks’ competitive position is nonlinear. At the first stage, when customers lack clear understanding of specific banks’ reliability, funding may become more expensive for small and medium-sized banks. However, after the majority of fragile players have gone out of business and customers have gained greater confidence in the reliability of small and medium-sized banks in operation, these banks’ competitive positions will substantially improve. The results of Mikhail Mamonov’s study shows that after the Bank of Russia has toughened banking supervision, banks have stepped up their efforts to prevent the emergence of problems with capital or have started to address them promptly.

Hence the long-term and, to a large extent, short-term results of supervision enhancement improve competition in the banking sector, providing incentives for players to improve risk-management. The extensive literature on the banking sector suggests that fragile banks or so-called zombie-banks hamper economic growth. Thus tougher supervision policy in the long run creates greater opportunities for a balanced economic growth.

In conclusion, I would like to thank the authors, reviewers and editors of this first issue. Your efforts have set a high standard for the renovated journal. I would also like to address potential authors once again: we expect you to submit new articles which would give us a better insight into the new trends in the economy as well as monetary and financial realms.
Neutral Interest Rates in CEEMEA — Moving in Tandem with Global Factors

Clemens Grafe, Goldman Sachs Global Investment Research*
Sara Grut, Goldman Sachs Global Investment Research*
Lorenzo Rigon, Stanford University*

The decline in long term interest rates in the first decade of this millennium and a slow response of output to very lower interest rates post the global financial crisis in developed countries has sparked renewed interest into estimates and drivers of the so called neutral rate. In this article we contribute to this discussion by applying an open economy version of the Kalman filtering approach pioneered by Laubach and Williams to a number of emerging economies of the European time zone. We find that the decline in neutral interest rate in most of the countries we analyse mirrors that elsewhere, but that unlike in more developed countries very little of that decline can be attributed to declines in potential growth rates. Instead we find that estimates of US neutral rates can explain the overwhelming part of the dynamics in neutral rates in the region and we provide estimates of the elasticities to US rates. The exception is Russia where we cannot attribute any significant part of the dynamics in neutral rates to either Russia’s trend growth or the US neutral rate. Instead, we conjecture that in the last two decades neutral rates have been mostly driven by persistent terms of trade shocks.

Keywords: neutral rate, US neutral rate, Russian neutral rate, terms of trade, terms of trade shocks, potential output growth, zero lower bounds

JEL: C5, E43, E52

1. Introduction

Central banks can typically add or withdraw stimulus from an economy by targeting some short-term market interest rate. By doing so, they usually attempt to meet a nominal target, which these days takes the form of an inflation target for most Central Banks. Apart from the current exchange rate the judgement of whether or not a Central Bank should raise rates involves a view on a number of variables such as potential growth, defined as the rate at which an economy can expand that operates at full capacity without adding to price pressures. The neutral

*The views expressed here are solely those of the authors and do not necessarily reflect those of others at Goldman Sachs International or Stanford University.
or natural rate of interest is then defined as the real interest rate that a Central Bank needs to target if it was to keep the economy at capacity. Assuming a Central Bank succeeds in keeping rates close and symmetric around its inflation target, it can be thought of as the intercept of the Taylor rule that appropriately describes the Central Bank’s reaction function minus the inflation target.

While the concept of natural or neutral interest rates is quite intuitive, the rate is not observable. Even worse and as outlined above the identification requires as inputs other unobservable variables such as a measure of spare capacity and potential growth. Consequently, any estimate will necessarily be imprecise and the concept has its critics as a useful tool for policy making. Milton Friedman for instance argued against natural rate–based policies in his presidential address to the American Economic Association, positing that “One problem is that [the policymaker] cannot know what the ‘natural’ rate is. Unfortunately, we have as yet devised no method to estimate accurately and readily the natural rate of either interest or unemployment. And the ‘natural’ rate will itself change from time to time.”

The latter is key, if we could assume that potential growth and the natural rate were time invariant, this would clearly help to identify both. However, any growth model illustrates that neutral rates are unlikely to be stable. The real interest rate even in the most basic closed economy non stochastic growth models is typically a function of time preferences, i.e. to which extent people value the future and hence want to save as well as the growth rate of total factor productivity, population growth and the intertemporal elasticity of substitution. In open economy models, global equivalents of the above will matter as well. In practice, the rate is likely to also depend on the regulatory framework, lending standards, quality of balances sheets and persistent terms of trade shocks among others. Many of these factors are unlikely to be time invariant and hence we cannot just calculate the rate by taking arbitrarily long averages. Neither are the factors influencing the neutral rate and their influence on rates easily observable.

That said even though we might not be able to write an algorithm that allows Central Banks to know the neutral rate at a given point in time, implicitly policy makers are arguably forced to make an assessment about the level of neutral rates when setting policy rates. Thus, though imprecise, having an estimate appears better than not having one.

While the concept of neutral rates is more than a hundred years old, having originally been proposed by Knut Wicksel (1898), it has received significantly more attention in the last decade for exactly the reasons, the rate is difficult to pin down. Until the turn of the millennium, long term averages of real rates did not display any obvious trend and hence simple averages were usually seen as good enough indicators of neutral rates. However, in the current millennium, long term interest rates net of the inflation targets started to decline. This process accelerated post the global financial crisis sparking renewed interest in applying more advanced econometric methods to estimate the path of neutral rates.
In this article, we contribute to this literature by applying a method first used in Laubach and Williams (2003) to some emerging countries of the European time zone (summarised under the acronym CEEMEA in what follows). There have been studies that have looked at individual countries in the region (Kreptsev et al. (2016) for Russia, Us (2018) for Turkey, Kuhn et al. (2017) for South Africa, Stefanski (2017) for CEE) but to the best of our knowledge this is the first contribution that uses the same framework across different countries which allows us to make some inferences that those papers couldn’t. Also with the exception of Kuhn et al. (2017) the papers use closed economy models while we use the small open economy as proposed by Berger et al. (2014) or Pedersen (2015).

For most countries our estimated dynamics for neutral rates displays a similar dynamics to that found in other countries. They fell from the beginning of the millennium but have largely stabilised from 2011-2012 onwards. We find two exceptions: In Hungary, our estimates suggest the neutral rate is still declining (though it remains above that in Poland for instance); More interestingly, Russian neutral rates fell in tandem with others early on, but while others continued to fall, Russian neutral rates have risen since the global financial crisis (GFC). Overall, neutral rates in Poland, Hungary and Romania appear to be converging to levels quite close to those found by others for small open economies in Europe, while in the Czech Republic and Israel neutral rates have effectively fallen to similar levels seen in the developed world (in the Czech Republic this is arguably owing to that country’s peg to the Euro).

In the standard model from Berger et al., that in line with Laubach and Williams (2016) models the neutral rate as a function of a country’s trend growth rate and a non-observable factor that is modelled as a random walk, we only attribute a very small part of the decline in neutral rates to the potential growth rate, a result that is very much in line for instance with Carrillo and Polly’s (2014) work on Mexico. However we find strong co-movement between the unobservable random walk variable and US neutral rates.

We hence complement the results of the standard model with those of a model where we use the US neutral rate as an exogenous input. We find that the US rate and its covariance with the unobserved factor in most cases explains more than 80% of the variation. One again, the exception is Russia, where we have to attribute all the variation to the unobservable factor.

The second model also allows us to make some inference on the impact of US neutral rates on the neutral rates in the countries we studied. We find quite high elasticities ranging from 4.1 in Turkey to close to 1 in the Czech Republic and Hungary and 0.2 in Russia though the latter is clearly only very imprecisely measured given the above.

---

2 The countries we look at are the Czech Republic, Hungary, Poland, Romania, Israel, Turkey, South Africa and Russia.
We fail to identify either Russia growth or US interest rates as drivers of Russian neutral rates. In fact the correlation between Russian neutral rates and trend growth appears to be negative. As we argue below, the correlation between Russian neutral rates and persistent changes in Russia’s Terms of Trade instead is quite strong and likely to dominate the other factors. Expectations of trending terms of trade in the future are likely to translate into a trend in future real exchange rates which will influence not only short term real interest rates but even longer term ones. If this conjecture proved right, a view of range bound terms of trade would imply a significant reduction in neutral rates in commodity producing countries like Russia in the years ahead.

2. Literature Review

Though the concept of neutral interest rates exists for more than a century, attempt to identify it intensified post the millennium when long term interest rates started to display what seemed a trend decline.

The strand of the literature that our contribution best fits into goes back to the seminal paper by Laubach and Williams in 2003, which proposes a joint estimation of the natural rate of interest, potential output, and its trend growth rate using a Kalman Filter. In their approach they model the natural rate for the US as a function of US growth and an unobservable variable meant to pick up the impact of changes other than growth such as shifts in savings rates due to changes in population growth, time preference, regulation etc. They estimate that between 1960 to 2003 the neutral rate in the US varied significantly between a peak of about 4.5% in the mid 60s and a minimum of 1.25% in the early nineties. However they found no secular trend with their estimate for the end of their observation period being close the period average at about 3%.

In their specification, the beta of the real rate vs potential output is close to one but a significant part of the variation in the trend growth rate is explained by the unobservable factor. In fact, on their estimates US trend growth is reasonably stable. Thus, as Hamilton et al. (2015) point out, neutral rates even for the US are statistically not as closely linked to US growth as is often thought.

An alternative strand of investigations into the neutral rate uses stochastic dynamic equilibrium models (e.g. Giammariou et al., 2003). These models have their advantages if the focus is on evaluating different monetary policy rules for instance in a cost/benefit analysis, but that comes at the price that the estimate of the neutral rate is likely to be quite model dependant. Also they are ill suited to incorporate structural parameters of the model that are not possibly not time invariant. Still the estimate Giammariou et al. (2003) obtained on the neutral

---

3 Consequently the decline in neutral rates since the GFC is quite plausibly not a sign of a sizeably permanent decline in trend growth or a secular stagnation (Summers, 2014) but by factors such as changes in regulation, long lasting balance sheet effects that hold back the recovery etc.
rate in the Euro area in 1994-2000 was very close to that obtained by Williams and Laubach over the same period, a period of relative stability in trend growth and long term interest rates.

Laubach and Williams followed up their earlier paper in 2015 extending the data set into the period post the Great Recession. While results for the earlier period confirm those in Laubach (2002), they found that the natural rate had declined post the Global Financial Crisis (GFC) to levels close to zero and remained at these levels even during the recovery phase. Their model ascribes about half of the decline to lower trend growth but other factors appear to be at work as well that are less easy to pin down.

Holston et al. (2016) apply the same techniques to other developed G7 economies, Canada, the Euro area and the UK. They find that neutral rates in all three declined similarly to the US and indeed display a substantial degree of co-movement, suggesting a significant influence of a global factor in determining the rates. Using a Cholesky decomposition the authors find a great deal of interdependence with shocks to the US neutral rate contributing about 35% of the variation in Canadian rate on a 10 year horizon and further but smaller contributions from the other two regions. While the US rate is less dependent on others even for the US they find in total about 20% of the variation being driven by shocks to other rates. The co-movement in rates driven by global factors as well as that even in theory, a country’s neutral rate must to some extent depend on global rates is necessarily a larger constraint on applying the Laubach Williams framework the more open and smaller the economy is.

Berger and Kempa (2014) present an extension of the framework to the open economy and apply it to Canada. Pedersen (2015) uses a similar framework and applies it to Denmark. We largely follow this open economy approach in our analysis.

More recently, researchers have started to apply the same approach to emerging economies. Perelli and Roach (2014) apply different filtering techniques including the approach discussed above to real interest rates in emerging economies (EM) finding that the secular decline of real interest rates in many EM’s already started in 2002 and prior to the GFC. Not surprisingly they find that common global factors play an important role in the determination of neutral rates in EM. Carrillo and Polly (2014) apply the Laubach approach to Mexico. While they similarly to the DM literature find a sharp decline in the neutral rate from about 3% pre-GFC to about 1% post the GFC, only a very small part of that decline can be attributed to the variation in Mexico’s trend growth rate with most of it being explained by the model omitted factor, again finding that the standard closed economy Laubach/Williams (LW) model is less well suited for open small economies.
Kreptsev et al. (2016) analyse the relationship of the equilibrium interest rate in Russia with potential GDP growth and global interest rates and extend the investigation into estimating neutral rates using the standard LW framework. Because short term money market rates were negative in real terms prior to the GFC, they use the short term bank lending rates which as the authors recognize themselves is problematic to use as a risk free rate and hence the results need adjustments. They find a neutral rate close to 3% for the lending rate in Russia but given a 250bps spread between money market rates and the lending rate used this would imply a neutral short term money market rate of 0.5%.

Kuhn et al. (2017) apply an open economy version of the LW framework to South Africa and use Bayesian techniques instead of the maximum likelihood method. They find no secular trend in the neutral rate till the GFC but a significant decline thereafter with neutral rates declining from around 3% pre GFC towards 1.5-2%. While prior to the GFC the upside pressure on neutral rates from rising trend growth were balanced by downside pressure of the non observable factor, post the GFC the decline is mostly associated with a decline in domestic trend growth. They attribute most of the dynamics in the non observable factor to changes in the South Africa’s CDS and changes in the EM current accounts which they use as a measure for excess EM savings putting downside pressure on global rates.

3. Model

Given our focus on emerging economies, we take two different approaches. First, we use the New Keynesian small open economy model proposed by Berger at al. (2014) and Pedersen (2015). The model includes four measurement equations for the following observable variables: real GDP (in logs), inflation, real interest rate and real effective exchange rate (REER). The first three measurement equations decompose the observable variables (with no observation noise) into two states: a trend state and a gap state. The fourth one models inflation deviations from a long term “regime” level.

The trend (or potential) real GDP follows a local linear growth model, adding to the past level of trend GDP a random innovation and a potential growth state, which in turn is modelled as a random walk plus noise. The real GDP gap is derived from an IS curve that includes the past two lags of the GDP gap, the real rate gap and two lags of the REER gap, plus a random shock.

Inflation is determined by a Phillips curve that combines the past level of inflation, the GDP gap, real exchange rate appreciation and a long term “mean” capturing the structural break between different inflation regimes.
Real Output:

RGDP Trend and Gap decomposition: \( y_t = y_t^N + x_t^Y \)

Local Linear Growth model for RGDP Trend: \( y_t^N = y_{t-1}^N + g_{t-1} + \varepsilon_t^{Ny} \)

…where RGDP Trend Growth is Random Walk \( g_t = g_{t-1} + \varepsilon_t^g \)

IS curve for GDP Gap

\[ x_t^Y = \alpha_0 x_{t-1}^Y + \alpha_1 x_{t-2}^Y + \alpha_0^r x_{t-1}^r + \alpha_1^r x_{t-2}^r + \varepsilon_t^{xy} \]

Inflation:

Open Economy Phillips Curve by regimes:

\[ \pi_t = \pi_{regime} + \beta_\pi \pi_{t-1} + \kappa_\pi x_{t-1}^\pi + \beta_q \Delta q_{t-1} + \varepsilon_t^\pi \]

Real Interest Rates:

R. Rates Trend and Gap decomposition: \( r_t = r_t^N + x_t^r \)

Natural Rate from Trend Growth… \( r_t^N = r_{t-1} + z_{t-1} \)

…and from Random Walk transitory level \( z_t = z_{t-1} + \varepsilon_t^z \)

Real Interest Parity for R. Rates Gap… \( x_t^r = c x_{t-1}^q + h_{t-1} \)

…plus AR(1) Risk Premium \( h_t = \rho h_{t-1} + \varepsilon_t^h \)

Real Effective Exchange Rate (REER):

REER Trend and Gap decomposition: \( q_t = q_t^N + x_t^q \)

Natural REER is Random Walk (from PPP): \( q_t^N = q_{t-1}^N + \varepsilon_t^{Nq} \)

REER Gap is AR(2) \( x_t^q = \delta_0 x_{t-1}^q + \delta_1 x_{t-2}^q + \varepsilon_t^q \)

Note: The first equation in each set is a measurement equation, the other equations are state equations. The observation error or state innovation terms \( \varepsilon_t \) are Gaussian white noises with \( \sigma^2 \) dispersion. In the alternative model, trend growth is replaced by US Natural Rate in the Natural Rate equation.

In the second approach, we include the US neutral rate (as measured by LW) as an exogenous determinant of the other countries’ neutral real rate, instead of the domestic trend growth. The reason for this choice is that a significant part of the trend dynamics in the first framework tends to be captured by the random walk component, suggesting this is indeed capturing global factors (this was also noted by Carrillo et al., 2014).
4. Methodology

To solve the model, we use a Maximum Likelihood approach. We acknowledge that the parameter estimation would ideally be done with Bayesian methods, as Maximum Likelihood methods face the “pile up” problem biasing the estimates. Pedersen (2015) indeed follows Berger and Kempa (2014) to set up a Bayesian Estimation. We constrain the MLE optimisation transforming the parameter space with sigmoid transformations in order to rule out solutions that would be inconsistent with:

— Stationarity (e.g. AR(1) parameters <||1||)
— Theoretical sign restrictions (e.g. REER appreciation is contractionary)
— Reasonable ranges of parameter values as implied by preliminary regressions on trends and gaps from univariate HP filters

The third restriction implies that the objective of the estimation is not so much that of achieving correct inference on the ‘true’ parameter values, which would require more sophisticated techniques (see Stock and Watson, 1998 and Kim and Kim, 2013). It is rather that of finding a set of parameter values that can reconcile a standard HP distinction of trends and gaps into a complete and theoretically grounded econometric model. A synthetic outline of the estimation method is as follows:

— First, a PELT algorithm is used to identify inflation regimes (i.e., dummies to identify structural changes in inflation). The minimum length of a regime is set to 20 quarters.
— Set up transformations of parameter space from model to estimation scale. Theoretical and stationarity restrictions on parameter space are further refined with “reasonable range” bounds derived as follows:

  • Use univariate HP filters to get preliminary trends and gaps
  • Use simple OLS regression estimates between the preliminary trends and gaps to get preliminary analogs of model equations
  • Use a (5%-95%) confidence interval around estimated parameters to get restrictions for the parameter space of the betas
  • Use dispersion of residuals to derive lower bounds for parameter space of the sigmas, addressing “pile up” problem with their maximum likelihood estimation (MLE)

Restrictions on parameter space are enforced with logistic transformation. In the case of AR(2) restrictions, which involve dependence, we use nested logistic transformations:

— Prior for states’ initial values and prior variances
   • From preliminary trends and gaps, and estimated variances
— MLE: minimization of negative log-likelihood
Simulated Annealing: from 10 starting points extracted from uniform distribution between upper and lower bound of parameter space. Computations are parallelized on all cores available to the machine. The ten final estimates are aggregated with an average weighted by a transformation of their relative likelihood

— For filtering and smoothing, we apply Kalman recursive formulas.

5. Data

We use the above methodology to estimate the neutral rates for the Czech Republic, Hungary, Poland, Israel, Turkey, South Africa and Russia. As inputs we use seasonally adjusted quarterly data for core inflation from 1990Q1 to 2017Q3 (annualised first difference of the log CPI excluding food and energy) provided by the OECD, GDP (log multiplied by 100) taken from the IMF and the real effective exchange rate calculated by the IMF. For the nominal interest rates, we use the average quarterly interbank rates provided by the OECD when available. Where this data is not available, we used the discount rate also provided by the OECD. For the real interest rates, we take the difference between the nominal interest rate and inflation expectations for the next year. We calculate inflation expectations as the average inflation over the four quarters ahead from a univariate AR(3) of inflation estimated over a 36 quarter rolling window.

6. Results

6.1. Domestic trend growth model — CEEMEA neutral rates are moving in tandem

Using the first modelling approach with trend growth rates, we find that neutral rates have declined and then stabilised in most countries across CEEMEA, with the exception of Hungary, where neutral rates are still declining (Figure 1), and Russia, where neutral rates have reversed their decline in 2009 and continued to rise since (Figure 2).

The broad decline in neutral rates across CEEMEA (Figure 3) is in line with studies done on the developed world. In DM, most authors find that neutral rates in the US specifically, but also in Europe, started to decline at the beginning of this century, and that the decline accelerated during and post the Global Financial Crisis only to stabilise in recent years (see Holston et al., 2016).

— Neutral rates have fallen in CEE-4 similarly to DM: In the Czech Republic, Hungary, Poland, and Romania, neutral rates are now 0.0%, 1.3%, 2.5% and 2.3%, respectively. Neutral rates in Poland and Romania have especially declined from elevated levels, with most of the decline
**Figure 1.** Neutral rates have declined and stabilized in Israel, Poland and Czech, but are still declining in Hungary

![Graph showing neutral rates declining and stabilizing in various countries](image)

*Note: Baseline Model Neutral Real Rate*

*Source: Goldman Sachs Global Investment Research*

**Figure 2.** Neutral rates have declined and are stabilising in South Africa, Turkey and Romania, but continue to rise in Russia

![Graph showing neutral rates declining and stabilizing in various countries](image)

*Note: Baseline Model Neutral Real Rate*

*Source: Goldman Sachs Global Investment Research*

**Figure 3.** The overall decline in neutral rates in CEEMEA is consistent with findings across DM countries

![Graph showing overall decline in neutral rates](image)

*Note: Baseline Model Neutral Real Rate*

*Source: Goldman Sachs Investment Research; Holston et al. (2016)*
seen before the GFC, where neutral rates were lower in the Czech Republic and Hungary (Figure 3). The declines in neutral rates since 2007 have been of similar magnitude to those estimated for the Eurozone and the US, but larger than those observed for Canada or the UK. In Hungary, neutral rate estimates are now in line with the levels observed in the UK and Canada, but remain distinctly above those observed in the Eurozone or the US. In the case of Hungary, neutral rates are still falling and appear to be converging towards the level of the Eurozone and the US, while in the Czech Republic this convergence appears to have taken place already.

— The decline in neutral rates in South Africa, Israel and Turkey is of a larger magnitude. In South Africa and Israel, neutral rate estimates have fallen from around 9.0% to 1.1% and 0.4%, respectively, thus to levels similar to those observed in the DM world. In Israel, the decline was gradual, while the decline accelerated for South Africa following the GFC. In Turkey, neutral rates fell from 14% to 3%, with most of the decline taking place after the GFC. While the level in Turkey remains well above that in DM countries, it is closing in on those in some CEE countries.

— Neutral rates continue to rise in Russia. The real outlier in the region is Russia, where the neutral rate has risen sharply from the negative values recorded in the pre-GFC period. Neutral rates reached a minimum around mid-2008 of -4.0% and have now reached 4.0%, with a continued gradual increase. In our view, the reasons behind this peculiar trend involve a combination of the change in the dynamics of the terms of trade, together with a structural break in monetary policy from a heavily managed float to inflation targeting. The rising oil prices pre-GFC, together with a managed exchange rate, kept real rates negative, while the falling oil prices since the GFC are sustaining the growth of the neutral rate as they have led to a trend real depreciation in the exchange rate. However, this also implies that the neutral rate in Russia will decline if oil prices fully stabilise.

As Holston et al. (2016) note in the estimates for the DM world, though they use closed economy models, the rates clearly show evidence of co-movement, suggesting an important role of global factors in shaping the dynamics of trend growth and neutral rates across countries. Similarly Carrillo et.al. (2014) found that most of the dynamics of the neutral rate in Mexico are not driven by the countries trend growth, but rather by the unobserved random process. In turn they also conjecture that this random process appears to co-move with longer-term US rates.
6.2. Introducing the US neutral rate works well as a determinant of CEEMEA neutral rates

We complement the above analysis by including US neutral rates as estimated by LW into our model instead of trend growth. Overall, the results are similar to the baseline model, with Turkey being the only country with markedly different results. In Turkey, the second model approach suggests an even steeper decline in neutral rates, to 0.0% instead of 3.0%.

Figure 4. Using the US neutral rate instead of the trend growth rate yields similar results …

Note: Neutral Real Rate (second model)
Source: Goldman Sachs Global Investment Research

Figure 5. … with neutral rates stabilizing in most of CEEMEA, except for Hungary and Russia

Note: Neutral Real Rate (second model)
Source: Goldman Sachs Global Investment Research
While the results are broadly similar, the coefficients on trend growth are generally weaker than the coefficient for the US neutral rate, suggesting that the baseline model is being driven by global factors, rather than domestic ones:

— **The impact of trend-growth on neutral rates is relatively small:** In the baseline specification, we find a coefficient for trend growth onto neutral rates which is close to zero for most economies (Figure 6), with the exception of Romania and South Africa, where the beta is 0.8 and 1.9 respectively. Thus, the low beta’s suggest that most of the variation in neutral rates are unexplained (Figure 8).

— **We find a high beta for the US neutral rate in driving neutral rates across CEEMEA:** Across the CEE-4, the coefficient to US neutral rate is

**Figure 6.** The beta of neutral rates to domestic trend growth is low across the region...

![Figure 6](image1.png)

*Note: Figure shows the annualised coefficient of trend growth to neutral rates (baseline model)*

*Source: Goldman Sachs Global Investment Research*

**Figure 7.** … while the beta to the US neutral rate is relatively high across CEEMEA

![Figure 7](image2.png)

*Note: Estimates for Alternative Model*

*Source: Goldman Sachs Global Investment Research*
around 1.0 for the Czech Republic and Hungary, and around 2.0 for Poland and Romania (i.e. a 1ppt change in the US neutral rate lowered neutral rates in these countries by 90bps and 200bps, respectively; Figure 7). Comparing this result with the magnitude of the trend decline in neutral rates over the two in-sample shows that the US rate, scaled by its coefficient, can account for around 50% of the decline in all four countries (Figure 9). For Israel, South Africa and Turkey the betas are around 2.0, 2.3 and 4.1 respectively, and the change in the US neutral rate can account for 66%, 80% and 85% of the neutral rate declines in the respective full samples. If we include the covariance between the US neutral rate and the risk premium, we find that more than 75% of the variation in neutral rates can be explained. Again, the true outlier

**Figure 8.** The random walk component captures most of the variation in neutral rates in the baseline model …

Source: Goldman Sachs Global Investment Research

---

**Figure 9.** … while the US neutral rate explains most of the variation in the alternative model

Source: Goldman Sachs Global Investment Research
is Russia, as its peculiar dynamic of real rates does not correlate with US neutral rate dynamics, and thus results in an insignificant coefficient (Figure 7).

Thus, the two separate approaches suggest that the trend US rate explains most of the different degrees to which neutral rates in CEEMEA have declined in the past. This suggests that global factors, such as the US neutral rate, are more important drivers of neutral rates in a small open economy than domestic growth.

6.3. Combining Trend Growth with the US Neutral Rate

To fully understand which factors drive the neutral rate (i.e., whether this is driven by domestic trend growth or global factors) we model both the US neutral rate and the trend growth together, capturing the covariance between the two. However, when we model both together, we get a negative coefficient for the domestic growth rate which is not straightforward to interpret (Figure 10). Part of the reason why domestic trend growth and the US neutral rates are difficult to disentangle is because of a high collinearity between the two.

Considering the combined model, which includes both trend growth and the LW US neutral rates in the neutral real rate state equation, we find the domestic neutral rate depends on the US neutral rate and a trend risk premium. The trend risk premium turn is composed by a random walk component — capturing the risk premium level — and a correction on this level that depends on domestic trend growth.

In the combined model, the coefficients for the US neutral rate are confirmed in multiple cases (Figure 10). For most of CEEMEA, trend growth results in a
significant correction of the risk premium (i.e., the coefficient on trend growth is negative), leaving space for a higher multiplier on US neutral rates.

The variance decomposition (Figure 11) shows that the random-walk component significantly declines in the combined model. In Russia, it is clear from the variance decomposition that only the combined model manages to effectively reduce the share of variance of the residual, and indeed the coefficient on US

**Figure 11. Combined model — Neutral Rates Variance Decomposition**

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>CZK</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HUF</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PLN</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RON</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ILS</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TRY</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ZAR</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RUB</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Source: Goldman Sachs Global Investment Research*

**Figure 12. Russia’s Neutral Rate appears to be highly correlated with the terms of trade**

**Note: Neutral rate (baseline model)**

*Source: Goldman Sachs Global Investment Research, Russia Statistical Office*
neutral rates, which was insignificant in the baseline model, becomes a sensible (and significant) 1.4 when domestic growth conditions are accounted for.

In Russia, we cannot attribute any significant part of the variation in neutral rates to either trend growth rates or US neutral rates. In fact in Russia we find a negative co-movement between neutral rates and potential growth. The most obvious source of this anomaly is Russia’s more variable terms of trade. A persistent terms positive (negative) terms of trade shock will lead to upward (negative) revisions in the potential growth rate. To the extent that these persistent positive (negative) terms of trade shocks enter expectations of future terms of trade they will lead to expectations of a trend real appreciation (depreciation) and depress (increase) the estimate of neutral rates. Hence the negative correlation between trend growth and neutral rates.

Indeed, a long-run weighted average of the neutral rate and the quarter-on-quarter change in the terms of trade shows a high correlation between the two, with the most recent improvement in Russia’s Terms of Trade not yet being picked up by our model estimates of the neutral rate yet, which we suspect is due to the end point problem of our filter.

**Figure 13.** Measures of output gap is negative in Israel, Turkey and South Africa, negligible in Russia, and positive in CEE-4

*Note: Figure shows average of 2016Q4-2017Q4;*

*Source: Goldman Sachs Global Investment Research*
6.4. Comparison to Output Gaps

Our estimate for neutral rates also gives an estimate for output gaps and real rate gaps across CEEMEA. These output gap measures are trend estimates (distinct from the production function-type estimates one would normally use), and allow us to gauge where policy is accommodative/restrictive relative to the business cycle from a pure statistical framework:

— Output gaps are currently positive in the Czech Republic, Hungary and Romania, but negative real rate gaps, which suggests monetary policy here is accommodative
— Turkey, South Africa, and Israel have negative output gaps and negative rate gaps
— In Russia, the current real rate is significantly above the neutral rate despite the output gap being closed, which suggests Russia is the only economy in CEEMEA with a restrictive monetary policy relative to its business cycle

7. Robustness

From an empirical point of view, the above results are highly sensitive to the restrictions in the parameter space, which are usually binding. In particular, one important restriction concerns the ‘noise to signal’ ratios, that is, the ratio of the standard deviations of the innovations of the trend component to that of the innovations of the gap component. As already argued, the objective is not so much that of correctly estimating the true parameters, but rather to find a parametrisation
that yields trend and gap series that have dynamics close to the univariate HP filtering while interacting among them in a way that is coherently described by a theoretically grounded multivariate model.

From a theoretical standpoint, the most relevant assumption for the analysis of the natural rates dynamics is the specification of the trend real rate state equation. We have already shown results for two different options: the baseline model using domestic trend growth and the alternative model using US neutral rates as the main driver. However, it is interesting to see how the results hold up to a generalization including both factors in the state equation.

8. Conclusion

To conclude, using a methodology to estimate neutral interest rates close to that previously employed by the Fed’s Laubach and Williams (2016), we find that neutral rates have stabilized at relatively low levels in most CEEMEA economies. Among the low yielders, neutral rates appear to be at levels close to those found in small open developed markets, like the UK or Canada. The exceptions are the Czech Republic and Israel, where neutral rates have fallen even further (and appear close to the levels in the US and Euro area). Among the high yielders, neutral rates have fallen sharply in South Africa and Turkey — with the former now lower than the latter. The main exception to this trend is Russia, where neutral rates appear to have risen. We have suggested that this peculiarity may reflect the interaction of a shift in Russian monetary policy with the dynamic of oil prices.

Our estimates for the neutral rates closely mirror the dynamics of the US neutral rate, which appears to be a better determinant of CEEMEA neutral rates than domestic trend growth (implying that global factors are more important than local factors in driving neutral real rates). The exception (again) is Russia, where terms-of-trade dynamics appear to be more important than the US neutral rate.

9. References


Bouis, R., Rawdanowicz, Ł., Renne, J.-P., Watanabe, S., and Christensen, A. (2013). The effectiveness of monetary policy since the onset of the financial crisis, OECD.


Impact of Banking Supervision Enhancement on Banking System Structure: Conclusions from Agent-Based Modeling

Alexey Ponomarenko, Bank of Russia
Andrey Sinyakov, Bank of Russia

The Bank of Russia’s policy for banking sector recovery and elimination of non-viable and unscrupulous banks from it receives experts’ attention and their mixed response. Our study, in line with the key points of critique, suggests that this policy can, in the medium term, bring down the level of monopolism in the banking system and improve its efficiency, while weakening medium-sized and small banks’ positions in the short term. Overall, long-term benefits from proactive supervision policy may significantly outweigh short-term losses owed to a rise in banking business concentration. To study the implications of proactive banking sector recovery we have constructed an agent-based model of the banking sector and calibrated its key variables using Russia’s banking sector data. Based on the model we have compared short-term and long-term effects under two supervision regimes, one more and the other less stringent. Model estimation suggests that short-term effects of proactive supervision policy weaken the positions of medium-sized and small banks, including those complying with supervisory requirements. However, as the banking sector gains in soundness, benefits from heightened confidence in these banks and the entire banking system generally exceed short-term losses. The share of medium-sized and small banks in loans and deposits of the banking system eventually proves to be higher than prior to supervision policy enhancement. The monopolism of the banking sector declines and price competition rises. The banking system becomes a more effective economy financing institution, shedding the excessive risk to individual and systemic stability, with the average risk of financed projects remaining unchanged. Meanwhile, medium-sized and small banks’ stability improves. We point out that agent-based models provide a helpful tool for studying a wide range of other issues vital to the Russian banking sector.

Keywords: banking supervision, banking system cleanup from “bad” banks, agent-based modeling, Russia’s banking sector

JEL: G28, G21, E47, C63
1. Introduction

Proactive policy that the Bank of Russia embarked on in 2013 to secure banking system recovery and its cleanup from fraudulent players had 332 banking licenses withdrawn as of June 2017. Note that a total of just 390 licenses were withdrawn in 2002–2012. The Deposit Insurance Agency’s payments to depositors expanded from 0.02% of GDP in 2012 to 0.7% of GDP in 2015.

These drastic measures by the Bank of Russia to straighten things out in the banking sector have however received a mixed public response. The critique generally boils down to the four points below, which have to do with both short-term and long-term effects of proactive supervision policy1:

1) What critics view as the massive license withdrawal is causing depositors’ outflow to large and government-controlled banks, damaging smaller, regional banks which stick to fair business practices. In economics this is termed externality2, in particular when a “system contagion” by a depositor flight has an adverse effect on the system as a whole3.

2) The fall in the number of banks will have an adverse long-term effect on competition in the banking sector. Since it is from private banks that licenses are withdrawn, a risk arises of increasing imbalances in favor of government-controlled banks which may then start to dictate their terms to depositors and borrowers, i.e., offer their services as monopolists at higher prices (higher interest rates on loans and lower rates on deposits).

3) Also, it is assumed that the predominance of government-controlled banks will make the banking system more vulnerable to future shocks. As this school of thought sees it, government-controlled banks are less adaptive to shocks but, aware that they will be bailed out in any event, are inclined to assume greater risks4, thereby increasing the banking system’s vulnerability to crises.

4) It is claimed that in the long term this will entail a deterioration in the banking system’s efficiency as an institution providing credit to the economy. Less extensive and varied tiers of the banking system and a decrease in the number of market niches where investment projects (borrowers) of various risk levels can find their creditors, may cause a credit shortfall in the economy.

It should be mentioned that economists have yet to reach a clear conclusion as to whether a smaller number of banks adversely affects competition in the banking sector or what the optimal level of this competition is5. The impact of banking sector recovery/cleanup on competition, financial intermediation efficiency, and

---

2 E.g., De Nicolo et al. (2012).
3 In the critics’ view, a similar situation was seen in the Republic of Tatarstan after Tatfondbank’s license was withdrawn. See media coverage of implications for the republic’s banking sector: http://www.banki.ru/news/columnists/?id=9750266 (in Russian).
4 The so-called too-big-to-fail problem.
long-term economic growth is studied in the literature in the context of the effects of proactive banking supervision (see, e.g., Kupiec, et al, 2017). Studies of the effect of tightening bank capital requirements indirectly suggest arguments for proactive banking supervision. Colliard (2015) examines an optimal banking supervision model and provides a literature review.

In order to analyze the benefits and disadvantages of proactive supervision policy in Russia, we have constructed a formal behavioral model of partial equilibrium for the banking sector, under which each individual bank seeks to maximize profit. The model was constructed using the agent-based modeling technique which is gaining popularity in economics and central banks’ policy analysis. Unlike the widely used DSGE models, this approach makes it technically simpler to model complex economic systems with a large number of agents and their types. In real life, for instance, households are not homogenous but differ in their income levels and tolerance for bank deposit risk, investment projects differ in their default probabilities, while banks are divided into groups by their lending policy types, i.e., the banking sector segmentation is present. Moreover, such models have an important advantage of offering the opportunity to analyze the dynamic interaction of all types of agents and the way independent actions of individual agent types explicitly form a market equilibrium (produce macroeconomic effects).

We calibrated the model against the data of Russia’s banking system so as to reflect the degree of concentration and the structure of the banking sector’s assets and liabilities from the perspective of how banks were grouped prior to stepping up banking sector recovery (mid-2013). We ran the model thus calibrated to compare the main characteristics of a banking system equilibrium under the two different supervision policy regimes: that of proactive banking sector recovery followed by the maintenance of tough capital requirements and that of loose supervision standards, with a focus on the characteristics of competition and efficiency in the banking sector.

Of course, we cannot claim that this approach to calibration allows an accurate quantitative assessment of results achieved by toughening supervision policy. Also, we have no reason to suggest that assumptions we use are realistic and comprehensively reflect the impact of supervision policy on the banking system. We did not aim to pinpoint the most vital problems and realistic premises. Although this study takes account of financial intermediation’s fundamental

---

6 For arguments for and against banking sector recovery policy, see, e.g., Speech by Dr Raghuram Rajan, Governor of the Reserve Bank of India, at the Confederation of Indian Industry’s (CII) first Banking Summit, Mumbai, 11 February 2016.
8 See, e.g., the section outlining the experience and areas of application of this approach by the Bank of England: Turrell (2016).
9 In particular, we do not believe that toughening supervision policy will affect the probability of emergence of new (scrupulous) banks.
features, such as information asymmetry in the loan and deposit markets, the possibility of a depositor flight or a bank run, including the possibility of the liquidity problem (depositor flight) turning into that of capital shortage (insolvency), we do not touch upon a wide range of issues, including moral hazard and risks to financial stability caused by banking system consolidation (too big to fail), the probability and implications of zombie-banks emergence, the viability of bailing out banks at the taxpayers’ expense, a balance between supervision and insurance as factors of market discipline, and the accent of supervisory activities on certain bank types.

Empirical studies suggest that explanations for processes developing in the banking system probably exaggerate the role of supervision policy. However, our approach allows illustrating a wider range of hypothetical effects which may potentially arise as supervision is enhanced.

The main results that we have obtained are as follows.

**Short-run consequences.** In the short run, the vigorous cleanup of the banking sector from “bad” banks stems from a growing perception of risks involved in putting money in all medium-sized and small banks (confidence crisis) and more exacting demands on the size of a bank as an observable and simple test of a bank’s risk of failure. As a result, a deposit outflow to large federal banks and, to a certain extent, even an outflow of funds from the banking system in general (for example into cash or foreign currency) are seen. This is how the negative externality of proactive supervision materializes. This results in the higher average deposit rate on the balance sheets of medium-sized and small banks, depressing their profit from liabilities management, as they earn this profit through obtaining liabilities at a lower price than a central bank’s key rate (rate curve) or the money market rate (rate curve). The profit contraction on the liabilities side results in a low level of medium-sized and small banks’ capital, preventing their lending growth. Because of the medium-sized and small banks’ capital replenishment problems, this group of banks faces greater financial vulnerability in the short term. The major federal banks find themselves in the opposite situation, expanding their capital thanks to a deposit influx and a lower cost of liabilities, which helps expand lending. As estimates suggest, this will eventually boost the share of the major banks in the banking system’s total deposits and loans. A rise in concentration

---

10 See, e.g., the description of Japan’s economic situation, Koo (2011).
11 This study did not aim to quantify direct public losses, while the given model specification takes into account such direct public losses. Indeed, from the macroeconomic perspective (general equilibrium) bank depositors are also the owners of firms as well as banks themselves. As we charge bank losses to the capital of other banks which can absorb such losses, we actually recognize depositors’ direct losses.
12 The last two questions may become the subject of a further study as part of this model class.
is also captured by the increasing Herfindahl index, often used as its measure. Competition weakening in the banking sector is adequately reflected by the increasing heterogeneity (dispersion of deposit rates across banks): as the role of a bank’s size grows, price competition recedes into the background, and the major banks earn extra profits owing to the abundance and low cost of liabilities. These profits are actually redistributed to them from smaller banks and from depositors. Another important short-term implication of recovery is a deterioration in the efficiency of the banking sector’s operation, as shown by a drop in the bank’s average risk exposure. Although the maximum loan risk exposure declines owing to the exit of “bad” banks which used to assume such risks most often, the segment of average-risk investment projects receives less resources. This happens because medium-sized banks catering to this segment are unable to provide loans due to their problems with capital caused by the depositor outflows.

**Long-term implications.** Many of the above effects are not only reversed upon the completion of the banking system cleanup period and the subsequent retention of proactive policy by the regulator but the banking sector’s performance in the key areas improves compared with an equilibrium found in the absence of a banking sector cleanup:

A. **Medium-sized and small banks’ deposit base expands compared with a long-run equilibrium seen in the absence of banking system recovery policy, painful in the short-term.** The share of deposits in large banks not only returns to the previous (prior to stepping up supervision) equilibrium, but declines lower – medium-sized and small banks enjoy a higher long-term depositors inflow as compared with the equilibrium under loose policy by the regulator. A reduction in the number of banks with negative capital enhances depositor confidence in medium-sized and small banks and the banking system in general. Having become generally more reliable and jettisoned the ballast of “bad” banks, the medium-sized and small banks shows a greater appeal to depositors than in the absence of banking system recovery by the regulator. As a result, the share of deposits in large banks dwindles in the long term.

B. **Medium-sized and small banks strengthen their capital positions, thereby boosting lending.** The share of medium-sized and small banks in new total lending grows. The key implication of depositors reversing their flow back to responsible medium-sized and small banks is their profit growth and hence recapitalization. Medium-sized banks’ rising potential to earn profit from liabilities translates into their growing ability to make profit

---

on the assets side. This enables nonfederal banks to boost lending to investment projects (facing capital constraints more rarely). Similarly to deposits, capital and lending expansion by medium-sized and small banks in a new long-term equilibrium surpasses levels seen in an equilibrium under loose policy by the regulator.

C. **The major banks’ monopolism decreases compared with a long-term equilibrium under loose supervision policy: profit is redistributed from the major banks to medium-sized and small ones.** A long-term rise in confidence in medium-sized and small banks driven by a vigorous cleanup of the banking system diminishes the role of a bank’s size as an important factor in choosing a bank for opening a deposit account. It is price competition and the quality of services that come to the fore. As medium-sized and small banks join this competition for depositors (from the position of “equality” with large banks) competition strengthens, leading to higher average rates on deposits in the banking system and more homogenous interest rates across banks. The Herfindahl index showing the concentration of the deposit and loan market drops by half.

D. **The banking system becomes a more readily available economy financing institution: excessive (frequently associated with asset stripping) credit risks are driven out of the system, while the share of bank-financed investment projects showing a moderate level of risk increases. Meanwhile, the system’s average credit risk remains unchanged.** Since medium-sized banks of the model on average pursue a softer lending policy than that of the major banks, the expansion of their share in total lending means greater availability of loans in the banking system, where minimum-risk borrowers, which as a rule implement the largest investment projects, are not the only ones that enjoy the opportunity to raise financial resources.

E. **Financial stability of the medium-sized and small bank group improves.** The share of banks with negative capital is reduced to zero in the long run, whereas it permanently remains unchanged under loose policy\(^\text{15}\). The system sheds the overhang of “bad” banks that used to persist previously, threatening to spark a domino effect and a massive crisis of confidence in scrupulous small banks as well as in the entire banking system even upon a minor shock. Medium-sized and small banks strengthen their capital positions and gain a fairer (more competitive) access to the deposit and

\(^{15}\) The model’s implicit estimate stands at 8% of banks at each moment in time. Note that it is not an aggregate total, because banks which have negative capital at a specific moment in time may subsequently recover if circumstances are favorable.
loan markets. The system as a whole becomes more sound than it would have been if there had been no cleanup of the banking system.

Further presentation of the paper is structured as follows. We will first provide a description of the model, setting rules of behavior for each of the four groups of agents, i.e., households, enterprises, banks, and the regulator (central bank). We will then report the main results of simulation of the model and, finally, will set out our conclusions regarding central bank policy.

2. The model and its calibration

This section presents only a brief description of the logic and structure of the model. A detailed description of the model and parameters calibrated is provided in the Annex.

We need to point out that the model makes use of agent-based modeling which economics took over from natural sciences. For the employment of agent-based models in economics and finance, including in central banks’ policy analysis, and their benefits and drawbacks, see Turrel in the Bank of England Quarterly Bulletin (2016) and Fagiolo & Roventini (2017).

The agent-based modeling requires the behavior of various types of the model’s economic agents to be defined, i.e., the way they make decisions in various situations (households or firms) – on a bank for opening a deposit account; banks – on whether or not to provide a loan to a company). This process aims to achieve the approximation of the behavior thus described to agents’ behavior in real life. Rules of behavior are therefore set for each individual agent in the model. Agents, on their part, vary in parameters, and agents’ heterogeneity is thus modeled. These parameters are assigned to agents as random values from a distribution with given characteristics. Agents’ emergence and default, the occurrence of various shocks, with which agents have to cope, are governed by probability laws. As a result their behavior patterns are studied via a multitude of simulations in which the law of large numbers manifests itself.

Our model constitutes a basic description of the loan and deposit markets in which banks operate. It can be regarded as a simplified version of the banking sector model presented in Chan-Lau (2017). Non-banking sectors are not however modeled, while demand for loans and deposits is determined by exogenous processes. The overall chart of the banking sector model is presented in Figure 1. In each period, agents act in the following order:

1) Borrowers declare default.
2) Banks with negative capital have their licenses withdrawn and are taken over by other banks.
3) Banks go over deposit rates and decide on their optimal level.
4) Depositors update their characteristics and make deposit placement decisions.
5) Interest on loans and deposits is paid and loan principals are repaid.
6) New projects are generated. Banks offer financing and add new loans to their loan portfolios.
7) Interest on interbank market liabilities is paid.
8) New banks are established.

We will consider agent types in the model and their operation in more detail below.

2.1. Banks

The key block of the model comprises banks which take deposits from households and firms and provide loans to companies showing various levels of return and matching default probabilities. We identify two sources of profit for the banks of the model: taking deposits at a lower than the key rate and providing loans at a higher than the key rate. Meanwhile, decisions on liabilities management in the model are generally not related to banks’ decisions on assets management. This reflects the idea that to obtain liquidity for issuing loans, banks can always go to the interbank market or turn to the central bank, but they need deposits only if these are less expensive than funding from the interbank market or from the central bank\textsuperscript{16}. The only constraint on lending expansion by the banks in the

\textsuperscript{16} That said, we disregard the interest risk issue, as we believe that the central bank’s policy is transparent enough and enjoys enough market participants’ confidence for all participants’ interest rate expectations to be anchored to the same shape of the interest rate curve.
model, if such new loans are in line with the bank’s lending policy regarding the risk level, is compliance with the capital adequacy requirement (an analog of H1). We therefore disregard the liquidity management issues.

Banks differ in the size of their capital (and, accordingly, assets) and lending policy types. It is assumed that at the initial point in time the small banks group has the loosest lending standards\(^ {17}\), issuing loans to high-risk projects. One feature of the model with respect to this third group of banks is that these banks continue issuing loans even upon attaining compliance with the capital adequacy requirements or breaking them.

Banks are modeled as follows. The economy contains a population comprised of \( j \) banks where each individual bank belongs to one of three subgroups. A bank’s balance sheet looks as follows:

<table>
<thead>
<tr>
<th>Assets</th>
<th>Liabilities</th>
</tr>
</thead>
<tbody>
<tr>
<td>Loans</td>
<td>Deposits</td>
</tr>
<tr>
<td></td>
<td>Net interbank liabilities</td>
</tr>
<tr>
<td></td>
<td>Capital</td>
</tr>
</tbody>
</table>

A bank earns interest on its assets and pays interest on its liabilities. The spread between these flows makes up its profit and capital increase. At each moment in time banks maximize profit, which is derived from the two sources below:

1. Profit from transactions in liabilities. A bank takes household deposits and places these funds in the interbank market (in a central bank), earning a spread between the key rate and a rate on deposits.
2. Profit from transactions in assets. A bank issues loans, obtaining liquidity for them in the interbank market (from a central bank). The profit is formed from the spread between a rate on the loan and the key rate.

The size of a bank is defined as a share of its assets in the total assets of the entire banking system. A rate on net interbank liabilities corresponds to the \( i \).

A bank sets rates on its products as follows:

- Loan rate variation is determined based solely on the risk level \( f_m \) of a project and is determined as \( i + \gamma_0 + \gamma_1 f_m \).
- In offering its rate on deposits, a bank seeks to outplay its competitors, bearing in mind that, other things being equal, a lower rate triggers a depositor outflow, while a higher one deprives a bank of a proportion of profit. The profit maximization by a bank on loan and deposit rates is modeled as an iterative process. When, in the first step, a bank has the opportunity to model its decision on the rates, it tests one of five strategies (running values in the range from low to high rates on deposits). Each

\(^{17}\) This condition holds only for the initial distribution of banks. In equilibrium, the initial bank size and its lending policy are not correlated as much: small banks have a chance of growing into large ones, while large banks may, under adverse circumstances, lose assets and become small even if they pursue prudent lending policy. We decided on classification by bank size rather than by lending policy, because, unlike the portfolio risk, size is an observable characteristic which helps depositors to make a decision.
bank offers a deposit rate from the \((i - \xi, i)\) range, which, given the current characteristics of other banks’ depositors and deposit rates would secure it a maximum profit.

As regards risk exposure and compliance with capital requirements, all the banks in the model are divided into three groups:

1) Large federal banks;
2) Medium-sized federal and regional banks;
3) Small regional banks.

The main difference between these groups of banks lies in their lending policies\(^{18}\). At the initial point in time, banks differ in size and their maximum tolerable default risk exposure for a borrower receiving loan, \(f_j^*\). Note that the characteristic of size can change depending on a bank’s performance, while risk tolerance of its lending policy remains unchanged.

The major federal banks are assumed to be the least inclined to finance high-risk investment projects, selecting only low-risk projects. As the model parameters were calibrated, roughly the top-ten banks were attributed to this group.

The banks from the second group are less particular in selecting investment projects and show a higher tolerance for project risks. In calibrating the model, the top-100 banks, excluding the top ten ones, fell under this category. The banks from the third group are prepared to finance any investment projects, regardless of their risk levels. These third-group banks are distinguished by their willingness to continue financing companies despite breaching the capital requirement (or even negative capital). These banks are found outside the top-100 group in the model.

Therefore, banks within each of these groups vary in size (that of assets and capital), while groups themselves are also different in terms of the size of banks within them. At the initial point in time, the banks from the first group are the largest ones, the second group includes smaller banks and the third group is represented by small banks. In calibrating the banks groups we included the top-ten Russian banks (operating a large nation-wide branch network) in the first group, the second group comprised top-100 banks, excluding top-10 institutions, the third group was made up of banks outside the top-100.

Apart from grouping banks by their lending policies, we assume that the loan market is segmented, meaning that not every bank can vie for lending to any emerging project. To reflect this effect, it is assumed that a bank can only offer financing to a newly emerged borrower with a probability of \(P_L\).

A bank ceases lending if its capital to loans ratio falls below the \(rr\) threshold. If a bank’s capital becomes negative, it loses its license with a \(P_B\) probability (determined by the sternness of supervision policy) and will stop operating (its balance sheet will be merged with that of another, randomly selected, bank

---

\(^{18}\) We note that this premise is not crucial to the general conclusion about the impact of supervision policy enhancement on competition. For details see Appendix 2.
unless this merger causes this new bank to fail to comply with the capital adequacy requirement).

At each moment in time, a new bank emerges in the system. The probability of an emergence of a third-group bank depends on the sternness of the regulator’s policy (the probability of the regulator visiting this bank for supervision). The tougher the policy, the lower the probability of new third-type banks emerging and the higher the probability that banks of the second type will emerge. Thus, in every period, a new bank of subgroup II is established with a \( P^{S1} \) probability and a new bank from subgroup III emerges with a \( P^{S2} - \delta_1 \times P^{R} \) probability. Random \( K \) deposits are invested in their capital, and net interbank liabilities appear on the assets side of their balance sheets.

2.2. Deposit market

On the liabilities side, banks compete for depositors, trying to attract them with the most lucrative rates (as long as these remain lower than the key rate). Depositors, on their part, differ as regards their risk tolerance. Risk-tolerant depositors place lower demands on a bank’s size (as an observable test of a deposit risk) than those less risk-tolerant. A key feature of such depositors’ behavior is that the news of a bank’s failure (its license withdrawal) heightens demands on a bank’s size and lowers confidence in the banking system. License withdrawal therefore triggers both deposit migration to other banks and deposit flight from the banking system (to alternative asset types outside the banking system, such as ruble or dollar cash), if demands on banks’ size increase substantially.

Formal modeling is performed in the following manner. There are \( n \) depositors on the deposit market, households or firms under the model. At each moment in time, depositors have options as follows:

1) Put money in the bank provided that at least one bank offers higher than required return in terms of rubles and the bank meets their minimum requirement for its size as shown by the size of its assets. The bank’s size is in this case an observable measure of this bank’s deposit risk. In this manner, all depositors are implicitly subdivided into two groups: risk-tolerant and risk-averse ones. Risk-averse depositors’ minimum requirement for a bank size will be higher (their goal is money safety rather than income).

---

19 This assumption reflects not only a negative news flow from license withdrawals but also depositors’ concern that their bank may also have its license withdrawn. As a result, they would have to spend time to recover their money from the Deposit Insurance Agency, and what is more, less than the entire amount of the deposit with accrued interest.

20 As a result, since households are firm owners and transactions in goods and services eventually place money in their hands, we can only refer to household deposits. To separate independent decisions by household in their respective capacities of depositors and firm-owning borrowers, the model separately examines households and companies — investment projects.

21 A bank’s capital in another version of the model.

22 Depositors thus take into account the lower probability of larger banks’ failure. Relying for instance on the adage “too big to fail”.

2) Use alternative investment opportunities (investments outside the banking system, including in cash).23

Every depositor has two characteristics:
— The minimum desired interest rate on deposits $i^D_n$.
— The minimum desired bank size $s^*_n$.

One important feature of the model is that depositors’ demands on a bank’s size depend on their level of confidence in banking system stability. In the model, bank failures (or license withdrawals) play a role of negative signals coordinating agents’ behavior. The stronger the signal (the more extensively bank failures occur), the lower the level of confidence in the banking system and the higher depositors’ demands on banks’ size. This mechanism allows modeling actually occurring migration of depositors from smaller to larger banks (from small regional to large regional banks and from large regional to federal banks) and out of the banking system (if a minimum desired interest rate at a bank of required size happens to be insufficient. This migration proves to be an important channel for bank failures to influence the liquidity situation and profitability, and hence compliance with the capital requirement at all banks of the same group by size. The cleanup of the banking system and declining occurrence of failures in turn improve confidence in banks and bring down demands on bank size, so there are increasingly more depositors willing to put their money in small banks offering high interest rates.

In each period, the parameter of the minimum desired bank size is adjusted to take account of the negative news flow $D_t$, which depends on the number of banking licenses withdrawn over period $L_t$. As a result, the current minimum desired bank’s size $s_{n,t}$ is determined as:

$$s_{n,t} = s^*_n + \alpha_1 D_t$$

$$D_t = L_t + \alpha_2 D_{t-1}$$

At each moment in time, bank deposits change owners so that the model could reflect transactions occurring in the real economy. Since the lending process automatically creates deposits, lending to companies increases household deposits in the model. This also aims to reflect the use of loans by companies in real life to pay for goods and services, with these funds eventually making their way onto household deposits (including salary deposits). In this context, to reflect the change of deposit holders in each period, parameters $i^D_n$ and $s^*_n$ are chosen randomly from the relevant distribution.

In each period, depositors with a $P^D$ probability revise their choice of a bank to hold their deposits. In this case they choose a bank with the highest interest rate from banks of a larger size than $s_{n,t}$. If this rate does not exceed $i^D_n$, the depositor decides to hold its money in foreign currency cash.

23 We do not look at investments in foreign currency, assuming that in an equilibrium adjusted for exchange rate movements returns prove to be equal.
A bank which has expanded (reduced) its deposits as a result of this deposit migration, proportionally decreases (increases) its net interbank liabilities.

The initial equilibrium (under a loose supervision policy) of depositors among the above banks groups was selected so as to match the distribution from the actual Russian banking system data prior to the supervision enhancement (the 2011–2013 average). As a result, we received the 55-35-10 distribution which reflects the share of household and corporate deposits in the Top 10 banks, the Top 100 banks excluding the Top 10 banks, and the rest of the banks by asset size as of March 2013.

2.3. Loan market

On the assets side, banks make a decision on whether to approve or decline loan requests from projects with various levels of return and a matching level of risk (which all the banks are able to observe). To simplify estimation, it is assumed that all the banks assess risk accurately. We disregard banks’ competition on the loan market via interest rates, simply assuming that a bank charges a rate on a loan which it approves. The rate is positively correlated with the level of risk (return) for this borrower’s investment project. This assumption aims to simplify computations without substantively changing the results. The important feature of the model, which brings it close to reality, is that loans issued by the banks are transformed into deposits. The model therefore reflects the economy’s money transactions to buy goods and services, in the course of which money changes owners within the banking system.

Formal modeling of the loan market is performed as follows:

Every company presents a risky investment project financed via bank loans. In each period, there appear investment projects for which borrowers want to get loans. Of these, projects are low-risk and projects are high-risk ones.

Projects have two characteristics, their distribution varying depending on the project category.

The first is the probability of default . At each subsequent step, there is such a probability of the borrower’s default, which would result in the relevant drop in a bank’s loan portfolio and capital. A high project default probability does not necessarily imply default. In the event of default, a bank fails to receive return on investment and the loan is not repaid. The value of collateral is assumed to be equal to zero and the entire loss is charged to the capital of the bank that financed the project24.

The second characteristic is the project rate of return . If there is no default, the company repays the loan and pays interest rate which the bank charged on the loan as it was provided.

24 This assumption only emphasizes that the lending process is not risk-free, the value of collateral cannot compensate a loan loss. Losses will still be impossible to avoid, even bearing in mind that compensation is possible for a certain percentage of loans.
The amount is determined as a share (\( \beta_1 \)) of the current money supply (the sum of all deposits and currency in circulation). The loan maturity is equal to 30 periods.

A borrower chooses a bank to finance the project randomly (with a probability proportional to the bank size) among banks willing to provide the loan (this is determined by the market segmentation and the bank’s risk tolerance). If the rate charged by the bank is above \( i_L^m \), no loan is issued and the project ceases to exist.

When a loan is provided, the size of a randomly chosen deposit increases by the relevant amount (for money supply creation through lending, see McLeay et al., 2014). The bank providing the loan increases its net interbank liabilities, while the bank receiving a newly-created deposit reduces them.

2.4. Regulator

And, finally, the model comprises a central bank which carries out an inspection of a bank with a given probability. It is assumed that upon finding negative capital the regulator withdraws the bank’s license. As a result, all the bank’s assets and liabilities go to another (randomly selected) bank whose capital allows it to absorb the loss\(^{25}\). At each moment in time, a small number of banks are established in the system. As a new bank is established, its founder is aware that the tougher the regulator’s policy, the shorter this bank’s life cycle will be should it fail to comply with capital requirements. So proactive supervision diminishes the probability of the newly-established bank being ranked as belonging to a “bad” bank type (the third group of banks).

The model reflects this in the following way. The central bank inspects this bank with a certain probability, which takes two values in the model: 0.2 prior to the autumn of 2013 and 0.8 after the autumn of 2013. The regulator’s visit to the bank under supervision enables breaches of capital requirements to be detected with certainty. If negative capital is found, the bank’s license is withdrawn and it is taken over by another bank (see above). If capital requirements are complied with, the bank continues operating as previously, while if a breach is discovered and capital is positive, banks from the first and second groups cease lending.

The regulator’s role in the model also consists in providing liquidity to banks as part of policy to manage interest rates of the money market (which is in this case identical to the central bank’s liquidity management instruments).

2.5. Initial conditions and simulation procedure

The parameters of every agent are drawn randomly from given distributions. The main calibrated parameters of the model can be found in Appendix 1. The

\(^{25}\) The maintenance of “bad” assets on the Deposit Insurance Agency’s balance sheet is essentially the same thing.
calibration of the model’s parameters was chosen so as to obtain in a large number of simulations equilibrium parameters close to those actually observed in Russia’s banking sector prior to stepping up supervision (the first half of 2013). These are the following structural parameters: deposit/loan distribution across the above groups of banks and the Herfindahl index used to assess the degree of concentration of the loan and deposit markets (see Appendix).

Thus, for initial conditions, a population of \( NB \) banks is generated whereas the calculation of the Herfindahl index assumes that each bank involved is the aggregate of the balance sheets of \( nb \) banks.

There are \( ND \) deposits of size \( D \) and \( NL \) low-risk loans of size \( L \) on the balance sheet of each bank. The size of capital is set at \( rr0 \) of total loans. Net interbank liabilities offset the balance sheet.

Also, \( NC \) depositors are created that hold \( C \) worth of financial assets in foreign currency cash.

Since the model is stochastic, the results obtained in different simulations will differ. However, with a large number of simulations, as various effects are averaged, systemic characteristics of an equilibrium should emerge in keeping with “the law of large numbers”. It is the comparison of these averaged systemic effects under the two different policy regimes that will produce our results.

We ran two 100-simulation series and, based on this calibrated model compared the principal characteristics of a banking system equilibrium under the two different central bank policy regimes.

The first policy regime is that of proactive banking sector recovery followed by the maintenance of proactive supervision policy and tough capital requirements. This implies a high probability of an inspector visiting a specific bank and discovering its problems, if any.

The second regime is the retention of a “loose” supervision regime, maintaining a low probability of a bank being inspected. This is equivalent to a low probability of detecting noncompliance, if any. This regime is assumed to have existed prior to mid-2013.

After setting the initial conditions and the beginning of the system’s dynamics in each simulation, we allow the system to “spend its life” in the two modes: under loose or proactive policy by the regulator. The system’s life consists of 250 periods (our calibration is so structured that we are unable to say whether these are months, quarters, or years). During the first 100 periods\(^{26} \), the system lives under the loose policy regime so as to let its stable properties become visible and allow the effects owed to the random setting of some initial conditions (for example, a specific sample distribution of projects by the risk level) to fade. Then, in 50 periods, the moment of policy change comes: the period of the banking system’s proactive recovery starts (Figure 2), following which the system is allowed to “live” another 100 periods, of which the first 30–50 can be regarded as a short-term period, and the following 50–70 — a long-term one.

In this case, comparison of the median under the two policy regimes based on a large number of simulations shows systemic differences. We set the number of

---

\(^{26}\) The figures below will present the dynamics of variables based on this set of periods.
simulations to 100, which, judging by the stability of the results, is sufficient for the “law of large numbers” to come into effect.

3. Results

Figures 2–11 present the averaged dynamic trajectories of the system’s key parameters prior to the start of banking sector cleanup and their dynamics after launching this policy. Short-term and long-term periods are presented separately. The short-term period is a time span from the 50th period (the time when banking supervision was stepped up) to the 80th–100th periods. The subsequent effects, until the 150th period, are regarded as long-term consequences of banking sector cleanup emerging in a situation of the banking sector’s adaptation to the regulator’s new policy. Along with them we presented 25-percent and 75-percent percentiles of these parameters’ distribution in the total number of simulations. The results of the simulations without supervision policy stepped up are shown in blue color and those with stepping up supervision — in red color.

Further on, we will make a distinction between short-term and long-term effects. We will first present a description of short-term implications for the banking sector.

3.1. Short-term effects of stepping up supervision

As Figures 2 and 3 show, enhanced supervision secures a fairly fast cleanup of the banking system from unscrupulous banks, which, in our case, are banks with negative capital.

**Figure 2.** Share of banks with licenses withdrawn over four periods in the total number of banks at the initial point in time

Source: authors’ estimates here and further
The short term sees increased risks of placing funds in all medium-sized and small banks with demands heightening on a bank’s size as an observable and simple test of bank failure risk. This causes the migration of deposits to large federal banks and, to a certain extent, even an outflow of funds from the banking system, for example, into ruble cash or foreign currency. As a result, the average rates on deposits on the balance sheets of medium-sized and small banks go up, diminishing their profit from liabilities management. Profit contraction implies a lower level of capital of medium-sized and small banks, preventing their lending growth. Due to the problems with capital replenishment by medium-sized and small banks, which under the model also on average pursue a riskier policy than large federal banks, this group of banks proves to be more financially vulnerable in the short term. The major federal banks find themselves in the opposite situation, expanding their capital thanks to a deposit influx and a lower cost of liabilities, which helps expand lending.

We estimate that eventually the share of the major banks in the banking system’s total deposits (Figure 4) and loans (Figure 5) will rise.

A rise in concentration is also captured by the increasing Herfindahl index, often used as a measure of concentration for deposit and loan markets (Figures 6–7). The weakening of competition for depositors is reflected by dropping rates on deposits in the banking system (Figure 8).

\[27\] This profit is earned by obtaining liabilities at a lower cost than the key rate (rate curve) of the central bank or the money market rate (rate curve).
Figure 4. Share of deposits in the first banks group (large banks)

![Graph showing the share of deposits in the first banks group (large banks).]

- **Strengthening oversight policies**
- **Unchangeable oversight policy**

Figure 5. Share of loans in the banks of the first group (large banks)

![Graph showing the share of loans in the banks of the first group (large banks).]

- **Strengthening oversight policies**
- **Unchangeable oversight policy**

Figure 6. Herfindahl index for deposit market

![Graph showing the Herfindahl index for deposit market.]

- **Strengthening oversight policies**
- **Unchangeable oversight policy**
The weakening of competition in the banking sector at first causes deposit rate nonhomogeneity (variations across banks) to rise slightly (Figure 9).

A higher rate dispersion reflects banks’ heterogeneity as regards their competitive positions on the deposit market. In a situation of depositors’ migration, large banks can afford to offer lower rates on deposits than all the others. By contrast, small banks, in the face of depositor outflows, have to offer higher rates as compensation for risk. So the heightening of the role of a bank’s size causes price competition to recede into the background, and the largest banks earn extra profit thanks to the abundance and low cost of liabilities. This profit is actually redistributed to them from smaller banks and from depositors. The heterogeneity suggests that in addition to price (a deposit rate) there are some other factors which affect depositors’ preferences. In a competitive banking system, these non-price factors (such as differences in banks’ size) are supposed to play a minor role.
Another short-term result of supervision enhancement is lower availability of loans to borrowers with a relatively high default risk, as the banks’ average credit risk exposure (Figure 10) and maximum credit risk exposure (Figure 11) decline. The maximum credit risk exposure declines because the group comprised of unscrupulous banks which pursue an extremely soft lending policy is forced out of the market. The average risk drop occurs because medium-sized banks catering to high-risk borrowers are unable to provide loans due to their problems with capital caused by depositor outflows. Meanwhile large banks (based on the model’s assumptions) are less inclined to lend to fairly risky projects despite these banks’ capitalization growth.
3.2. Long-term implications of supervision enhancement

One important result of proactive supervision is that in the long term the number of license withdrawals falls to virtually zero after the banking system's cleanup versus the persistence of roughly the same number under loose policy (Figure 2). This is not so much a result of proactive policy as that of the system itself getting more sound — the regulator's scope of operation narrows.

Indeed, as proactive supervision policy also hampers the emergence of “bad” banks, the share of banks with negative capital declines to zero in the long term versus its staying permanently at around 8% under a loose policy (Figure 3). The system eventually becomes more sound, shedding the formerly permanent “overhang” of “bad” banks threatening with a domino effect and a massive crisis of confidence in honest small banks and in fact the entire banking system following an even minor news-generated shock. Hence the financial stability of the system as a whole improves.

The drop in the number of banks with negative capital and license withdrawals enhances depositors’ confidence in medium-sized and small banks as well as the banking system in general. As a result, depositors’ demands on a bank's size (as a test of deposit risk) lessen, with depositors starting to migrate both to the banking system and back to medium-sized and small banks from large ones. This causes the share of deposits in the banks of the first group, i.e., the major banks, to shrink around 15 percentage points (Figure 4). The key result is that the share of deposits in the major banks not only returns to the previous equilibrium but declines lower than it used to be prior to policy toughening, with medium-sized and small banks

![Figure 11. Maximum loan default risk](image-url)
enjoying a rising long-term depositor inflow compared with the equilibrium under the regulator’s loose policy. Having become more reliable and jettisoned the ballast of “bad” banks, this group of banks shows a greater appeal to depositors compared with a situation where the regulator fails to clean the banking system from unscrupulous participants.

The Herfindahl index shows a similar long-term trend, pointing to a concentration downturn on the deposit side (Figure 6). The deposit market concentration drops practically by half.

The by-effect of depositors reversing their flow back to responsible medium-sized and small banks is these banks’ profit growth and hence their capital expansion (recapitalization). Medium-sized banks’ rising potential to earn profit from liabilities translates into their growing ability to make profit on the assets side. This enables nonfederal banks to boost lending to investment projects (facing capital constraints more rarely). Similarly to deposits, capital and lending expansion by medium-sized and small banks in a new long-term equilibrium surpasses levels seen in an equilibrium under loose policy by the regulator. As a result, the share of medium-sized and small banks in total lending goes up by around 10 percentage points (Figure 5).

The industry concentration downturn is reflected by the Herfindahl index for the loan market declining below its equilibrium seen in the absence of a vigorous banking sector cleanup (Figure 7).

The long-term confidence growth with respect to medium-sized and small banks as the banking system is cleaned from weak players diminishes the role of a bank’s size as an important factor for depositors in choosing a bank for opening a deposit account. Price competition comes to the fore, with the quality of service also taken into account. As medium-sized and small banks join this competition for depositors (from the position of “equality” with large banks) competition toughens, leading to higher average rates on deposits in the banking system (Figure 8).

The banking system becomes more homogenous in the process. The fall in the deposit rate variation across banks is clearly pronounced (Figure 9), with the cross-bank rate dispersion dropping by half. This homogeneity adequately indicates the toughness of price competition, as it shows the low significance of other factors securing advantages for some (large) banks as regards the cost of deposits. The deposit rate increase along with the rise in rate uniformity means that extra (monopolist) profit from the low cost of deposits to large banks, which is typical of an environment where “junk” banks operate, is redistributed in favor of medium-sized and small banks in the new equilibrium emerging after the cleanup of the system.

Banks’ maximum risk exposure (maximum default probability of all projects approved by banks) stays at a lower level in a long-term equilibrium (Figure 11). This shows that the banking system does not assume excessive risks after the
removal of banks using such practices, with their managers often knowing eventual results (fraud and asset stripping) in advance.

Meanwhile the banks’ average risk exposure (weighted average default probability for approved investment projects) is restored to the levels found under loose supervision policy (Figure 10). As banks’ maximum risk exposure decreases, this rise in the average risk exposure reflects an increase in the proportion of lending by medium-sized banks which provide loans to projects with a moderate but higher risk than those financed by the major federal banks.

Another important conclusion from Figures 10–11 is that despite the average risk staying on the same level due to the expansion in the share of loans from riskier banks, the regulator’s proactive policy protects the system from accumulating financial stability risks (a larger number of banks breaching the capital requirement).

Finally, the fact that the medium-sized banks in our model tend to practice looser lending standards than the largest ones, their increasing share in total lending also means greater availability of lending, where it is not only minimum-risk borrowers that have the vast opportunities to raise financial resources. Excessive risk meanwhile is driven out of the banking system: such projects are rarely approved. Conventional banks are not an important source of financing high-risk projects (venture financing) throughout the world. Hence our result is quite realistic in this respect.

4. Conclusion

We have built up an agent-based model for the banking sector to analyze the impact of proactive policy of the banking sector cleanup on the main characteristics of the banking system, including the parameters of competitive environment. Our general conclusion from the study is as follows:

1) The enhancement of banking supervision offers a variety of important long-term benefits. It promotes competition, which eventually benefits scrupulous banks, improving their financial stability. Benefits are enjoyed by both depositors and borrowers, especially small and medium-size enterprises (SMEs).

2) To lessen short-term negative effects, banks’ negative influence on one another needs to be minimized.

3) In light of our findings, it is especially important that during the period of proactive recovery of the system small and medium-sized banks send their customers signals that their respective banks are ranked as “good” rather than “bad”, unscrupulous ones. This, in particular, can be achieved through a more open information disclosure by banks relying on fair business practices, their greater willingness to secure transparency, especially regarding compliance with the capital requirements and engagement in financing the real economy.
4) Since greater openness does not in itself eliminate the problem of disclosing false information on banks’ operation by fraudulent market participants, this places more stringent requirements for rating banks, including by the regulator, in terms of financial reporting quality, followed by public disclosure of this information. The focus should be placed on giving the public unbiased information about credit institutions, without fear that disclosure of this information could entail problems for the banking system or trigger depositor flight from some banks.

This paper does not examine the entire range of potential effects which may accompany proactive supervision policy. In particular, the issues of moral hazard and risks to financial stability stemming from banking system consolidation (too big to fail) and strengthening of banks with government stakes, the acceptability of bailing out banks at taxpayers’ expense, a balance between supervision and insurance as factors of market discipline, orientation of supervisory activities to certain bank types, along with issues relating to the origin, spreading and impact (including systemic one) of bank panic. Moreover, this study did not aim to model a great multitude of other, no less important, problems: the influence of the money and foreign exchange markets on banks’ operation, change from a liquidity deficit to a liquidity surplus. We show that object-based models can act as a practical and effective tool for studying the issue examined here and other similar subjects.

Appendices are available at: www.cbr.ru/money-and-finance

5. References


De Nicolo, G., Favara, G. and Ratnovski, L. (2012). Externalities and Macroprudential Policy. IMF Staff Discussion Notes, № 12/05.


Bank’s Hidden Negative Capital Before and After the Senior Management Change at the Bank of Russia

Mikhail Mamonov, Center for Economic Analysis and Short-Term Forecasting

What was the size of hidden negative capital (HNC) in the Russian banking system before the Bank of Russia’s senior management changed? Has the financial regulator’s new, tougher, supervision policy launched in mid-2013 been productive in dealing with problems accumulated over recent years? This study attempts to provide first answers to these questions using Heckman selection models and solving a possible regulator’s optimization problem using the data on Russian banks over the period from December 2009 to May 2017. The study results suggest that before the change of the Bank of Russia’s senior management in mid-2013, the average level of HNC at banks operating in Russia was very high: 14% of the banking system’s total assets on each specific date, and fairly permanent in time. But as early as half a year after the Bank of Russia’s senior management changed, the overall size of HNC started contracting rapidly, coming close to 4% by mid-2016. In the last 12 months of the time span under study the level of the HNC stabilized at 4%. These estimates point to the high effectiveness of the new senior management’s supervision policy. This policy has produced a strong indirect positive effect: large part of fragile banks have started addressing their problems without waiting for the Central Bank’s intervention. The effectiveness of the Bank of Russia’s new policy cannot be attributed to other factors such as macroeconomic stabilization or improvement and/or rising profitability of bank operations. Rather, when supervision policy was stepped up, these other factors were acting in the opposite direction.

Keywords: banks, hidden negative capital, fraudulent accounts, Bank of Russia, supervision policy

JEL: G21, P23, P34, P52

---

1 The study was carried out under a research grant № 16-36-60037 mol_a_dk from the Russian Fund for Fundamental Research (RFBR) in 2017.

The author is grateful to Oleg Solntsev and Anna Pestova for establishing a creative environment and discussing the study concepts. In addition, the paper was substantially improved thanks to comments by Ion Lapteacru and Simona Nistor in the course of a special session of the 19th Annual Conference of the International Network for Economic Research (INFER) at the Bordeaux University, France. The author acknowledges discussion of the paper with the participants of research seminars at the Laboratory for Macroeconomic Research of National Research University Higher School of Economics and the Finance Department of Higher School of Economics, II World Congress on Comparative Economics “1917–2017: Revolution and Evolution” (St Petersburg), and the 34th Symposium on Money, Banking and Finance at Nanterre University, France.
1. Introduction

In recent years, the issue of hidden negative capital (HNC) at Russian banks has come to the fore in discussions on improving banking system stability, overshadowing such topics as nationalization of financial institutions, the falling number of regional banks, introduction of the countercyclical capital buffer or rising interest rates on banking loans to the real economy. Under the standard definition used in the literature and among financial regulators, HNC means a negative difference between a bank's total assets and its total liabilities (James, 1991; Kang et al., 2015; Cole and White, 2015), which is thus close to the concept of negative net worth. Why can such a negative difference arise and how long can it persist? The answer to the first question is that any economy, industry or an individual firm may, at any time, encounter negative productivity, income allocation or insolvency shocks (Bernanke and Gertler, 1989) forming a negative financial cycle spiral persistent in time (Kiyotaki and Moore, 1997; Chen, 2001). The answer to the second question is that the period of HNC's persistence may be rather long for at least two reasons. On the one hand, from the perspective of the bank capital theory (Diamond and Rajan, 2000), a bank can operate with HNC until it loses the confidence of depositors, who, because of information asymmetry, may be unaware of its problems, continuing to replenish its liabilities, but as this asymmetry is overcome, an immediate deposit run and a license withdrawal by the regulator follow. (Diamond and Dybvig, 1983). On the other hand, any economy is subject to regulatory forbearance (examples from the US are in Wheelock and Wilson, 2000; Kang et al., 2015), with a regulator being unable to withdraw licenses for a multitude of reasons ranging from the high cost of overcoming information asymmetry to possible political pressure and a government budget deficit (Brown and Dinç, 2011).

The problem of banks' HNC in Russia could not fail to arise because of certain features of the system itself and those of economic agents.

First, economic agents – banks and their borrowers in this case – have a higher risk tolerance (compared with developed countries) leading to both higher returns on equity and higher volatility of these returns (implying more frequent and massive bank failures).

Second, the banking system itself is more dynamic due to its young age and undersaturation of key markets for banking services because of it.

Third, the Russian economy has an increased vulnerability vis-à-vis external shocks.

Fourth, the regulatory environment is not static, proposing to the system various innovations, which create adaptation costs for banks. All these peculiarities set the stage for the spreading and escalation of shocks, which, as pointed out in the above theoretical studies, give rise to HNC and make it impossible for the regulator to detect them immediately.

2 We note that certain relationships can be found between any of these topics and the HNC, which are to be addressed in further studies.
According to estimates using the Bank of Russia Bulletin data, from mid-2013 to mid-2017, the regulator discovered 2 trln rubles of HNC at a total of 227 banks (around 2.5% of the banking system’s assets). In our previous study we estimated the overall size of yet-undetected HNC at 3.6%–6.8% of the banking system’s assets as of mid-2016, the statistical detection threshold standing at 5% (Mamonov, 2017b). This study provides a more in-depth analysis and poses the following questions:

1. What was the total amount of HNC accumulated by the time the Bank of Russia’s senior management changed?
2. How stable was their overall level prior to mid-2013?
3. Did this level start declining after the Bank of Russia’s senior management had changed, and if so how fast and what it stemmed from (a change in macro-, micro-, regulatory environment or altogether)?

Answers to these questions can shed light on whether the Bank of Russia’s policy of cleaning the banking system from unfair players was effective or not. To provide answers to these questions, we will, following Mamonov (2017b), use the methods of Heckman selection models, allowing to evaluate the size of what is hidden and thereby address the problem of the sample selection bias (Heckman, 1979). Specifically, by using a fixed sample of banks with HNC detected by mid-2016 (the fragile group), we will estimate the size of HNC at banks that were operating in each month from the start of 2010 to mid-2016 (backward forecasting) and from mid-2016 to mid-2017 (forward forecasting), i.e., in the variable reference group. Thus, what distinguishes this study is that rather than fix the sample composition at the end of the observation period (mid-2017) it forms the sample within the available time span, closer to its end (mid-2016), for the current experience in detecting HNC to be sufficient for their backward and forward (for the immediate future) forecasting. It is well known that balance sheet fraud gets ever more sophisticated, as banks are well aware that, learning from past experience, the Bank of Russia keeps improving its supervision policy.

The results of this study suggest that the efforts of the Bank of Russia’s new senior management to uncover HNC at banks have borne fruit: apart from the direct positive effect (uncovering of HNC and license withdrawals) this has also produced an indirect positive impact: banks operating in Russia have started scaling down balance sheet fraud without waiting for the Bank of Russia’s intervention.

This paper comprises the following sections: Section 2 explains the methodology of the study. Section 3 describes the data used. Section 4 presents the results of regression analysis and estimation of HNC at banks still operating during the period. Section 5 concludes.

---

3 At the time when this paper was prepared official statistics on the major developments of the second half of 2017 — uncovering HNC at some of the top-10 banks (the Otkritie Financial Corporation and Binbank), were still unavailable.
2. Methodology of detecting and backward and forward forecasting of banks’ hidden negative capital

Detection of HNC, to the best of our knowledge, has not been addressed in academic literature on financial stability. Previous research considered determinants of negative capital in already failed banks (James, 1991; Kang et al., 2015; Balla et al., 2015; for detailed reviews see Mamonov, 2017a, b). Therefore, hereinafter we will concentrate on the methodology of HNC detection at different points of time.

In our previous study, we described the functional of Heckman selection models which allows — at each given probability threshold — evaluating the size of HNC at any of still operating banks through comparison with the sample of banks where the regulator has previously uncovered such HNC (Mamonov, 2017b). We will refer to the sample of still operating banks as the reference group, while banks where HNC have already been detected will be termed the fragile group. In this study we will fix the fragile group composition as of mid-2016 and vary the reference group’s composition in the December 2009 – June 2016 time span for the purpose of backward forecasting of HNC and in the July 2016 – May 2017 horizon to perform forward forecasting of HNC. In this manner, we will run a series of $7 \times 12 + 6 = 90$ Heckman selection models – one model for each month in the period from December 2009 to May 2017. Heckeman selection model will be given by:

$$D_{it} = \alpha_1 + \sum_{l=1}^{L} \beta_{1.l} \cdot BSF_{l,it-k} + \delta \cdot \ln{T}_A_{it} + \rho_1 \cdot MACRO_{j,t} + \varepsilon_{1,it+\tau} > 0$$ (1)

$$\frac{NNW_{it+\tau}}{Capital_{it-1}} = \alpha_2 + \sum_{l=1}^{L} \beta_{2.l} \cdot BSF_{l,it-k} + \rho_2 \cdot MACRO_{j,t} + \varepsilon_{2,it+\tau}$$ (2)

where $i$ is the index of a bank which may belong to the fragile or reference group; $t$ the varying time of model estimation ($t = 1, 2, \ldots, 90$). $NNW_{it+\tau}$ (Negative Net Worth) is the size of the HNC (bln rubles) uncovered by the Bank of Russia $\tilde{t} + \tau$ time after the license withdrawal ($\tau = 1, 2$ months) from the banks of the fixed fragile group, or the value to be estimated for still operating banks from the variable reference group. Therefore, for banks from the fragile group $t = \tilde{t}$, where $\tilde{t} = 43, 44, \ldots, 79$, which corresponds to the time span from June 2013 (the Bank of Russia’s senior management change) to June 2016 in which we fixed the composition of the fragile sample. $Capital_{it-1}$ is the stock of a bank’s equity capital on its balance sheet. $D_{it}$ is the likelihood of detecting nonzero HNC (a latent variable). $BSF_{l,it-k}$ are microeconomic control variables reflecting differences across banks by parameters such as the structure of their assets and liabilities, the size of business and risk tolerance, the scale of balance sheet fraud based on indirect signs (an increased reliance on taking household deposits combined with an increased reliance of the business model on corporate lending, an increased turnover of funds in
correspondent accounts with the Bank of Russia along with a decreased repo liabilities turnover (for details see Mamonov, 2017b). $k = 3$ months, similarly to the previous study. \(\text{MACRO}_{j,t}\) is a \(j\)-th control variable for the macroeconomic change (annual GDP growth rate, unemployment, etc.). $\varepsilon_{1,lt+\tau}$ and $\varepsilon_{2,lt+\tau}$ are regression errors of the selection equation (1) and size equation (2), respectively.

It should be reminded that Heckman model is deemed to be identified, while an incomplete sample bias is deemed proved if the following two conditions are met: a variable is found which has a substantive and statistically significant effect on the selection process but has no impact on size; a statistically significant correlation of errors $\varepsilon_{1,lt+\tau}$ and $\varepsilon_{2,lt+\tau}$ is shown. If these conditions are met then consistent estimates of coefficients of equations (1)–(2) are only possible using the Heckman one-step maximum likelihood (ML) procedure or the effective Heckman two-step method. In the former case, equations (1)–(2) are estimated within a system, and it takes certain time to reach convergence, which is not guaranteed in the general case. In the latter case, equation (1) is estimated as a usual probit-model, used to calculate the nonselection hazard which is equal to the ratio of the density function value to that of the distribution function in each observation point. Then this variable is added to equation (2) as a supplementary regressor and its statistical significance is tested. If one of the two conditions is not satisfied, a consistent estimation is possible using any available technique. We will use the two-step Heckman method in the Stata package which does not require time for convergence to be reached.

Thus equation (1) prescribes the selection rule, i.e., the procedure for attributing each bank in either the fragile or reference group, while equation (2) determines the relationship between the relative size of the HNC and the fundamentals. Equation (2) enables the HNC size to be estimated in each of the 90 months for each bank from the reference group. The mechanism is so structured that this will be a nonzero value, but it is probabilistic and should be compared with the probability of detecting HNC which is estimated by equation (1). This means that we need a probability threshold, and if it is exceeded we will believe that the size of the HNC estimated by equation (1) is indeed nonzero. This in turn requires threshold quantification.

We will quantify the probability threshold using the out-of-sample forecasting procedure and, unlike our previous study, the estimation of the regulator’s utility function for each possible value of this threshold, i.e., ranging from 1% to 100%. Formally, we will move from the upper bound of $\tau$ period backwards (end-2015), carry out estimation using model (1)–(2), calculate the size of HNC at the banks from the reference group and compare them with banks where HNC were actually found and license withdrawn in the subsequent half a year. In addition, we will estimate type I and type II standard errors – cases when the model predicted nonzero HNC but they were not found and those when the model did not predict nonzero HNC but the regulator detected them. The regulator’s utility function will
be defined as one where the regulator wishes to maximize the share of accurately predicted events of HNC presence and absence, i.e., the share of correct attribution of banks to each of the two groups under analysis. We do not know the regulator’s preferences regarding the importance of banks’ correct attribution to the first or second group, which makes us prescribe them in a judgement-based manner. We will assume that the regulator may be (a) indifferent regarding the attribution of banks to the first or second group (weights \( \gamma_1 = \gamma_2 = 0.5 \)), (b) attaches greater importance to the correct attribution of banks to the fragile group (\( \gamma_1 = 0.75 \) and \( \gamma_2 = 0.25 \)), (c) views the correct attribution to the reference group as having a greater value (\( \gamma_1 = 0.25 \) and \( \gamma_2 = 0.75 \)). Therefore, the problem of maximizing the regulator’s utility function is given by:

\[
U(p, \gamma_1, \gamma_2) = \gamma_1 \frac{A(p)}{A(p) + B(p)} + \gamma_2 \frac{D(p)}{C(p) + D(p)} \rightarrow \max
\]

under constraints \( \gamma_1 + \gamma_2 = 1 \)

\( 0 \leq \gamma_1 \leq 1 \) and \( 0 \leq \gamma_2 \leq 1 \)

where \( U(p, \gamma_1, \gamma_2) \) is the value of the regulator’s utility function depending on the chosen \( p \) threshold of uncovering HNC and the regulator’s preferences regarding the attribution of a bank to one or the other group, \( \gamma_1 \) and \( \gamma_2 \). The events \( A(p), B(p), C(p) \) and \( D(p) \) are standard for the literature on bank failures (DeYoung and Torna, 2013) and the signaling approach (Kaminsky and Reinhart, 1999). Their meaning is disclosed in Table 1.

### Table 1. Predictive characteristics of selection models for bank HNC required for constructing the regulator’s utility function

<table>
<thead>
<tr>
<th>HNC is predicted at ( t_0 ) time using detection probability threshold equal to ( p ) (( p = 1%, \ldots 100% ))</th>
<th>( \gamma_1 )</th>
<th>( \gamma_2 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>yes</td>
<td>A(p)</td>
<td>B(p)</td>
</tr>
<tr>
<td>no</td>
<td>C(p)</td>
<td>D(p)</td>
</tr>
</tbody>
</table>

In solving the optimization problem (3)–(5) we obtain the value of \( p^* \) probability threshold and those of the regulator’s preferences \( \gamma_1^* \) and \( \gamma_2^* \) providing for the maximum regulator’s utility function.

Then, using the obtained \( p^* \) optimal threshold, backward and forward forecasting of the values of HNC for banks from the reference group will be carried out. These estimates will then make it possible to analyze the effectiveness of the Bank of Russia’s banking sector recovery policy.
Concluding this section, we point out that, to our knowledge, Heckman type (1) and type (2) selection models were only used to analyze bank HNC in Balla et al. (2015) using US bank data for the periods of the 1980s and 2000s banking crises. That study however focused on the consistent estimation of negative capital determinants rather than that of their hidden size at American banks that were still in operation. With regard to Russia’s banking system, Drobyshevsky and Zubarev (2011), Karminsky and Kostrov (2013), Peresetsky (2013) and many others studied the probability of bank failures using logit- or probit-models but did not seek to find out the size of HNC at banks whose licenses had been withdrawn.

3. Dataset

Official data on the size of HNC revealed by the Bank of Russia at troubled banks are published in the Bank of Russia’s Bulletin some time after a bank’s failure in case the actual value of assets is smaller than the value of liabilities in each respective bank. This study uses the same list of banks with HNC found after the Bank of Russia’s senior management changed in mid-2013 until July 2016 as Mamonov (2017b) did. These are 167 banks from the fragile group fixed in time for the purposes of our study (see Section 2). The reference group varying in time for the same purposes was made up of around 1020 banks at the beginning of the time span under analysis, December 2009, dwindling to 576 banks at the end of the period, May 2017. Data for forming the composition of control factors in equations (1) and (2) were obtained from standard sources, which are Forms 101 containing monthly balances of credit institution assets and liabilities that are disclosed on the Bank of Russia website. Then the outliers – data from the 1st and 99th percentiles for all the relative variables in equations (1) and (2), were eliminated. As a result, there were 152 banks left in the fragile group, while the resulting number for the reference group was 661 banks at the start of the time span under analysis, 735 in mid-2013 and 508 at the end of the period (the significant contraction in the number of observations at the start of the period stemmed from the shortage of observations for some individual indicators).

Descriptive statistics are presented in Table 2 as of the start of the period in question, mid-2013, when the Bank of Russia’s senior management changed, and as of the end of this time span. Since Heckman selection model is estimated simultaneously using data from both groups, the descriptive statistics are presented as a total for the fragile and reference groups. Moreover this study does not seek to identify differences between these groups (this was done in the previous study, Mamonov (2017b)). We are however interested in the variables themselves because they determine the predictive power of the selection model (1)–(2) and, accordingly, forecasts for the size of HNC. Therefore, we will look at the variables in somewhat more detail.

4 See: https://www.cbr.ru/credit/forms.asp
### Table 2. Descriptive statistics: before, during and after the Bank of Russia's senior management change (as a total for the reference and fragile groups)

<table>
<thead>
<tr>
<th>Bank-level explanatory variables, lag = 3 months before license withdrawal</th>
<th>1 January 2010 (I)</th>
<th>1 July 2013 (II)</th>
<th>1 June 2017 (III)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Average</strong></td>
<td><strong>Standard deviation</strong></td>
<td><strong>Average</strong></td>
<td><strong>Standard deviation</strong></td>
</tr>
<tr>
<td>Household deposits &gt; median × Corporate loans &gt; median</td>
<td>100.6</td>
<td>305.2</td>
<td>141.0</td>
</tr>
<tr>
<td>Corporate loans &gt; median × Profit/capital &lt; median</td>
<td>-4.4</td>
<td>17.0</td>
<td>-12.1</td>
</tr>
<tr>
<td>Loan loss provisions &gt; median × NPL &lt; median</td>
<td>5.0</td>
<td>10.7</td>
<td>5.3</td>
</tr>
<tr>
<td>Monthly turnovers in corresp acnts with Bank of Russia, %</td>
<td>61.9</td>
<td>286.2</td>
<td>299.5</td>
</tr>
<tr>
<td>Monthly turnovers of repo liabilities, %</td>
<td>1.1</td>
<td>4.1</td>
<td>1.5</td>
</tr>
<tr>
<td>Household loans / Assets, %</td>
<td>15.0</td>
<td>14.4</td>
<td>15.6</td>
</tr>
<tr>
<td>Absolutely liquid assets / Assets, %</td>
<td>14.6</td>
<td>14.5</td>
<td>11.7</td>
</tr>
<tr>
<td>Corporate securities / Assets, %</td>
<td>0.8</td>
<td>2.4</td>
<td>1.1</td>
</tr>
<tr>
<td>Loans due from interbank market / Assets, %</td>
<td>6.5</td>
<td>10.0</td>
<td>8.3</td>
</tr>
<tr>
<td>Foreign liabilities / Liabilities, %</td>
<td>4.2</td>
<td>10.0</td>
<td>4.2</td>
</tr>
<tr>
<td>Capital / Assets, %</td>
<td>20.4</td>
<td>12.5</td>
<td>18.3</td>
</tr>
<tr>
<td>Loan loss provisions / Loans, %</td>
<td>12.9</td>
<td>11.5</td>
<td>13.4</td>
</tr>
<tr>
<td>Annual asset growth rate, %</td>
<td>11.0</td>
<td>56.0</td>
<td>20.3</td>
</tr>
<tr>
<td>Assets (log.)</td>
<td>1.2</td>
<td>1.8</td>
<td>1.7</td>
</tr>
<tr>
<td>Number of observations: min–max</td>
<td>817–834</td>
<td>888–915</td>
<td>660–689</td>
</tr>
</tbody>
</table>

*a* The values are normalized by the value of assets, %

*Note: In the parentheses under the coefficients, their robust standard errors are given.>*c* (<c>) — the censored variable: it takes values if they are higher (lower) than *c* and 0 otherwise.
As shown in our previous study (Mamonov, 2017b), the censored variable has the greatest explanatory power as it reflects all events of banks’ increased reliance on both taking household deposits and, at the same time, an increased share of corporate loans in their total assets (compared with median values for the banking system). As can be seen in the table of descriptive statistics below, this concurrence only became more pronounced in time. This could not help affecting the profitability of banking operations, as shown by the negative dynamics of the second variable in the table. This variable reflects all the events of banks showing both an increased share of corporate loans in their total assets and a decreased ROE (compared with median values for the banking system). With the profitability of operations falling, incentives to falsify balance sheets could not help growing. These strengthening incentives are evidenced by the change in the third variable in the table. This is a censored variable reflecting all events of both an increased loan loss provisioning (as a percentage of assets) and a decreased share of non-performing loans (as a percentage of loans). These situations raise suspicion, because, if the quality of loans is better than average for the system, why set aside larger loan loss provisions than average for the system. If the situation improves in the banking sector, this variable should move towards zero. Looking at the fourth and fifth variables, we note a manifold expansion in the turnover of funds in banks’ correspondent accounts with the Bank of Russia and repo liabilities. Among more straightforward indicators (from the 6th to the 14th) we note a twofold rise in loan loss provisioning, representing loan quality deterioration, as well as the sluggish movement of assets at both ends of the period under study (an effect of the period being “framed” with the crises).

4. Evaluation of aggregate HNC using Heckman selection models

4.1. Regression analysis results

A series of 90 sequential estimations of the Heckman selection models (1)–(2) was conducted for each respective month within the time span under analysis, from December 2009 to May 2017. Table 3 presents a part of these estimates: at the end points of the time span and in mid-2013, the time of the Bank of Russia’s senior management change. The first column of the table contains all of the 14 microeconomic control variables, for which descriptive statistics were presented in Table 2 above. In the second and third columns there are estimates of the equations of selection (1) and size (2) for the lower boundary of the time span under analysis (model I), in the fourth and fifth — those for the time of the Bank of Russia’s senior management change (model II), in the sixth and seventh — those for the upper boundary of the time span (model III). The lower part of the table contains the

---

5 The rest of the estimates for the other 87 months are not shown for space considerations but are available upon request.
relevant statistics: the number of banks in the reference and fragile groups, the median probabilities in the respective bank groups, and the correlation of regression errors in the equations of selection and size. As is clear from the table, we did not use macroeconomic variables in our final estimations, because we found that their inclusion in the first two years of the time span under analysis causes the estimates of HNC size to move upwards substantially with the coefficient estimates being formally consistent and statistically significant. All the variables used were three-month lagged to provide a signal window.

Table 3. Selection models of bank HNC: estimates at the edges of the available time span and at the time of the Bank of Russia’s senior management change

<table>
<thead>
<tr>
<th>Bank-level explanatory variables, lag = 3 months before license withdrawal</th>
<th>1 January 2010 (I)</th>
<th>1 July 2013 (II)</th>
<th>1 June 2017 (III)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Selection</td>
<td>Size</td>
<td>Selection</td>
<td>Size</td>
</tr>
<tr>
<td>1. Household deposits &gt; median × Corporate loans &gt; median</td>
<td>0.0019*** (0.0003)</td>
<td>0.0013** (0.0006)</td>
<td>0.0011*** (0.0041)</td>
</tr>
<tr>
<td>2. Corporate loans &gt; median × Profit/capital &lt; median</td>
<td>-0.029*** (0.005)</td>
<td>-0.015* (0.008)</td>
<td>0.004* (0.002)</td>
</tr>
<tr>
<td>3. Loan loss provisions &gt; median × NPL &lt; median</td>
<td>0.007 (0.007)</td>
<td>-0.005 (0.018)</td>
<td>0.020*** (0.005)</td>
</tr>
<tr>
<td>4. Monthly turnovers in corresp accounts with Bank of Russia, %</td>
<td>0.000 (0.000)</td>
<td>0.002*** (0.001)</td>
<td>0.000 (0.000)</td>
</tr>
<tr>
<td>5. Monthly turnovers in repo liabilities, %</td>
<td>0.020 (0.015)</td>
<td>-0.104** (0.048)</td>
<td>0.015 (0.012)</td>
</tr>
<tr>
<td>6. Household loans / Assets, %</td>
<td>0.006 (0.005)</td>
<td>-0.026 (0.017)</td>
<td>-0.008* (0.004)</td>
</tr>
<tr>
<td>7. Absolutely liquid assets / Assets, %</td>
<td>-0.005 (0.007)</td>
<td>-0.011 (0.024)</td>
<td>-0.003 (0.006)</td>
</tr>
<tr>
<td>8. Corporate securities / Assets, %</td>
<td>0.006 (0.024)</td>
<td>0.064 (0.066)</td>
<td>-0.015 (0.018)</td>
</tr>
<tr>
<td>9. Due from IB market / Assets, %</td>
<td>-0.003 (0.008)</td>
<td>-0.035 (0.028)</td>
<td>-0.026*** (0.008)</td>
</tr>
<tr>
<td>10. Foreign liabilities / Liabilities, %</td>
<td>-0.061*** (0.016)</td>
<td>0.080 (0.049)</td>
<td>-0.014 (0.009)</td>
</tr>
<tr>
<td>11. Capital / Assets, %</td>
<td>-0.028*** (0.008)</td>
<td>-0.130*** (0.027)</td>
<td>-0.016** (0.007)</td>
</tr>
<tr>
<td>12. Loan loss provisions / Loans, %</td>
<td>0.030*** (0.007)</td>
<td>0.011 (0.021)</td>
<td>0.006*** (0.002)</td>
</tr>
</tbody>
</table>

* values are normalized by the value of assets, %
The key estimation results are as follows. First, Heckman selection models successfully address the problem of identifying bank HNC throughout the 90 months of the time span under analysis, but two conditions of identification were not met for about 20 months of the 2012–2013 period (the assets variable in the selection equation and the correlation of errors in the equations of selection and size were not significant, see Section 2). This means that consistent coefficient estimates can be achieved using not only the effective two-step Heckman procedure but also by other available methods. We present Heckman estimates for all the segments of the time span under analysis for comparability reasons. In comparing the estimates of detection likelihood and the size of the predicted HNC, we can see that at the time span borders the estimates of HNC detection likelihood were substantially higher than those in the middle of the time span: around 50% in December 2009 and May 2017 versus 32% in June 2013, while the estimates of the
median HNC size were roughly stable in time. This in itself suggests that (a) HNC as a phenomenon were inherent in the Russian banking system at least 3.5 years before the senior management change at the Bank of Russia; (b) there was a very high probability of identifying these HNC and (c) the signs of the median HNC size declining are yet to be seen (meaning that the banking sector cleanup from unscrupulous players has yet to be completed).

Second, the estimates themselves allow ranking all the variables by their statistical significance for explaining HNC. Variables 1, 2, 11, and 13 are, for instance, significant throughout the time span, variables 3 and 9 — in its second half, variables 4, 5, 10, and 14 — at both edges of the time span. We are therefore refining the findings of our previous study (Mamonov, 2017b) and concluding that variables I and II are the most stable in time indicators of HNC. They signal that the likelihood and size of HNC were the higher throughout the time span, the higher, relative to the system's averages, were the share of retail deposits in liabilities and that of corporate loans in assets (variable I) and the larger was the shortfall of a bank's own resources (variable II).

Third, we have also obtained noteworthy results regarding other variables. Specifically, if a bank had an increased share of both loan loss provisions and, at the same time, a reduced share of nonperforming loans (relative to the system in both events), then, beginning from the second half of the time span under analysis, this consistently evidenced the existence of a “hole” in its capital but was not very informative for evaluating its size (variable 3). Meanwhile, the level of loan loss provisioning (variable 12) itself was positively correlated with the selection likelihood in the first half of the time span and negatively in its the second half. This may indicate banks’ changing attitude towards loan loss provisioning thanks to the new policy of the Bank of Russia’s new senior management to eliminate the banking sector’s problems accumulated over the previous years. Before the management change, banks’ decisions concerning provisioning may have been quite arbitrary and, therefore, positively correlated with the likelihood of HNC. By contrast, in the second half of the time span under study, when banks began to understand that the Bank of Russia may visit them any time for inspection, provisions began to perform their function. As a result, if provisions have already been set aside, a bank is better protected from micro- or macroeconomic shocks and features a proportionally lower likelihood and a smaller size of HNC. Another finding regards liquidity. If a bank has an adequate supply of liquidity (variable 7) and is a net creditor in the interbank market (variable 9), the likelihood of HNC becomes lower, and the size of this HNC if it has been found, – smaller. Finally, there is a countercyclical relationship between the annual change in bank assets and the likelihood and size of HNC. For instance, during macroeconomic crises (or in the vicinity thereof) an increased asset growth rate is positively correlated with both of the above indicators for HNC, while times of economic growth see their negative correlation, as very much expected.
4.2. Calibration of probability threshold under the maximization of the regulator’s utility function

The preceding stage of this study has provided estimates of two panel data series: those of detection probability and size of HNC at banks still in operation in each of the 90 months of the time span under analysis. Now we need to determine what HNC sizes can be deemed indistinguishable from zero and which cannot. Under the methodology employed (see Section 2), we will carry this out using the detection probability threshold computed based on the regulator’s utility function (3).

The results of solving the regulator’s optimization problem are presented in Figure 1 for three prescribed sets determining the regulator’s preferences (unknown to us) regarding the importance of attributing a bank to the reference or fragile group.

Figure 1. Results of the regulator’s utility function estimation

Weighted average sum of shares of accurately predicted events of HNC presence (“1”) and absence (“0”); the weights are equal to the preset values of the regulator’s unknown preferences; a 6-month forecast period.

The first set, where the regulator views correctly attributing banks to both the fragile and reference group as equally important (weights \(\gamma_1 = \gamma_2 = 0.5\)), yields the utility function whose maximum is achieved at a detection probability level of 8% \((p = 0.08)\).

The second set, where the regulator views correctly attributing a bank to the fragile group as more important (weights \(\gamma_1 = 0.75, \gamma_2 = 0.25\)), determines the regulator’s utility function whose maximum utility corresponds to a softer threshold than in the preceding case, that of 4%, \((p = 0.04)\).

The third set, in which the regulator views correctly attributing a bank to the reference group as more important (weights \(\gamma_1 = 0.25, \gamma_2 = 0.75\)), represents a utility function whose maximum expectedly goes to the upper probability bound at 100% \((p = 1)\). This option is noninformative for our further analysis. But we note that at a threshold of 13% \((p = 0.13)\) there is an intersection of all of the three options, meaning that all the choices result in the same level of the regulator’s utility.
Thus, we have three candidates for the role of the optimal threshold $p^*$. We will now compare the predictive capability of selection model (1)–(2) for each of them and choose the best of the available options. The comparison results are shown in Table 4, in the first panel of which the predictions of the presence/absence of HNC (values $A(p)$, $B(p)$, $C(p)$ and $D(p)$) are shown and the shares of accurately predicted events of HNC presence/absence are computed, while the second panel shows estimates of HNC size for each case given in the first panel and those of the shares of accurately predicted HNC sizes in their actual total size identified by the regulator in the first half of 2016.

Table 4. Bank HNC: out-of-sample forecast for a 6-month time span

<table>
<thead>
<tr>
<th>HNC in model</th>
<th>Yes</th>
<th>No</th>
<th>Yes</th>
<th>No</th>
<th>Yes</th>
<th>No</th>
</tr>
</thead>
<tbody>
<tr>
<td>Threshold = 4%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Threshold = 8%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Threshold = 13%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Panel 1: Prediction of HNC detection

<table>
<thead>
<tr>
<th>Actual HNC in 1H 2016</th>
<th>Yes</th>
<th>No</th>
<th>Yes</th>
<th>No</th>
<th>Yes</th>
<th>No</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes</td>
<td>42</td>
<td>1</td>
<td>38</td>
<td>5</td>
<td>27</td>
<td>16</td>
</tr>
<tr>
<td>No</td>
<td>522</td>
<td>172</td>
<td>395</td>
<td>299</td>
<td>264</td>
<td>430</td>
</tr>
<tr>
<td>Share of accurately predicted events of HNC presence, %</td>
<td>97.7</td>
<td>88.4</td>
<td>62.8</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Share of accurately predicted events of HNC absence, %</td>
<td>37.7</td>
<td>56.0</td>
<td>73.4</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Panel 2: Prediction of HNC size, bln rubles

<table>
<thead>
<tr>
<th>Actual HNC in 1H 2016</th>
<th>Yes</th>
<th>No</th>
<th>Yes</th>
<th>No</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes</td>
<td>253</td>
<td>402 / 1</td>
<td>161</td>
<td>190 / 213</td>
</tr>
<tr>
<td>No</td>
<td>6443</td>
<td>2309</td>
<td>1131</td>
<td></td>
</tr>
<tr>
<td>Size of accurately predicted HNC, %</td>
<td>62.8 / 62.7</td>
<td>84.6 / 39.9</td>
<td>74.9 / 32.9</td>
<td></td>
</tr>
</tbody>
</table>

For reference:
Actual size of HNC detected by regulator over 1H 2016 = 403.4 bln rubles

Note: Panel 2 presents the results of an out-of-sample forecast of the HNC size using selection model (1)–(2). For instance, at a threshold of 4% the model predicts 253 bln ruble-worth total HNC detected by the regulator at 43 banks in the 1st half of 2016. The actual size of these HNC totals 402 bln rubles. The model failed to predict a 1 bln ruble HNC at one bank. So the share of accurately predicted HNC over this period stood at 62.8% = 253 / (253 + 402) × 100 for banks which the model correctly identified and 62.7% = 253 / (253 + 402 + 1) × 100 for all events which the model identified and failed to identify. The rate of success for the other two threshold values is calculated in a similar manner.
As the estimation results show, on the one hand, a threshold value increase goes along with a pronounced contraction in the share of accurately predicted HNC and their size, and, on the other hand, this is accompanied by a rise in the share of accurately predicted events of their absence. If we were to decide on an optimal threshold of 4% this would produce a prediction of almost totally accurate prediction of HNC in still operating banks but also a fairly low share of accurately predicted events of their absence – slightly above 1/3. If, conversely, we were to use the most rigid of the three thresholds, 13%, we would see the opposite situation. Hence an intermediate threshold of 8% looks like a reasonable compromise. It yields a still high share of accurately predicted HNC but already a hefty 56% share of accurately predicted events of their absence. We will therefore settle on \( p^* \) as an optimal threshold. At this threshold, the share of predicted size of HNC in their overall volume for the 38 banks at which the model correctly detected the presence of HNC (panel 1) stood at 85%. In other words, the model fairly accurately predicts the size of HNC if it detects its presence at a bank. The model, however, failed to recognize HNC at the other 5 banks (panel 1), whereas the regulator actually found there were such HNC for a total of 213 bln rubles, i.e., these are quite large banks. So, the share of accurately predicted volume of HNC in the overall size of both those the model detected and those it failed to detect stands at just 40%. Undoubtedly, this is a limitation in terms of our analysis, which should be addressed in further studies. This suggests that the problems of large banks are hidden from the regulator much more carefully and are much harder to find out statistically – from a sample comprising primarily small banks (large banks have access to a wider array of accounting manipulations compared with small banks, including transborder manipulations with the balances of correspondent accounts at foreign banks, etc.). For now, we will not try to refine the results we have obtained but will assume that subsequent estimates of the overall HNC size may be too conservative. However, this exercise is informative in any case, as it shows the least amount of losses society suffers because of unfair competition in the banking system.

4.3. Change in banks’ HNC and the probabilities of their detection before and after the Bank of Russia’s senior management changed: is the situation improving?

Having estimated Heckman selection models and quantified probability thresholds, we can now move on to constructing dynamic series of HNC and probabilities of detecting thereof over the December 2009 – May 2017 period. In each of these 90 months, the resulting total of HNC included all observations for which model (2) predicted a nonzero HNC and model (1) predicted an at least 8-percent likelihood of detecting it. The results of this estimation are presented graphically in Figure 2 below as a percentage of the banking system’s total assets on
The total size of HNC is estimated by comparing profiles of banks from the fragile (with HNC) and control (without HNC) groups (Section 2). The fragile group was made up of all banks for which the size of HNC could be found in The Bank of Russia Bulletin by mid-2016 and which were subjected to standard outlier filtration. In addition to the dynamics of HNC, two points in time are indicated: mid-2013, when the Bank of Russia’s management changed, and mid-2016, the time when the fragile group was made up. This is the key point of this study. In analyzing the dynamics, we obtained relative to the above two points in time, we can draw the following conclusions:

First, prior to the change in the Bank of Russia’s senior management, the average level of HNC was very high — 14% of the banking system’s total assets, and quite stable in time but showed pronounced volatility, varying between 10% and 16%.

Second, as early as a year after the Bank of Russia’s senior management changed, the level of HNC started to decline, falling to 4% by mid-2016, i.e., by four times over three years. Meanwhile mid-2015 saw a surge from 5% to 10% — this was exactly when the model started to predict HNC at the Otkritie Financial Corporation and a number of other Russian financial institutions, some of which successfully dealt with their problems — or hid them in an even more sophisticated manner.

Third, in the last 12 months of the time span under analysis the level of HNC stabilized at 4%.

How are these empirical findings to be interpreted? In our view, they support the conclusions of numerous studies to the effect that problems have accumulated for years in Russia’s banking system given the lack of drastic policy measures to eliminate unscrupulous players from the banking sector. It is noteworthy that before the senior management change at the Bank of Russia the amount of HNC was quite substantial and stable in time, let alone close to zero. The phenomenon of HNC is not something newly emerged amidst dim economic prospects or Western sanctions against Russia. The above numbers point to the strong effectiveness of the new senior management’s measures, which helped reduce the
relative level of HNC fourfold. Moreover, the impact of their reduction can be decomposed into direct and indirect effects. As is known – and pointed out in the Introduction — over the time the new management were taking action, detected HNC totaled an equivalent of 2.5% of the banking system’s assets, while the overall effect amounts to 16% – 4% = 12%. This suggests a substantial indirect effect: in anticipation of supervision policy toughening, banks which still operated but already encountered problems had to start addressing them on their own by cleaning their balance sheets from toxic assets and/or scaling down their loans to related parties without waiting for the Bank of Russia to show up. In other words, banks which were able to deal with their problems on their own have already done so, while the ones incapable of resolving the situation will face the regulator. As the change in the relative level of HNC suggests, the indirect effect has already petered out, meaning that banks which keep hiding HNC will likely be found fairly fast.

We will now deal with the distribution dynamics of probabilities of detecting HNC (Figure 3). It can be seen that in the period from mid-2013 to early 2017 the median probability exceeded the selected 8% optimal threshold, i.e., more than half of banks in operation may have concealed their HNC. It is only towards the middle of the period under analysis that the median probability fell below the threshold along with the probability value at the lower percentiles of distribution. On the one hand, this clearly points to the cleanup of the banking system, but on the other hand, indicates that the regulator will have to take greater efforts to uncover such HNC at banks which continue concealing their problems. It is noteworthy that probability values at higher percentiles of distribution, by contrast, rise in time. This shows that model (1)–(2) “learns” with time, making it easier for the regulator to uncover HNC if banks rely on similar methods to hide them.

**Figure 3.** Estimating the probability of detecting HNC at still operating banks (excluding the top-5 government-owned banks)

---

* The total size of HNC is estimated by comparing profiles of banks from the fragile (with HNC) and control (without HNC) groups (Section 2). The fragile group was made up of all banks for which the size of HNC could be found in The Bank of Russia Bulletin by mid-2016 and which were subjected to standard outlier filtration.
In conclusion we will briefly speculate on factors which may have contributed to the above-mentioned positive indirect effect of the Bank of Russia’s new policy. Table 5 presents average values of HNC at banks still in operation for the three periods analyzed above and two possible factors unrelated to the Bank of Russia’s measures. The first of them reflects aggregated macroeconomic conditions, the second represents the profitability of the banking sector itself. Did the second and third periods (II and III) really see a macroeconomic upturn versus the first period (I), which could have triggered a profitability rise (I)? As calculations show, the HNC contraction in the second period took place concurrently with both a macroeconomic downturn and a drop in the profitability of banking operations. In the third period, the situation with profitability was not different from that in the first but the macroeconomic environment at least stabilized at a low level. This means that the indirect effect in question is of course partly due to factors unrelated to the Bank of Russia’s measures. But this part is hardly large enough to account for the magnitude of the indirect effect. We are inclined to believe that this positive indirect effect, after all, stems from the Bank of Russia’s measures.

### Table 5. Bank HNC, profitability of the banking system’s assets, and GDP growth rate

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Total size of HNC at still operating banks (excluding top-5 government-owned banks), % of banking system’s total assets</td>
<td>14.2</td>
<td>9.6</td>
<td>4.2</td>
</tr>
<tr>
<td>12-month moving average profit / Banking system’s total assets, % after loan loss provisions (ROA)</td>
<td>1.8</td>
<td>1.0</td>
<td>1.1</td>
</tr>
<tr>
<td>12-month moving average GDP growth, %</td>
<td>3.6</td>
<td>−0.5</td>
<td>0.5</td>
</tr>
</tbody>
</table>

### 5. Conclusion

What was the size of HNC in the Russian banking system before the Bank of Russia’s senior management changed? Has the financial regulator’s new, tougher, supervision policy launched in mid-2013 been productive in dealing with problems accumulated over recent years? This study attempts to provide initial answers to these questions using Heckman selection models and solving a possible regulator’s optimization problem based on December 2000 – May 2017 Russian bank data. We analyzed two groups of banks: the fragile and reference ones. The former was made up of all credit institutions for which The Bank of Russia Bulletin published official data on uncovered HNC; we fixed this sample in time. The latter, on all dates from December 2009 to May 2017, comprised all banks in operation,
some of which were hiding HNC from the regulator. We varied this sample in time relative to the fixed fragile group and attempted to find out the size of their HNC.

We can sum up our findings as follows. Before the Bank of Russia’s senior management changed in mid-2013, the average level of HNC at banks operating in Russia was very high: 14% of the banking system's total assets on each specific date, and fairly permanent in time. However, as early as half a year after the Bank of Russia’s senior management changed, the overall size of these HNC started contracting rapidly, coming close to 4% by mid-2016. In the last 12 months of the time span under study the level of the HNC stabilized at 4%. These estimates point to the high effectiveness of the new senior management's tough supervision policy. This policy has produced a strong indirect positive effect: part of fragile banks have started addressing their problems without waiting for the Central Bank to show up. The effectiveness of the Bank of Russia’s new policy cannot be attributed to third factors such as macroeconomic stabilization or improvement and/or rising profitability of bank operations. Rather, when supervision policy was stepped up, both of the above third factors were acting in the opposite direction.

6. References


The Folk Theorem of Decreasing Effectiveness of Monetary Policy: What Do the Data Say?'

**Ugo Panizza,** *The Graduate Institute, Geneva and CEPR*

**Charles Wyplosz,** *The Graduate Institute, Geneva and CEPR*

It is increasingly claimed that unconventional monetary policies are subject to decreasing effectiveness in supporting growth and raising the inflation rate. There are good reasons to believe that the effects of further asset purchases by central banks and of moving the interest rate deeper in negative territory progressively decline. But has it been happening? This paper attempts to provide an answer. Looking at the Eurozone, the UK, the US and Japan, it uses different approaches (local projection and Bayesian VAR) on different sub-samples. The evidence is mixed. Policy interventions proxied by the shadow policy rate seem to be subject to the decreasing effectiveness hypothesis. However, this is not the case for QE announcements.

**Keywords:** unconventional monetary policy, zero lower bound, shadow policy rate

**JEL:** E52, E58

1. **Introduction**

After several years of experimentation, a growing number of observers are convinced that unconventional monetary policies are becoming increasingly less efficient (see, e.g., Goodhart and Ashworth, 2012; Krishnamurthy and Vissing-Jorgensen, 2013; Roubini, 2014; Blanchard, 2016a; Oakley, 2016). Central bankers, on the other hand, generally consider that these policies work well and eventually deliver what they are designed for (Draghi, 2016; Gagnon et al., 2016), although

---

1 We are most grateful to Tomasz Wieladek for having generously shared data and for help with coding. We also wish to thank the National Institute of Economic and Social Research and Kiyohiko Nishimura for providing data. We acknowledge useful comments by Jesper Lindé, Lawrence Ball, Joe Gagnon and Athanasios Orphanides.
the lags may be much longer than with traditional monetary policy (Engen et al., 2015). Former central bankers are less sanguine, though (Bernanke, 2016; Ueda, 2012). A poll of British-based economists (Armstrong et al., 2015) shows that opinions on the issue are about equally split. This paper attempts to determine which view is backed by (admittedly short) data.

Nearly ten years after the Global Financial Crisis, growth remains subdued and inflation is significantly below target in most advanced economies. It may be that this state of affairs is unrelated to monetary policy. Slow growth is sometimes justified by the secular stagnation hypothesis or to on-going widespread deleveraging, while low inflation is related to a possibly flatter Phillips curve. Responses to these possibilities do not belong to central banks. Yet, the central banks of the US, Japan, the UK, and the Eurozone (and others as well) are actively engaged in trying to support growth and raise inflation with unconventional policies. It matters greatly whether they have the tools to achieve these objectives. In the presence of decreasing effectiveness, they would have to act increasingly aggressively (Orphanides and Wieland, 2000). Ominously, however, if economic agents think that unconventional polices do not work, inflation expectations could decrease, as they have in Japan. In that case, the task of central banks could become hopeless.

Skeptics argue that unconventional policies have drawbacks. They fear the mispricing of risk, the scarcity of safe assets and, more generally, distortions created by interest rates that remain too low for too long. While central banks seem to believe that the risks are small and the effects large, skeptics take the view that the balance between costs and benefits — or between risk and returns — is increasingly tilting against unconventional policies. This view implies that fiscal policies should take over. Given the high indebtedness of many governments in advanced countries, this means shifting from one series of drawbacks to another one. An often mentioned way-out is helicopter money, fiscal expansions financed by money creation.

For these reasons, it is helpful to assess the decreasing effectiveness hypothesis. While the sample size is limited, there are good reasons to attempt an early evaluation. This is a live issue which is also key for assessing the potential role to be played by fiscal policies. Moreover, unconventional policies are not meant to be pursued for years onward. They have already stopped in the US, where the issue is when and how to exit. This means that the sample size is unlikely to increase much further.

Given the short sample horizon, we need to use monthly data, which inevitably restricts the range of variables that can be examined (for instance, in the case of the

---

2 Bernanke states that “there are signs that monetary policy in the United States and other industrial countries is reaching its limits.”
Eurozone and Japan we need to use industrial production instead of GDP). Moreover, macroeconomic phenomena are slow. Standard monetary policy, for instance, is understood to affect the real economy after several quarters and unconventional policies may take longer to make a difference (Engen et al., 2015). It follows that monthly data may be more noisy than desirable. This explains why most studies so far have looked at the effects of unconventional policies on financial markets, which react fast. These studies have detected strong effects, but that is only the first step in actually fostering growth and raising inflation. We, instead, follow Weale and Wieladek (2015), Garcia Pascual and Wieladek (2016), and Haldane et al. (2016) and directly examine the impact of unconventional policies on growth and inflation. We contribute to the literature by providing a direct test for the decreasing hypothesis and experimenting with alternative measures of unconventional policies. We finds some evidence of decreasing effectiveness although the statistical significance of these results is limited, marred as it is by the short time span under study.

The paper is organized as follows. The next section briefly reviews the channels of effects of unconventional policies, and informally provides arguments for the decreasing policy effectiveness hypothesis. Section 3 summarizes recent empirical research designed to measure the effects of unconventional policies and explains our own strategy. Section 4 presents a battery of tests designed to determine the empirical validity of the hypothesis in the case of the US, Japan, the UK, and the Eurozone. The last section concludes.

2. Channels of Effects and the Decreasing Effectiveness Hypothesis

A growing body of literature describes and evaluates how unconventional monetary policies affect the economy and identifies the potential channels though which these policies operate\(^3\). This section examines reasons why decreasing effectiveness could set in. It follows the traditional analysis of standard monetary policy that considers the direct interest rate channel, the confidence channel, the portfolio balance channel, the bank credit channel, the sovereign credit risk channel, and the exchange rate channel.

Decreasing effectiveness means that the effect of the relevant instrument declines as it is used more intensively. It is an assertion that the relationship \(y = f(x, Z)\) between the instrument \(x\) and the sought-after effect \(y\) is not

\(^3\) The earlier literature focused on quantitative easing. Contributions include Bernanke et al. (2004) and Krishnamurthy and Vissing-Jorgensen (2013). Recent reviews include Borio and Zabai (2016), Fratzscher et al. (2014), Santor and Suchanek (2016), and World Bank (2015).
necessarily linear and, at any rate, becomes concave downward when \( x \) increases sufficiently, that is \( \partial f/\partial x > 0 \) for low (normal) values of \( x \) and \( \partial^2 f/\partial x^2 < 0 \), at least beyond some threshold\(^4\). In the present case, we will estimate \( \partial f/\partial x \) over different samples and check whether the instrument loses potency over time. We do not attempt to identify the channels involved, leaving it for further work.

2.1. Non-standard monetary policies

Unconventional monetary policies start when the policy interest rate approaches the Zero Lower Bound (ZLB). Having reached the ZLB, a number of central banks have looked for other ways to lift the inflation rate toward its target and promote growth.

The two main instruments have been Quantitative Easing (QE) and negative interest rates. Central banks have supplemented these instruments with forward guidance, essentially committing to keep the policy rate at the ZLB or below for a long period of time and announcing a precise schedule of asset purchases spread over several quarters. We look at these two main instruments and their channels of transmission and ask whether there exist plausible reasons for entertaining the presence of decreasing effectiveness.

Negative interest rates

Negative interest rates are often seen as a straightforward extension of standard policy, whereby there is no reason to see zero as particularly meaningful. In addition, what matters for economic activity is the real interest rate, which can be and has been negative independently of the level of the nominal rate. In fact, negative nominal interest rates are all but straightforward (Coeuré, 2016).

Cash money bears zero interest, which is why it is presented as a dominated asset. The usual explanation of why money exists at all is that it yields an implicit positive return in the form of transaction services. It follows that any asset that can be freely exchanged for money must offer a higher, i.e. a positive yield. Since central banks normally set the interest rate by intervening on very short-term maturity assets, they should not be able to bring the policy rate below zero. We now know that this conclusion is incorrect. The reason is that cash is costly to hold in large amounts because of storage and insurance costs. These costs, however, are finite and not very large. The implication is that there is a limit to how negative interest rates can be: the zero lower bound can be breached, but there exist an

\(^4\) In this example, an increase of \( x \) is meant to capture an expansionary unconventional monetary policy. If the instrument is the policy interest rate and the outcome is GDP growth or inflation, we should write \( \partial f/\partial x < 0 \) for normal values of \( x \) and \( \partial^2 f/\partial x^2 < 0 \), at least below some threshold.
effective lower bound (ELB). The position of this bound will remain unknown until it is breached. Even so, it is likely to differ from country to country depending on the structure of the banking system, taxation, regulation, etc.

Quantitative easing

Quantitative easing (QE) includes many procedures and official names\(^5\). The defining purpose is to inject, without sterilization, large amounts of liquidity by buying assets held by banks or by investors according to a preannounced purchase program. The assets usually are chosen to be safe, normally Treasury debt. Central

![Figure 1. The Fed's balance sheet and CPI, January 2003 = 100](image)

\(^5\) QE is called Large Scale Asset Purchases (LSAT) in the US, Asset Purchases Program (APP) in the Eurozone, Quantitative and Qualitative Easing (QQE) in Japan. Related interventions are Long term Refinancing Operations (LTRO) in the Eurozone, but these are not QE proper, as they focus on improving intermediation. The increase in the size of the balance sheet is a, possibly unwanted, by-product, that is demand, not supply-driven.
banks have also bought privately issued assets either because they wanted to enhance monetary policy transmission by removing risk from bank balance sheets or because they wanted to reduce the borrowing costs of corporations. The counterpart of these purchases is base money creation.

The relevant monetary theory principle is the assertion that money and the price level grow hand in hand. QE, therefore, should lead to an acceleration of inflation. Figure 1 shows the case of the US. The upper chart indicates that the relationship was largely verified until QE was started, and then broke down. The lower chart focuses in more detail over the period 2003-2006. It shows that the link started to deteriorate before the crisis. The gradual deterioration of the relationship and its full-blown breakdown suggest serious limits to elementary principles.

Central banks can expand the monetary base at will when the interest rate is at the ZLB (Borio and Zabai, 2016), but monetary policy requires that monetary aggregates also increase, or that interest rates decline along the yield curve, or both, as discussed below. The link from the monetary base to the monetary aggregates is the well-known multiplier. If, as shown in Figure 2, the multiplier declines toward zero, a key link disappears.

The link to prices and output operates through increased spending. Under normal conditions, spending responds and the Phillips curve mechanism delivers the expected effect. Yet, there may be conditions when monetary aggregates fail to raise spending and when the effect of economic activity on inflation may be limited. The apparently feeble impact of economic activity on inflation is the object of current research. Ball and Mazumder (2011) find that the slope of the Phillips curve declines when inflation is low and stable. Blanchard (2016b) points to the role of strongly anchored expectations and to the general imprecision of the

Figure 2. M2 and the Fed’s Balance Sheet, January 2003 = 100

Source: FRED, Federal Reserve Bank of Saint Louis
relationship. To understand why the economy may fail to respond to monetary aggregates, or responds with decreasing effectiveness, it is necessary to examine the various channels of unconventional monetary policy.

2.2. The direct interest rate channel

A lower interest rate is expected to increase spending by households and firms by encouraging borrowing as it affects the intertemporal price of consuming and saving. Additionally, the exchange rate is expected to depreciate, which should raise net exports.

Low and negative interest rates

Long-term rates matter more than short-term rates. As a consequence, the short-term policy rate can have an effect on economic activity only if lowers the whole yield curve. This, in turn, requires that market expectations concerning the future evolution of the policy rate, both its level and the duration of the current policy stance, respond to monetary policy decisions. When the policy rate is perceived to be close to the ELB, the only direction of substantial future changes in the policy rate is upward. This situation prevents a further lowering of the longer end of the yield curve. This is why central banks have developed the signaling channel as they wish to indicate that the policy rate will remain low or negative for a long period. This has worked well, as indicated e.g. by Bean et al. (2015). Yet, the presence of a lower bound suggests that there is increasingly less room for the level effect, leaving only the duration effect.

However, the adverse effects of low or negative interest rates could hamper the duration effect. These adverse effects include aggressive risk-taking by investors and financial institutions eager to achieve better returns, more difficult price discovery (Santor and Suchanek, 2016), the risk of disintermediation, and reduced bank profitability (Brunnermeier, 2016). If these adverse effects do not last for too long, they are likely to be harmless. However, persistent risk-taking, mispricing and reduced profitability are bound to become toxic and, possibly, contractionary (Brunnermeier and Koby, 2016). At best, steps taken by financial institutions to mitigate growing vulnerabilities will reduce monetary policy effectiveness, possibly reversing the impact when markets increasingly expect that the policy rate will be raised and the yield curve will steepen. At worst, market fragility may lead to a crisis.

Even if the yield curve can be kept flat, possibly in negative territory along the ELB, the question is what happens to the economy. An important issue is why the previous interest rate reductions — those that brought the policy rate to the ELB —
have failed to lift spending. A variety of reasons have been invoked. Potential borrowers may be over-indebted and potential lenders may be over-leveraged. Lower rates shift income from lenders to borrowers, which may help with over-indebtedness but at the expense of over-leveraging. Under these conditions, it is unclear why lowering the interest rates further would be more successful than previous reductions.

In addition, intertemporal substitution increases spending today at the expense of spending tomorrow. To sustain spending next year, a new lowering of the interest rate is needed. When the end of lowering interest rates is nearing, because the ELB is close by, further intertemporal substitution becomes impossible. Keeping the interest rate low does not provide continuing stimulus. Indeed, time must come when spending falls because it has been brought forward previously.

Another aspect is the pass-through of lower wholesale interest rates to rates applied to consumers and firms. Bean et al. (2015), among others, report that the pass-through is limited when the policy rate is close to zero. In some countries, banks cannot charge a negative interest rate. Where they may, they are reluctant to upset commercial relationships with customers. In addition, a flat yield curve erodes the profits from maturity transformation. If competition is imperfect, banks may elect to not lower their lending rates much.

Quantitative easing

Large asset purchases by the central bank raise asset prices and therefore lower their interest rates. The intention is to encourage borrowing or share issuance to prompt more spending by consumers and firms through lower borrowing costs and wealth effects. As stated, quantitative easing does not look different from standard open market operations. In that case, the impact from lower short-term policy rates to relevant longer-term rates is driven by market expectations of future policy rates, which can be strengthened through signaling. Other asset classes are affected by portfolio rebalancing, see next channel. However, quantitative easing differs from standard policy because central banks can choose which assets and amounts they purchase, thus acting directly on particular interest rates and asset classes.

By and large, the reasons why policy effectiveness might be decreasing are similar to those examined in the previous section. An additional consideration is that the volume of assets to be bought is finite. If the central bank holds a significant share of some assets, the relevant market becomes shallow and therefore less efficient. It follows that there is limit to what central banks can purchase, aptly called the Effective Quantitative Bound (EQB) by Santor and Suchanek (2016). As
we move closer to the (unknown) EQB, markets may anticipate that QE is reaching its limits. Furthermore, the associated uncertainty can raise risk premia across the board.

Another often noted side effect is the emergence of asset price bubbles. Mechanically, QE raise asset prices. These prices can remain at their fundamental level as long as they are matched by the present value of future earnings evaluated at prevailing low interest rates. Yet, because QE is by definition temporary, asset prices are expected to fall back once liquidity is being removed. Thus the long-run resale value must be lower than current prices, with considerable uncertainty as to when exit will take place and at which pace. The result is that the wealth effect may become smaller as asset prices rise and that the expectation of a subsequent crash may build up as QE unfolds.

2.3. The portfolio balance channel

As any other price change, a lower policy rate is transmitted to other returns through a general equilibrium effect as investors rebalance demand for all asset classes. The portfolio balance channel captures investors’ willingness to take on more risk when the policy rate declines and risk premia decline across the board.

Low and negative interest rates

Textbook models show that policy rate reductions lead to portfolio reallocation which ultimately result into higher asset prices. The question is whether this mechanism becomes less effective when the policy rate approaches zero or enters into negative territory. In other words, what can challenge the presumption that changing the policy rate triggers asset price changes and generalized portfolio reallocation when the rate is announced to remain low, or even negative, for long?

A first question concerns the pass-through from the policy rate to asset returns. While sovereign bond yields have responded to policy rate cuts and sometimes went into negative territory, corporate bond yields have not fully followed. Moreover, nominal illusion could limit the ability to issue bonds with negative rates below a certain threshold.

Finally, in principle, asset prices are the present discounted value of future earnings. These (mostly implicit) calculations are problematic when the discount rate is negative (or even close to zero), if only because the present value may

6 For instance, over 2008–2016, the spread between Aaa 10 year US corporate bonds and 10 US Treasuries has ranged between 150 and 250 basis points, with an average of 185 basis points. The long run (1953–2016) average of this spread is 93 basis points and the 1953-2008 average is 76 basis points.
become unbounded. This is one reason why markets fear “bubbles”, in fact major uncertainty rather a “bubblish” equilibrium. A possible reaction is that the attendant uncertainty results in higher risk premia and therefore a muted response as the de facto discount rate does not decline as usual. Although, as indicated below, there is substantial evidence that the pass-through has remained active, these features suggest that its size may become increasingly smaller resulting in decreasing effectiveness.

Quantitative easing

As some assets are withdrawn and their prices increase, investors rebalance their portfolios by acquiring other assets. If the supply of these other assets does not increase to meet higher demand, their prices increase and their yields decline. This is indeed a key channel of transmission. Central banks rely upon this channel to spread the impact of QE through lower term and risk premia.

The possible downside is that risk becomes mispriced and that investors — including banks, insurance companies and pension funds — ride up the flattened risk-return schedule with improper understanding of risk being taken or by accepting risk because of pressing search for yield. As QE unfolds, more risk may be taken. This may not result in immediate loss of effectiveness but into subsequent financial difficulties once QE stops and is reversed.

2.4. The confidence channel

Unconventional policies are meant to signal that the central bank stands ready to take exceptional action when other standard means are no longer available. In addition, the ability to target specific assets implies that the central bank can address specific financial concerns, as with TARP in the US. Both readiness to act and the emergence of new instruments could help dispel fears on financial markets and beyond that serious problems remain unaddressed (Curdia and Woodford, 2011). The confidence channel is expected to boost spending, both directly by improving expectations, and indirectly by reducing interest rate risk premia.

Since what matters is the signaling impact of the policy measures, we can consider negative rates and QE together. Both policies face the same challenge that the room for maneuver declines as the policies unfold. Negative rates are limited by the unknown ELB, QE is ultimately limited by the size of the markets. These Knightian limits are bound to gradually erode policy effectiveness. They could even make nonstandard policies ineffective if faced with the perception that the end of the road is in sight and that the central bank has reached its limits and can no longer implement confidence-building measures.
One specific aspect of QE is that asset purchases can be used to repair malfunctioning markets where demand is reduced by fears of impending crisis. When confidence has returned in the affected markets, QE stand to be less efficient. This was the case in 2007–2009. Indeed several papers (Krishnamurthy and Vissing-Jorgensen, 2013; Fratzscher et al., 2014) detect such an effect early on, but none once markets were stabilized.

2.5. The bank credit channel

The literature on the credit channel (or financial accelerator, Bernanke and Gertler, 1995) argues that ample availability of liquidity encourages banks to increase lending, over and above its impact on the interest rate, meaning that this channel is specific to QE. When banks hold large amount of cash, they have a choice of holding it as part of their reserves at the central bank, or to lend it to customers. Under the assumption that bank-lending rates exceed central bank deposit rates, it is expected that banks will prefer to lend to customers. Easier credit terms, in turn, are expected to encourage spending by both households and firms.

The bank credit effect could be muted when banks and/or their customers are highly leveraged. Thus, in contrast to the confidence channel, QE might become more effective when the situation improves.

Decreasing effectiveness may arise if nonstandard policies lead banks to chasing increasingly less reliable borrowers. In that case, either lending rates will rise or lending will start decreasing. Indirect evidence is provided by Figure 2, which plots the size of the balance sheet of the Federal Reserve — already shown in Figure 1 — and the monetary aggregate M2. The ratio between these two measures is the money multiplier, which captures the banking system lending response to liquidity injections by the central bank. As is well known, the money multiplier has declined considerably since the start of QE: the ratio of M2 to the size of the Fed’s balance sheet has declined from an average of 8.1 over 2003–2007 to 3.7 over 2009–2012 and 2.8 since 2013.

2.6. The sovereign credit risk channel

The interest rate paid by the Sovereign usually serves as the basis for all interest rates in the country. This is because sovereign debt is, rightly or wrongly, treated as a safe(r) asset. It follows that any policy action that lowers ceteris paribus the interest rate on the sovereign debt is potentially equivalent to a direct reduction of the policy rate. Government debt purchases by the central bank in effect reduce public sector indebtedness because debt service on purchased instruments implies
payments to the central bank, which rebates profits to the Treasury. Equivalently, the public sector swaps interest-bearing debt for zero interest-bearing money. Insofar as a lower debt reduces the sovereign risk premium, this is an additional channel.

The channel faces limits, though. The temporary nature of QE means that the corresponding amounts of public debt are not fully written down; debt service is reduced only as long as QE is in place. The resulting reduction in risk premia is bound to be reversed as the end of QE becomes nearer. The “taper tantrum” of 2013 could be a manifestation of this effect (Foerster, 2014).

2.7. The international portfolio balance and exchange rate channels

Finally, the portfolio rebalancing process means that residents will acquire foreign assets, possibly in amounts commensurate with the central bank purchases. This should lead to a depreciation of the exchange rate. In small open economies, the exchange rate channel is usually the main channel of standard monetary policy.

Low and negative interest rates

In principle, the effect of the policy interest rate on the exchange rate is independent on its size.

Quantitative easing

The experience so far is that the large countries have adopted QE policies more or less at the same time. As a result, their effective exchange rates have not moved much (Caballero et al., 2016; Eggertson et al., 2016). More significant movements have affected a host of other countries, but this is not an issue related to this paper.

3. Overview of the Evidence

A growing number of studies are attempting to evaluate the effectiveness of unconventional monetary policies. With a very short sample period, much of the early work has relied on high frequency observations and focused on the effect of nonstandard policies on various asset prices or on measures of asset price volatility.

The impact of macroeconomic variables, which move much more slowly, has been examined only recently⁷.

⁷ A number of studies are summarized in Den Haan (2016) and Borio and Zabai (2016).
Given the unsettled economic and financial situation, a key challenge is to disentangle the influence of policy decisions from other contemporaneous shocks, including those that affect financial markets and lead central banks to enact non-standard policies. As noted by Borio and Zabai (2016), two methods have been used.

First, using various types of VAR analysis, some authors test whether non-standard policies affect the variables of interests. To be valid, these tests should not include observations that predate the adoption of non-standard policies since these policies are designed to significantly alter the operations of financial markets. Identification is particularly challenging because policy announcements are likely to be more important for asset markets than actual implementation. For instance, daily or monthly asset purchases may have little impact once the schedule of purchases has been announced. This is why more recent papers have focused on policy announcements and event studies.

### 3.1. Effects of QE on Financial Markets

Empirical evidence from the US, the UK and the Eurozone strongly backs the existence of an effect of QE on asset prices and interest rates at all maturity. In the case of the US, Krishnamurthy and Vissing-Jorgensen (2013) provide early evidence that the portfolio balance channel has been effective. This evidence has been confirmed by a large number of studies, including Engen et al. (2015), Gagnon et al. (2011) and Wu and Xia (2015). Qualitatively similar results are reported for the UK by Kapetanios et al. (2012) and the Eurozone by Frazscher et al. (2014).

Other channels have also been identified. Bauer and Rudebsuch (2014) separate out the portfolio balance channel from the interest rate channel in the US and conclude that both have been operative. Altavilla and Giannone (2015) reports evidence on the interest channel via private forecasters in the US and the Eurozone. There is also some evidence that the risk channel operates in the US and the Eurozone (Carpenter et al., 2013).

The bank liquidity channel is supported by studies that focus on the early phase of the financial crisis (e.g., Krishnamurthy and Vissing-Jorgensen, 2013, for the US, Darracq-Paries and de Santis, 2013, in the case of the Eurozone). Hong (2016) documents the lengthening of the maturity of bank loans in the US.

### 3.2. Effects of QE on Output and Inflation

Fewer studies so far looked at the ultimate criterion of success of non-standard policies, namely whether they boost output and inflation. An important roadblock...
has been the measurement of unconventional policies. Macroeconomic variables cannot be measured at high frequency so that the small sample period available implies that too few observations are available. The solution is to extend the sample back to periods when unconventional policies were not in use. In principle, one can measure both the policy rate and the size of central bank balance sheets over long periods but it is likely that the adoption of unconventional policies represents a regime change.

A breakthrough has been the estimation of a shadow policy rate by Wu and Xia (2015). They use the term structure to model the shadow policy rate as a function of standard factors and allow this rate to diverge from the near-zero actual policy rate. Of course, the estimated shadow rate is subject to a number of assumptions. Using this shadow rate, Ball et al. (2016) shows that QE in the US has reduced shadow policy interest rates by 200 to 300 basis points and conclude that there must have some output and inflation effect. Churm et al. (2015) perform a similar exercise for the UK.

More direct evidence is provided by a few studies that include output and/or inflation in various VAR formulations. Baumeister and Benati (2010) use the long-term interest spread at unchanged policy rate as a measure of QE in a time-varying SVAR that also include GDP growth and the GDP deflator inflation. They report a significant macroeconomic impact in the US, UK and the euro area. Weale and Wieladek (2015), which is the paper closer to ours, use Bayesian VAR (BVAR) model that includes monthly data for GDP, CPI, asset prices and long-term rates and show that asset purchase announcements have a statistically significant effect on prices and GDP in both the US and the UK. Garcia Pascual and Wieladek (2016) apply a similar methodology to the Eurozone and also find a significant effect of quantitative easing on prices and output. Haldane et al. (2016) evaluate the early evidence on quantitative easing (besides the UK, US and Eurozone, they also study the cases of Canada, Japan, and Sweden) corroborate the finding that quantitative easing has significant effects on prices and output. They also show that the QE is more effective in the presence of financial frictions.

### 3.3. Negative Interest Rates

The experience with negative interest rates is even more recent than with QE, making it challenging to seek formal evidence. The focus so far has been on the transmission of negative rates on financial conditions. Bech and Malkhozov (2016) informally find that negative policy rates have been transmitted to most other interest rates, with some exceptions like mortgage rates. Coeuré (2016) describes

---

9 Along similar lines, Vitale (2016) shows that monetary policy is more powerful when firms face financial distress.
the ECB policy and accepts that adverse effects may set in because banks may eventually have to change their business models and, mostly, because there is still a minimum interest rate.

3.4. Non-Linearities and the Decreasing Effectiveness Hypothesis

A number of papers have started to explore the decreasing effectiveness hypothesis, following the initial hint by Goodhart and Ashworth (2013) and the early observation by Krishnamurty and Vissing-Jorgensen (2012). Barnichon et al. (2016) examine whether the financial accelerator mechanism may lead to non-linearities. They find that increased credit supply has a stronger output effect in periods of weak growth when credit is constrained than during period of fast growth. They use data for the US on a long sample that includes the post-2008 period. Their results can be seen as implying that QE may have been more effective early on when financial markets were impaired and illiquid than when abundant liquidity had already been created (this finding is consistent with the results of Haldane et al., 2016, mentioned above). Similar results are reported for Japan's experience in the 2000s by Ueda (2012). For the QQE experiment started in 2013, Nishimura (2016) develops the view that “policy exhaustion” has set in, mostly in Japan but also in the US.

Kapetenios et al. (2012) conduct a counterfactual analysis of the Bank of England policies. They use three different VAR models (Bayesian over different windows, Markov switching, and time-varying parameters) estimated over long periods. They compare the impulse responses of GDP and inflation when long-term yields are reduced by 100 basis points. While these authors use time-varying estimates up to September 2010, they do not focusing on detecting changes within the QE period.

To the best of our knowledge, this paper is the first attempt to detect changes within the QE period.

4. Empirical Results for the US, UK, Eurozone and Japan

There are two challenges associated with testing the hypothesis that unconventional monetary policies have decreasing effectiveness on economic activity. First, the short time span of these policies prevents us from using quarterly data, as it is commonly done in the monetary policy literature. Second, there is no commonly agreed quantitative indicator of unconventional monetary policies.

---

10 The contribution by Demertzis and Wolff (2016) is closest to our own investigation in its intentions. Informally, it interprets the evidence as indicating that the ECB’s QE had a positive effect on investment spending after its announcement in 2015 but that the increased decided in early 2016 did not produce such an effect.
We address these challenges by using monthly data and two different proxies for non-standard monetary policies. Even with monthly data, we have limited degrees of freedom (especially when we try to assess whether the parameters vary over time). The point estimates may thus be sensitive to small changes in the sample or estimation technique. We address this issue by comparing the results of two different econometric approaches.

4.1. Measuring unconventional monetary policy

The policy stance of the central bank is normally measured with a policy interest rate (like the Fed Fund rate in the US) or with a monetary aggregate (such as M1 or M2, as originally suggested by Friedman and Schwartz, 1963). The presumption is that these instruments can be controlled by the central bank and are stably related to economic activity (Leeper et al., 1996). Unconventional monetary policy starts when the central bank loses control over these instruments, either because the policy rate approaches the effective lower bound (Figure 3a, b) or because the main monetary aggregates no longer respond to the injection of base money (Figure 2). This is why the evaluation of unconventional monetary policies requires alternative indicators of policy stance.

Our first indicator of unconventional policy is the “shadow” (i.e., non-observable) policy rate derived from the yield curve. The shadow interest rate is an attempt at measuring how the policy rate would look like if the central bank had been willing (and able) to adopt a negative policy rate. One way to recover this shadow rate is to build a statistical model of the lower end of the yield curve and then use the model to build a counterfactual estimate for the overnight rate.\footnote{An alternative approach is to use a large number of variables and then estimate the shadow rate by applying dynamic factor analysis to these variables (Lombardi and Zhu, 2014).}

The idea of using the yield curve to build a shadow term structure is due to Black (1995) and was recently operationalized for the US, UK, and the Eurozone by Xia and Wu (2016).\footnote{Other papers that use the term structure to estimate a shadow policy rate include Christensen and Rudebusch (2015) and Krippner (2013).} Xia and Wu (2016) show that when the US policy rate is above 50 basis points, their estimate for the US shadow rate coincides with the observed rate (Figure 3a). The coincidence, instead, is not perfect for the UK and the Eurozone. In the case of Japan, instead, our shadow rate data start in 1995 and coincide with the period in which the Bank of Japan started adopting a policy of very low interest rates (0.5% between 1995 and 2001 and 0.1% between 2001 and 2006).\footnote{We source shadow rate for the Eurozone, UK and US from Xia and Wu (2016) and Japanese shadow rates from IMF (2015, Box 3.1).}

Xia and Wu (2016) also show that the shadow rate can be used to assess the effectiveness of monetary policy when the observed rate approaches zero. One advantage of using the shadow rate as a measure of policy stance is that this
Figure 3a. Actual and Shadow Policy Rate

Figure 3b. Shadow Policy Rate and Adjusted Shadow Policy Rate
indicator is directly comparable with policy rates in normal times. One challenge related to the use of this measure is that the shadow rate is model-specific and that different specifications can yield alternative measures of the shadow term structure (Christensen and Rudebusch, 2015).

Another potential problem with the shadow rate is that any event that affects the slope of the yield curve will also have an impact on the estimated shadow rate, even if this event has nothing to do with monetary policy actions. While the unobservable shadow rate is not a policy instrument (it is a proxy for a series of unconventional monetary policy instruments, including the announcement of asset purchases), it can be treated as a policy instrument if we assume that policymakers target the shadow rate and reverse-engineer non-standard monetary policies to achieve this target\(^{14}\). Changes in the shadow rate could therefore be decomposed into changes linked to policy decisions and changes that are not directly related to policy decisions and are instead due to other factors that shift the slope of yield curve. To separate these two components, we compile a monthly dataset of monetary policy announcements (they could be changes in the policy rate, announcements of asset purchases, forward guidance, announcements about changes in monetary policy strategy) and build a proxy for the policy component of the shadow rate by regressing the shadow rate on a set of dummy variables that take a value of one during months for which there was a monetary policy announcement.

Formally, suppose that in a given country we identify \(N\) policy announcements, then we estimate the following model:

\[
SR_t = \alpha + \sum_{i=1}^{N} \beta_i D_{i,t} + \varepsilon_t
\]

where, \(SR_t\) is the shadow rate in month \(t\) and \(D_{i,t}\) is a set of dummies associated with policy announcements (we include as many dummies as policy announcements)\(^{15}\).

Finally we use the predicted values of this model to build an adjusted shadow rate that only varies in months characterized by a policy announcement. The adjusted shadow rate plotted in Figure 3b is defined as:

\[
ASR_t = \hat{\alpha} + \sum_{i=1}^{N} \hat{\beta}_i D_{i,t}
\]

We conduct all of our tests using the adjusted policy rate described in Equation (2). The results, however, are robust to using the unadjusted policy rate.

\(^{14}\) We would like to thank an anonymous referee for pointing this out.

\(^{15}\) Assume that there are 3 monetary policy announcement, the first in month \(t\), the second in month \(t+s\), and the third in month \(t+s+z\). \(D_1\) will take value 0 from month 1 until month \(t-1\), it will then take value 1 from month \(t\) to month \(t+s-1\), and value zero from month \(t+s\) to month \(T\). \(D_2\) will take value 0 from month 1 until month \(t+s-1\), it will then take value 1 from month \(t+s\) to month \(t+s+z-1\), and value zero from month \(t+s+z\) to month \(T\). \(D_3\) will take value 0 from month 1 until month \(t+s+z-1\) and value 1 from month \(t+s+z\) to month \(T\).
For our second indicator of unconventional policy, we use the announcements of large scale asset purchases (quantitative easing) by the Federal Reserve and the Bank of England. While this indicator has the advantage of being based on an observable measure of unconventional policy (unlike the shadow policy rate, it is not model specific), it has the disadvantages of not being directly comparable with standard monetary policy instruments and of not incorporating policy measures (such as forward guidance or the “whatever it takes” speech of July 2012) that go beyond and above asset purchases.

As the experience of the European Central Bank with quantitative easing is too short for our own purposes of evaluating potential decreasing effectiveness to this policy (see, however, Garcia Pascual and Wieladek, 2016, for an evaluation the macroeconomic effects of quantitative easing in the Eurozone), we experiment with the size of balance sheet of the European Central Bank. We do the same for Japan, but for the opposite reason. Since 1999, the Bank of Japan has sporadically expanded its balance sheet but changing clarity in its communication severely hampers the use of QE announcements by the Bank of Japan.

Figure 4 shows the evolution of the balance sheets of the US Federal Reserve, Bank of England, and European Central Bank. The figure also plots the data for asset purchase announcements.

**Figure 4. QE Announcements and Central Bank Balance Sheets**
4.2. Econometric Strategy

We follow Weale and Wieladek (2015) and identify the output and price effects on non-standard policy with a Bayesian VAR with sign restrictions. In the Appendix, we summarize results obtained with an alternative approach, local projections.

There are tradeoffs related to the use of these estimation techniques. Local projection requires identifying the policy instrument to solve the standard endogeneity problem. The Bayesian VAR has the advantage of allowing modeling the complex interactions among all variables but it absorbs a larger number of degrees in freedom, which is problematic as the sample size of QE periods is limited. Multicollinearity among the variables included in the VAR can amplify the problems associated with limited degrees of freedom. These problems could be attenuated by imposing tight priors in the BVAR estimates. However, tight priors also have costs because they may not let the data speak.

Specifically, we use monthly data to estimate the following reduced-form VAR model:

\[ Y_t = A + B_1 Y_{t-1} + B_2 Y_{t-2} + E_t \]  \hspace{1cm} (3)

Where \( Y \) is a vector of the following six variables: (i) the log of monthly GDP; (ii) the log of the consumer price index (CPI); (iii) the log of real equity prices (measured by dividing a stock market index by the CPI) (iv) the yield on ten-year government bonds, (v) the log of the effective exchange rate and (vi) a measure of unconventional monetary policy. With respect to this last variable, as mentioned above, we use alternatively the adjusted shadow policy rate and announcements of asset purchases. \( B_1 \) and \( B_2 \) are matrixes of parameters associated with the lagged dependent variables and \( E \) is a vector of normally distributed residuals with mean zero and variance covariance matrix \( \Sigma \).

In order to interpret the results of the reduced form VAR, it is necessary to impose a set of identifying assumptions that allow recovering the structural parameters from the estimates of \( B_1, B_2 \) and \( \Sigma \). We adopt the same sign restrictions as in Weale and Wieladek (2015) in assuming that expansionary policy at the zero lower bound signals that expansionary unconventional policy reduces long-term interest rates and increases real equity prices at impact and also for the month that follows the policy action. We also follow their strategy of identifying demand and supply shocks: with a positive demand shocks both output and prices increase, and this also leads to an increase in the long-term interest rate and real equity prices;

\[^{16}\text{These problems are compounded by the fact that monthly GDP data are noisier than the quarterly data which are normally used to estimate the effect of monetary policy.}\]
with a positive supply shock, income increases, prices decrease, and long-term interest rates and real equity prices increase.\footnote{\textsuperscript{17}}

When we estimate the Bayesian VAR with announcements of asset purchases, our model is almost identical to that of Weale and Wieladek (2015). There are, however, three notable differences. First, we assess the decreasing returns hypothesis by estimating the model over different subsamples. Second, we augment their model with the log of the effective exchange rate (their model only includes 5 variable). While including an additional variable has a cost in terms of degrees of freedom, we feel that it is important to control for the exchange rate which has been found to be an important channel of effect of conventional and unconventional monetary policies (see, e.g. Caballero et al. 2015)\footnote{\textsuperscript{18}}. Third, we study more countries (or group of countries, in the case of the Eurozone) and have a slightly longer sample (approximately one extra year of data).

When we estimate how the adjusted shadow rate affects prices and GDP, we add one restriction to those listed above. Specifically, we restrict the contemporary impact on prices and output of a reduction of the adjusted shadow rate to be non-negative.\footnote{\textsuperscript{19}} While this additional restriction rules out that an expansionary monetary policy (as captured by the adjusted shadow rate) will lead to lower output or prices on impact, it does not rule out the possibility that expansionary policy will have a negative effect on output and prices starting one month after the implementation of the policy. As we maintain this restriction for all subperiods for which we estimate the model, the restriction does not affect the relative magnitude of the output and price responses across estimation periods (which is what we are interested in).

We use the same uninformative priors as Weale and Wieladek (2015) who, in turn, follow Uhlig (2005) in assuming that the priors are drawn from a Normal–Wishart density multiplied with an indicator variable that takes a value one when the impulse response satisfies the sign restriction. Therefore, the priors and identification strategy are not exogenously chosen for each country. They are instead based on the country-specific impulse response functions that satisfy the sign restrictions.

Testing for the decreasing effectiveness hypothesis

Going back to the discussion of Section 2, we use different subsample to look at how \(\frac{\partial f}{\partial x}\) varies when \(x\) surpasses a certain threshold or to check if the value of \(\frac{\partial f}{\partial x}\)\footnote{\textsuperscript{17}} Weale and Wieladek (2015) experiment with 4 alternative identifications schemes. The first is a simple Cholesky ordering, the second imposes a series of sign restrictions, the third jointly uses sign and zero restrictions, and the fourth uses sign and variance decomposition restrictions. In our baseline model we use the second identification scheme, but we also experiment with the third identification scheme.\footnote{\textsuperscript{18}} As we are only interested in estimating the impact of nonstandard policy on the other 5 variables, the inclusion of the exchange rate does not require additional identification restrictions.\footnote{\textsuperscript{19}} If we do not impose this restriction, we obtain imprecise estimates (i.e., much larger confidence bands) across subperiods.
is affected by the duration of the nonstandard policy. There are at least two possible ways to test for the decreasing effectiveness hypothesis. One possibility is to explicitly allow for non-linearities by, for instance, including the square of the adjusted shadow rate or squared asset purchases\textsuperscript{20}. Another possibility is to check whether the effectiveness of unconventional monetary policy has decreased over time by splitting the sample into different sub-periods. We opt for the second option because it is difficult to explicitly include non-linearities in the VAR framework described above. Moreover, it would be problematic to work with the square of the shadow policy rate as the shadow rate can take on negative value and a quadratic model would imply that a shadow rate of, say, 2\% is as effective as a shadow rate of minus 2\%.

When we work with asset purchases announcements in the US and the UK, we first estimate the model using data for the full QE period (2009M1-2015M11 for the US and 2009M1-2016M5 for the UK) and then use thresholds that are based on the size of asset purchases. We choose sub-periods of similar length but that also coincide with different phases of unconventional monetary policy. For instance, for the US the first period covers QE1 and QE2 and the second period encompasses Operation Twist and QE3. In the case of the UK, instead, the first period covers QE1 and the second period QE2 and QE3.

In the case of Japan, instead, we use data on central bank balance sheet (always scaled by GDP) for the first (2001M1-2008M7) and second (2008M8-2016M5) QE experiments and then we further split the second period into sub-periods of similar length (2008-2013 and 2012-2016). We also experimented with the balance sheet of the European Central Bank and different sample splits. However, either the program did not converge or produced unstable results.

When we work with the policy rate and the adjusted shadow policy rate, we start by estimating the model with the actual policy rate in the pre-nonstandard policy period (1995-2008 for the US and the UK, 1999-2008 for the Eurozone, and 1978-2000 for Japan). Next we estimate the model for the post Global Financial Crisis period (2008-2016) for the US, UK, and Eurozone, and for the full QE period for Japan (2001-2016). Finally, for the US, UK, and Eurozone we split the nonstandard policy subsample in two sub-periods of similar length (2008-2012 and 2011-2016) and for Japan we split the sample between the first and second phase of nonstandard policy (2001-2005 and 2005-2016)\textsuperscript{21}.

Note that when we estimate the models for the US, UK, and Eurozone we include data up to 2012M6 in the first sub-period and start the second sub-period in 2011M6. This small overlap between the two subsamples allows us to have at

\textsuperscript{20} For the adjusted shadow rate one could define the following variable: $\SASR = N \times ASR^2$, where $ASR$ is the adjusted shadow rate and $N = -1$ when $ASR < 0$ (with $N=0$ when the adjusted shadow rate is nonnegative).

\textsuperscript{21} We also estimated models where we split the 2005-2016 period in two subsamples of similar length but did not find any interesting result.
least 50 observations for each period but can potentially bias our estimates against finding different responses to monetary policy in the two periods.

In describing the results of the various exercises described above, we will compare the magnitude of the responses of output and prices to monetary policy across sub-periods and also discuss whether the responses are statistically significant in one period and not in another one. Our confidence bands often overlap across periods and formal tests show that the difference of the impulse response functions across sub-periods is rarely statistically significant. This lack of a statistical significance could be due to one of two factors: (1) there is no difference between the two periods; or (2) our short estimation period and the fact that in some cases there is an overlap between the two samples do not give us enough power to precisely estimate the difference across periods. The fact that the difference between impulse responses is not generally statistically significant across sub-periods is an important caveat to our conclusion.

Data

As mentioned above, all of our estimations use monthly data with different starting points and up to 2015M11 for the US, 2016M5 for the UK, Eurozone and Japan. GDP data for the US are from Macroeconomic Advisers, GDP data for the UK are from the National Institute of Economic and Social Research, and we use industrial production for the Eurozone and Japan (from the Eurostat and the Bank of Japan, respectively).

Price index (CPI) and policy rate data for the US and the UK are from the FRED database maintained by Saint Louis Fed. For the Eurozone, instead, we use data from the European Central Bank (we measure the policy rate with the ECB refinancing rate) and for Japan we use Bank of Japan data.

The shadow policy rates are for the US, UK, and Eurozone are from Xia and Wu (2016) and for Japan are from IMF (2015), data on announcements of asset purchases for the US and UK are from Weale and Wieladek (2015) and data on the balance sheet of the ECB and Bank of Japan are from the respective central banks.

Finally, data on equity prices and long term interest rates are from Datastream. For the Eurozone, we measure long-term interest rate using 10-year German bunds and equity prices using a weighted average of the French, German, Italian, and Spanish stock markets\(^2\).

\(^2\) We use the following weights: German DAX 0.3, French CAC 0.3, Italian MIB 0.2, and Spanish IBEX 0.2. These weights reflect the different capitalizations of these 4 stock markets before the global financial crisis. For the UK and the US, we use the FTSE 100 and S&P 500 indexes, respectively. We use the German bunds because our intention is to control for how unconventional policy actions are reflected in the long-term risk-free rate. It may be, however, claimed that the objective of non-standard policy in the Eurozone was to reduce the risk premium in the periphery. Our results are robust to substituting the German Bund rate with the average yield of German Bunds and Italian 10 year government bonds.
4.3. The effect of unconventional policies as proxied by the adjusted shadow policy rate

We look at the impulse response functions of the log of GDP and of the log of the price level to a 1 percentage point decrease in the shadow policy rate at time $t = 1$. The figures that follow display the mean response (continuous line) along with the 68% Bayesian confidence intervals (dotted lines) drawn from 500 Monte-Carlo simulations.

Preliminaries: the interest and exchange rates

In order to understand the results, we first need to understand how the data capture the behavior of the adjusted shadow policy rate. Figure 5 shows its own response to a one-off shock to itself. For each country, the two upper charts correspond to the pre-and post-financial crisis periods, respectively. The results differ across countries. In the US and the UK, conventional cuts in the actual policy interest rate are long lasting: a policy change is sustained over the horizon period of 24 months and the confidence interval is narrow. Unconventional policies are both more variable and less precisely estimated. The adjusted shadow policy rate first continues to be reduced before rising after about one year in the US and 6 months in the UK where it becomes higher than initially, although not significantly so. The pattern is the opposite in the Eurozone, with an immediate reversal in the period before the crisis and a stable pattern after 2008. In Japan, the policy is reversed over both sub-sample periods.

The two lower charts focus on the post-crisis period. Except for Japan, this period is divided in two equally sized subsamples. In all countries, the adjusted shadow policy rate reduction is reversed in the first sub-period, and this is also the case for the second sub-period on Japan and the UK. The reduction is maintained in the US and the Eurozone. Keeping in mind that the adjusted shadow policy rate (as estimated from the yield curve and signaled by policy announcements) includes actual interest rate cuts and the battery of unconventional policies once the actual rate approaches zero, the presence of reversals can be interpreted as a sign of lack of credibility. Under that interpretation, the unconventional policies of the US and the Eurozone during the last sub-period can be seen as more credible than in Japan and the UK. This pattern may have an impact on the output and price effects through asset prices and the confidence effect.

Such an interpretation is partly confirmed by the evolution of the effective exchange rate displayed in Figure 6. A cut in the shadow policy rate is expected to lead to a depreciation, which is partly confirmed. In all countries and periods, the initial depreciation is progressively reversed, however. This is particularly the
case in the post-crisis period. A possible explanation is that many countries adopted similar policies and thus neutralized potential beggar-thy-neighbor effects. This interpretation is supported by the observation that the initial depreciation was stronger in the US and the UK, which moved earlier than the Eurozone and Japan.

The solid lines plot the median responses of the shadow policy rate to a 1 percentage point decrease of the shadow policy rate. The dashed lines plot the 68% Bayesian credible confidence set. The impulse response functions are computed using a Bayesian VAR with the sign restrictions described in the text and the following variables: log of GDP (industrial production for Japan and the eurozone), log CPI, long term interest rate (10 year bond rate), adjusted shadow policy rate, real equity prices, log effective exchange rate.
Output and price effects

Figures 7 reports the response of log GDP to an adjusted shadow policy rate shock — as before, a 1 percentage point cut — and Figure 8 reports the response of log CPI. They are structured as the previous figures.

Except for Japan, the output effects are smaller during the post-crisis period (upper panels) and this is mostly due to decreasing effectiveness in the most recent period (lower panels). The same observation applies to price effects, although they seem to be stronger after 2011 in the UK. For these countries, therefore, these results are consistent with the decreasing effectiveness hypothesis. It is also the case...
that the price effect is smaller after the financial crisis but, except for the Eurozone, there is no indication of decreasing effectiveness over the most recent sub-period, quite the contrary actually. This difference in results may be due to the fact that the European Central Bank was late in implementing unconventional policies. It could also reflect that the ECB shifted its policy in 2012 by adopting a more active stance (“whatever it takes”). Hence, the 2011 breakpoint may not be appropriate for the Eurozone.
One feature, common to all cases except the Eurozone, is that GDP first rises strongly and then declines in the more recent sample, not in the earlier one. Table 1 provides the estimated impact at two different lags: when it is at its maximum and at lag \(12\). This hump-shaped pattern is partly explained by the evolution of both the adjusted shadow rate itself and, to a lesser degree, of the effective exchange rate. There is no such evidence for the Eurozone, possibly due to the late shift of the policy stance, both in bringing the policy rate down to zero — and then below zero — and in adopting explicit QE.

The results are similar to those reported by Weale and Wieladek (2015) with their second identification scheme, which we are using here.

---

**Figure 8.** Price response to the shadow policy rate

The solid lines plot the median responses of log CPI to a 1 percentage point decrease of the shadow policy rate. The dashed lines plot the 68% Bayesian credible confidence set. The impulse response functions are computed using a Bayesian VAR with the sign restrictions described in the text and the following variables: log of GDP (industrial production for Japan and the eurozone), log CPI, long term interest rate (10 year bond rate), adjusted shadow policy rate, real equity prices, log effective exchange rate.

The solid lines plot the median responses of log CPI to a 1 percentage point decrease of the shadow policy rate. The dashed lines plot the 68% Bayesian credible confidence set. The impulse response functions are computed using a Bayesian VAR with the sign restrictions described in the text and the following variables: log of GDP (industrial production for Japan and the eurozone), log CPI, long term interest rate (10 year bond rate), adjusted shadow policy rate, real equity prices, log effective exchange rate.

---

23 The results are similar to those reported by Weale and Wieladek (2015) with their second identification scheme, which we are using here.
Table 1. Increase in GDP and CPI

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>GDP</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Max</td>
<td>0.21*</td>
<td>0.19*</td>
<td>0.07*</td>
<td>0.18*</td>
<td>0.09</td>
</tr>
<tr>
<td>Lag 12</td>
<td>0.18*</td>
<td>0.14*</td>
<td>0.05*</td>
<td>0.12*</td>
<td>0.06</td>
</tr>
<tr>
<td>Lag max</td>
<td>24</td>
<td>6</td>
<td>13</td>
<td>9</td>
<td>5</td>
</tr>
<tr>
<td>CPI</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Max</td>
<td>0.17*</td>
<td>0.23*</td>
<td>0.04*</td>
<td>0.08*</td>
<td>0.08*</td>
</tr>
<tr>
<td>Lag 12</td>
<td>0.16*</td>
<td>0.12*</td>
<td>0.03*</td>
<td>0.08*</td>
<td>0.07*</td>
</tr>
<tr>
<td>Lag max</td>
<td>17</td>
<td>5</td>
<td>5</td>
<td>6</td>
<td>6</td>
</tr>
</tbody>
</table>

* The effect is statistically significant at the Bayesian 68% confidence level.

Notes: Maximum changes in log GDP and log CPI and changes at lag 12 following a 1 pp reduction of the adjusted shadow policy rate.

Taken together, these results suggest a nuanced view of the decreasing effectiveness hypothesis. There is little to support for the view that unconventional policies do not work or that their effectiveness has declined over the most recent period. On the other hand, there is some evidence that the effect is wearing out faster. This is associated with a more rapid reversal of the initial exchange depreciation and of the initial decline of the adjusted policy rate, possibly because markets see a declining room for maneuver, as discussed in Section 2. The weakening of the exchange rate channel may instead be due to the fact of joint expansionary policy stance of the largest central banks.

4.4. The effect of large asset purchases (QE)

We now move to our second indicator of unconventional monetary policy. For the US and the UK, we measure unconventional monetary policy with the announcements of asset purchases by the US Federal Reserve and the Bank of England. For Japan, instead, we use the balance sheet of the central banks because the long period of sporadic unconventional policies in Japan is not easily interpreted. As mentioned above we were unable to estimate the model for the European Central Bank.

Figure 9 presents the impulse response function of the exchange rate to an asset purchase of 1% of GDP. For the US and the UK, the three panels correspond respectively to the whole period of unconventional monetary policies, and to the first and second sub-periods. As before, these sub-periods have a 12 month overlap. For Japan, we present results for the periods before and after the financial crisis in
In the case of the US, the policy announcements lead to an appreciation of the effective exchange rate across all periods. These effects are not statistically significant, except for the early part where they also are somewhat stronger. This apparently surprising result can be seen as consistent with previous results reported in Section 3.4. The quoted literature indicates large effects on a wide range of financial asset prices, interpreted as an indication that QE was instrumental in alleviating market dysfunctions in the wake of the financial crisis. Of interest is the contrast with Figure 6, which indicates that cuts in the adjusted shadow rate led to an exchange rate depreciation. The shadow policy rate measures the policy rate that would be observed in the absence of an effective lower bound given the slope the top row, and to two equally sized and partially overlapping post-crisis sub-periods.

In the case of the US, the policy announcements lead to an appreciation of the effective exchange rate across all periods. These effects are not statistically significant, except for the early part where they also are somewhat stronger. This apparently surprising result can be seen as consistent with previous results reported in Section 3.4. The quoted literature indicates large effects on a wide range of financial asset prices, interpreted as an indication that QE was instrumental in alleviating market dysfunctions in the wake of the financial crisis. Of interest is the contrast with Figure 6, which indicates that cuts in the adjusted shadow rate led to an exchange rate depreciation. The shadow policy rate measures the policy rate that would be observed in the absence of an effective lower bound given the slope
of the yield curve. The adjusted rate incorporates all policy announcements. Therefore it differs from asset purchase announcements as it incorporates how all policy announcements affect the slope of the yield curve possibly through portfolio reallocation and expectations on future policy actions. This suggests that lower policy interest rates alone, no matter what their levels – positive or negative – are, produce standard effects on the exchange rate. However, the expansion of central bank liquidity may lead to an appreciation if its main effect is to reduce financial market disruptions. Since QE can combine both characteristics, it is not surprising perhaps to see in Figure 9 that the overall impact on British and Japanese exchange rate is not statistically different from zero.

The estimated output and price effects of QE — announcements in the case of the US and UK, actual purchases in the case of Japan — are displayed in Figures 10

Figure 10. GDP response to QE announcements

* Response of industrial production to BoJ balance sheet

The solid lines plot the median responses of log GDP to an announcement of a 1 percent of GDP asset purchase (BoJ balance sheet in the case of Japan). The dashed lines plot the 68% Bayesian credible confidence set. The impulse response functions are computed using a Bayesian VAR with the sign restrictions described in the text and the following variables: log of GDP (industrial production for Japan), log CPI, long term interest rate (10 year bond rate), announcements of asset purchases over GDP (BoJ balance sheet over GDP for Japan), real equity prices, log effective exchange rate.
and 11. In the US and UK, the output and price effects are larger in the most recent period, thus apparently running against the decreasing effect hypothesis. Yet, it must be noted that these effects are less precisely estimated and not always significant. The fact that these effects are mostly significant in the early period of unconventional policies confirms the previous observation that the liquidity effect of QE works in the presence of financial market disruptions, not as a monetary policy instrument. In Japan, the effects are not significant, even though the estimates are larger over the most recent period when QE was enacted very explicitly.

We next focus on comparisons with Figures 7 and 8. In line with the reasoning above, we interpret the difference between the adjusted shadow rate and QE as

* Response to BoJ balance sheet

The solid lines plot the median responses of log CPI to an announcement of a 1 percent of GDP asset purchase (BoJ balance sheet in the case of Japan). The dashed lines plot the 68% Bayesian credible confidence set. The impulse response functions are computed using a Bayesian VAR with the sign restrictions described in the text and the following variables: log of GDP (industrial production for Japan), log CPI, long term interest rate (10 year bond rate), announcements of asset purchases over GDP (BoJ balance sheet over GDP for Japan), real equity prices, log effective exchange rate.
representing the effects of all other unconventional policies, chiefly policy interest rate cuts until the effective lower bound is reached and central bank signals afterwards. It is not surprising, therefore that the impulse response functions shown in Figures 10 and 11 display smaller output and price effects in the US and UK, and that the differences are larger for the early period when the policy rates were being cut. In Japan, difference is also larger over the most recent sub-period, an indication that the change is policy in 2013 made a difference.

5. Conclusions

Three main conclusions emerge. First, unconventional policies work. Second they work less well than conventional policies. Third, there is limited evidence that their effectiveness has declined over time. Whatever evidence exists, it may be a confirmation of earlier studies that find these policies to be helpful in the presence of financial market disruptions, less so when the markets have recovered.

Table 2. Evidence on decreasing effectiveness

<table>
<thead>
<tr>
<th>Measure of unconventional policy</th>
<th>QE Announcements</th>
</tr>
</thead>
<tbody>
<tr>
<td>Adjusted Shadow Policy Rate</td>
<td></td>
</tr>
<tr>
<td>QE Announcements</td>
<td></td>
</tr>
<tr>
<td>Econometric approach</td>
<td></td>
</tr>
<tr>
<td>Local projections</td>
<td></td>
</tr>
<tr>
<td>BVAR</td>
<td></td>
</tr>
<tr>
<td>Local projections</td>
<td></td>
</tr>
<tr>
<td>BVAR</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Country</th>
<th>GDP-Adjusted Shadow Policy Rate</th>
<th>CPI-Adjusted Shadow Policy Rate</th>
<th>GDP-QE Announcements</th>
<th>CPI-QE Announcements</th>
</tr>
</thead>
<tbody>
<tr>
<td>US</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
</tr>
<tr>
<td>UK</td>
<td>yes</td>
<td>yes</td>
<td>no</td>
<td>no</td>
</tr>
<tr>
<td>EZ</td>
<td>yes</td>
<td>mixed</td>
<td>no</td>
<td>n/a</td>
</tr>
<tr>
<td>Japan</td>
<td>yes</td>
<td>no</td>
<td>n/a</td>
<td>no</td>
</tr>
</tbody>
</table>

a Answer to the question: is there evidence that the effects have declined over the more recent period of unconventional monetary policies? Answer based on visual inspection of the results.

b Based on the precision of the estimate and the profile of the response.

At the methodological level, we find that the results are sensitive to the way unconventional policies are measured, to the criterion (GDP or CPI), and to the econometric approach. Table 2 summarizes our results, whether we measure policy with the adjusted shadow policy rate with QE announcements or actions, and whether we use BVARs or local projection as indicated in the Appendix. When we use the adjusted shadow policy rate our evidence is broadly consistent with decreasing effectiveness both when we use local projections and Bayesian VARs.
However, when we use QE announcements, we rarely find evidence that is consistent with the decreasing effectiveness hypothesis (with the exception of a price effect for the US and the UK which, however, does not appear to be robust across estimation methodologies).

A yes/no assessment may be too rough to do justice to the overall profile of impulse response functions, however. We need to look at the whole horizon. A careful analysis of these profiles suggest that during the more recent period the impact of unconventional policies has not declined, or has even increased in several instances, but that the effect wears out fast.

We also need to keep in mind that the impulse response functions look at the effects of a one-off effect of a shock to the chosen measure of unconventional policy. It matters a great deal how this measure evolves over time. A key argument in favor of the decreasing effectiveness hypothesis is that the policy space is shrinking when the policy interest rate reaches its effective lower bound and when the balance sheet has already been considerably enlarged. We find some evidence that the shadow policy rate, which includes market expectations, tends to reverse faster over the most recent period. This pattern is mirrored in the evolution of the exchange rate, a potentially important channel of monetary policy.

Many caveats apply. First, it may be too early to assess the decreasing effectiveness hypothesis. Facing short samples, we have been led to use partially overlapping sub-periods, which works against finding supporting evidence. In addition, the limited sample size leads to large standard errors. Furthermore, macroeconomic phenomena are low frequency events; using monthly data increases the sample size but does not erase the fact that we do not have much perspective. Second, we do not test the hypothesis formally. We simply compare visually the economic responses to policy actions. Third, in line with much of the literature, we look at the responses of GDP and CPI levels, not at their growth rates. This is a natural next step. Fourth, some other variables may need to be included. We have experimented with a measure of financial market volatility (the VIX index) and the results did not change significantly. Finally, the sensitivity of results to the measure of policy and to the econometric approach calls for much more experimentation.

Appendix is available at: www.cbr.ru/money-and-finance

6. References


Kiley, M. (2015). Low Inflation in the United States: A Summary of Recent Research. FEDS Notes, Federal Reserve Board,


Reflections on Central Banking, Protectionism and Globalization

Dr. Jacob A. Frenkel,
Chairman, JPMorgan Chase International,
Chairman of the Board of Trustees, G30,
Former Governor, Bank of Israel

This year marks the 10th anniversary of the eruption of the global financial crisis of 2008-2009. It provides a good opportunity for reflections on the recent evolution of Central Banking Theory and Practice, as well as on the challenges for globalization, on the emergence of protectionism and on the growth of populism. This brief paper offers few reflections on each.

1. Central Banking

1.1. Basic Principles

Over the past decades central banking theory and practice evolved. The economic system, the nature of markets and the structure of economic policies, have changed and have brought about fundamental changes in central bank structures and policies. The mandate of central banks has been modified, the policy instruments have evolved, and the degrees of accountability and transparency have risen. These developments resulted in the current state of affairs in which central banks have become “the only game in town”. While this “distinction” is somewhat flattering, it is neither desirable nor sustainable.

A superficial observation of central bank policies today in comparison with its policies a decade or two ago, may yield the conclusion that at present we are in a fundamentally different universe than we were in the past. Based on this superficial observation one may reach the conclusion that the “new world” requires a fundamentally new approach to central banking, that old concepts are obsolete, that we can throw the old textbooks away and replace them by the new ones. I strongly believe that this would be the wrong conclusion. Even though the economic system has evolved, the basic premises and principles of central banking and the benefits from globalization and open markets, which have been established on the basis of many decades of experience, remain robust, useful, and relevant.

Not withstanding the fact that the economic circumstances and the legal frameworks differ across countries and have changed over time, the general principles of central banking remain valid and intact. They include a specification of the mandate of the central bank, with a typical focus on the attainment of price

---

1 This article represents the report the author made at the Gaidar Forum in Moscow in January 2018.
stability and financial stability, a medium term perspective, independence from short-term political pressures, autonomy in executing the mandate, responsibility for the smooth operation of the payments system, accountability and, in extreme cases, playing the role of the “lender of last resort”. Modern central banking requires that the formal authority of the central bank is specified and protected by an explicit central bank law. These characteristics have evolved through time and by now they have become universal. Specifically, the recognition that countries with strong central banks exhibit on average a better economic performance than those with weak central banks, have resulted in a convergence of views about the role of central banks in a modern society.

The development of capital markets and the growing integration among such markets have introduced new characteristics to central banking. They resulted in an increased interdependence among economies, thereby increasing attention to international policy coordination and cooperation. In contrast with the relatively slow adjustments in the markets for goods and labor, capital markets adjust very quickly. Changes in capital markets reflect not only current policy actions, but also expected policy actions. As a result, the roles of reputation, consistency, and credibility have become extremely important. The growing importance of capital markets and the expanded roles of banks and the financial sector, have contributed immensely to the performance of the economy but, at the same time, have also increased its degree of vulnerability. Thereby, the functions of supervision and regulation of the financial industry have become a critical factor governing the stability of the economic system.

1.2. The Global Financial Crisis

The Asian economic crisis and the Russian default which occurred during 1997–1998, served as a wake-up call and illustrated vividly the importance of capital markets and the critical role that financial stability (and instability) play. However, developments during the subsequent decade have revealed that the lessons from the experience of the late 1990s (especially following the Asian crisis and the Russian default), have not been fully learned. Many economies (especially among the industrial countries) have neglected to apply appropriate risk management to their macroeconomic and financial conduct. An exceedingly rapid rise in debt and leverage and a mispricing of risk have increased the vulnerability of the economic system and have ultimately resulted in the financial crisis of 2008–2009. In 2009 the crisis brought the world economy to a halt, global growth vanished and the level of output in the industrial countries shrunk by 3.4 percent (see Figures 1-2).

The results of this crisis, which erupted as a “perfect storm”, are still with us. They have resulted from poor risk management, the failure of bank supervision, the lack of enforcement of the regulatory framework, the deficiencies of the regulatory system itself, the vulnerability that arises from fiscal laxity and excess
leverage, from the distortions that arise from lack of flexibility of the economic system, and the like.

1.3. Unconventional Policies

In order to avoid a complete meltdown following the outbreak of the 2008–2009 financial crisis, governments in most countries have increased dramatically their budget deficits (fueling further the excessive levels of debt) and central banks all over the world have lowered interest rates to unprecedented low levels (see Figure 3).
Thereby, conventional monetary policies reached their limit. Since the economic recovery has not yet been in sight, central banks needed to resort to unconventional measures. The size of the balance sheets of the central banks has expanded at a very rapid rate and the composition of assets in the central bank’s balance sheet has also changed dramatically. Instead of holding safe and highly liquid short-term government treasury bills, most of the central banks have accumulated a wide range of lower-quality and less liquid securities such as mortgage-based securities and the like. Clearly, unconventional challenges needed to be met by unconventional measures. Since the various economies suffered from a “balance sheet shock” it was understood that the period of adjustment would not be short. However, in retrospect, it is fair to say that very few policy makers and market participants expected the period, during which central banks would have had to resort to unconventional measures, to be so long. The lengthy period during which the unconventional policies have been in place has created the risk that these unconventional policies, which were initially believed to be just a temporary detour from the conventional policy regime, might become the new paradigm, and thereby constitute the new set of conventional policies.

There is a complete consensus that in the United States, following the outbreak of the crisis, the initial monetary policy response by the Federal Reserve has been both essential and successful. Quantitative easing (QE) was appropriate and effective. With the passage of time however, when the first round of quantitative easing (QE1) did not bring about the full recovery, especially due to the fact that a large share of the burden of policy adjustment fell on the shoulders of the Federal Reserve, another round of QE was adopted — QE2. Monetary policy became overburdened and when the desired recovery did not yet fully materialize, the next round of expansion took place — QE3. In retrospect, each round was productive,
but its effectiveness diminished—the law of ‘diminishing returns’ took hold. Across the Atlantic Ocean, in Europe, the European Central Bank (ECB) also engaged in its own version of QE. The challenges in Europe were more severe, since there are many diverse governments, diverse economic conditions, labor markets that are less flexible, and banking union that is still not in place. A similar challenge characterized the situation across the Pacific Ocean in Japan. The economic strategy in Japan (Abenomics), has been based on the “three arrows” – monetary policy, fiscal policy, and structural policy — but most of the burden fell on the shoulders of the Bank of Japan; only one of the three arrows has been fully operational. In most of the industrial countries of the world, monetary policy has become overburdened and has become the “only game in town”. These developments are illustrated in Figure 4.

Figure 4. Total Assets of Key Central Banks

<table>
<thead>
<tr>
<th>Current Assets, Billions of $</th>
<th>Current Assets, % of GDP</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total</td>
<td>14,306</td>
</tr>
<tr>
<td>of which</td>
<td></td>
</tr>
<tr>
<td>Fed</td>
<td>4,437</td>
</tr>
<tr>
<td>ECB</td>
<td>5,247</td>
</tr>
<tr>
<td>BoJ</td>
<td>4,622</td>
</tr>
</tbody>
</table>

Source: Bloomberg; last observation: Fed: Dec 6, 2017; ECB: Dec 8, 2017; BoJ: Dec 10, 2017

1.4. Normalization

Everyone recognized that it would be desirable to restore “normalization” and to bring about the conditions that would enable higher central bank interest rates. Much of the debate however, has been whether the economies are strong enough to warrant the start of normalization. Since the macroeconomic situation differs
significantly between the US, Europe and Japan, it is clear that normalization should not be initiated simultaneously. In this regard, the US has best positioned to start the journey of normalization. Economic growth clearly improved (see Figure 5), unemployment has declined dramatically (see Figure 6), the duration of unemployment has also declined and inflation (net of energy prices) is also on the rise, approaching the two-percent target. While the US normalizes the course of monetary policy by raising short term interest rates and subsequently also reducing gradually the size of the FED's balance sheet, Europe and Japan are not yet ready to normalize their unconventional monetary policies. As a result, there is likely to be a period of time during which monetary policies in the US, Europe and Japan

**Figure 5. Real GDP Growth, Select Countries**

![Graph showing real GDP growth for select countries](image)


**Figure 6. Unemployment Rate: US and Euro Area**

![Graph showing unemployment rates for US and Euro area](image)

*Source: Eurostat and BLS; last observation for Euro area October 2017, For US, November 2017*
would differ from each other. These diverging paths of monetary policies are likely to be reflected in corresponding changes in exchange rates whereby the US dollar appreciates relative to the other currencies. Such exchange rate changes are desirable. They reflect a healthy component of the international adjustment mechanism, and should not be resisted. Furthermore, the appreciation of the US dollar relative to the Euro should also shorten the period of time that the ECB requires before it also initiates normalization.

Much of the discussion about normalizing the course of monetary policy, have focused on the negative consequences of higher interest rates. An excessive focus on the cost of normalization rather than its benefits increases the risk that such normalization will be initiated too late. In order to restore balance to the discussion regarding normalization of monetary policy, it is important to recognize that maintaining an exceedingly low interest rate and delaying the process of normalization is also costly. It entails economic cost that need to be taken into account. In what follows, I list some elements of such cost:

1) The low-interest rates induce investors to seek alternative ways to generate returns. By chasing after yield, investors end up assuming higher risk, which might be mispriced.
2) The low-interest rates bring about an inflation of stock prices and may generate a financial bubble.
3) Corporations divert their efforts towards stock buy-backs instead of allocating their resources to investment in plant and equipment.
4) The inflated financial markets create a disconnect between the real and the financial sectors of the economy.
5) The low-interest rates encourage excessive leverage and thereby increase the vulnerability of the financial system.
6) The low-interest rates and the flat yield curve result in negative consequences for the financial-services industry. This includes: banks, insurance companies, and pension systems.
7) The transmission of the effects of monetary policy through the economy operates through the financial system; a weakened financial system reduces thereby, the effectiveness of monetary policy.
8) The low-interest rates provide an artificial stimulus to interest-sensitive sectors, such as housing. Since this sector is typically a low-productivity sector, it results in an overall reduction in the productivity of the economy.
9) The excessive reliance on monetary policy enables governments to postpone the necessary fiscal and structural measures. The postponement of these measures reduces the flexibility of the economy, and thereby reduces productivity and growth.

This partial list of the negative consequences of excessively low rates of interest, suggest that a delayed normalization is costly and that one should always bal-
ance these costs against the cost of normalization. Obviously, most central banks are fully aware of these considerations, but on balance it seems that the public and the political debate of these issues put a greater weight on the arguments highlighting the cost of normalization than on the arguments highlighting the benefits from normalization. If such a bias exists, it is likely that when the process of normalization does take place, it will be implemented too late and might proceed too slowly. An undue delayed normalization may entail significant cost.

1.5. Inflation Targets

The risk that normalization might be excessively delayed is enhanced by the fact that recently all the major central banks put extraordinary emphasize on achieving their two-percent inflation target. This is true of the Federal Reserve, the Bank of England, the ECB, the Bank of Japan and others. There are many reasons (some of them not under the control of the central bank) as to why, in spite of the extraordinary efforts, the rates of inflation in the industrial world have been very low and have gotten ‘stuck’ below the two-percent target. If, ‘no matter what’, interest rates are to be maintained at exceedingly low levels, for as long as the rate of inflation is below two-percent, then there is a significant possibility that financial stability might be at risk. The BIS has forcefully voiced this concern and I believe that it would be prudent to pay a significant attention to the concern regarding financial stability.

Inflation targets have proven to be highly successful as a strategy for disinflation in countries with high inflation. In such cases the path of the inflation targets is the compass, which provides credibility and transparency to the multi-year policy goal of achieving a gradual reduction of inflation towards medium-term price stability. For such countries it is less critical if a range of which the mid point is 2.0 percent, or 1.5 percent, or 2.5 percent defines the long-term goal of ‘price stability’. This is not, however, the policy challenge faced today by the United States and Europe (as well as by many other industrial countries). These countries have already achieved medium term price stability; their aim is to preserve it by keeping their inflation rate within the price-stability target range. For these countries it matters a great deal if the mid point of its target range is 2.0 or 1.5 or 2.5 percent.

Setting the inflation target at the rigid level of two percent (or slightly below two, in the case of the ECB) represents an interesting convergence of views among the major central banks. The various statements of the major central banks, which are intended to provide the markets with forward guidance, illustrate this convergence. Below are samples of recent statements by the Federal Reserve, the Bank of England, the ECB and the Bank of Japan. Each one of these statements refers to the two-percent inflation target.

**Forward Guidance by the Federal Reserve**: “Inflation on a 12-month basis is expected to remain somewhat below 2 percent in the near term but to stabilize
around the Committee’s 2 percent objective over the medium term. Near-term risks to the economic outlook appear roughly balanced, but the Committee is monitoring inflation developments closely. In determining the timing and size of future adjustments to the target range for the federal funds rate, the Committee will assess realized and expected economic conditions relative to its objectives of maximum employment and 2 percent inflation…” (FOMC statement, December 13, 2017).

**Forward Guidance by the Bank of England:** “The Bank of England’s Monetary Policy Committee (MPC) sets monetary policy to meet the 2% inflation target, and in a way that helps to sustain growth and employment <…> The MPC still expects inflation to peak above 3.0% in October, as the past depreciation of sterling and recent increases in energy prices continue to pass through to consumer prices… The steady erosion of slack has reduced the degree to which it is appropriate for the MPC to accommodate an extended period of inflation above the target” (Minutes of the MPC meeting November 2, 2017).

**Forward Guidance by the ECB:** “<…> we confirm that from January 2018 we intend to continue to make net asset purchases under the asset purchase programme (APP), at a monthly pace of €30 billion, until the end of September 2018, or beyond, if necessary, and in any case until the Governing Council sees a sustained adjustment in the path of inflation consistent with its inflation aim. <…> Our monetary policy decisions have preserved the very favorable financing conditions that are still needed for a sustained return of inflation rates towards levels that are below, but close to, 2%...” (ECB President Mario Draghi, Introductory statement, Dec. 14, 2017).

**Forward Guidance by the Bank of Japan:** “The Bank will continue with «Quantitative and Qualitative Monetary Easing (QQE) with Yield Curve Control,» aiming to achieve the price stability target of 2 percent, as long as it is necessary for maintaining that target in a stable manner. It will continue expanding the monetary base until the year-on-year rate of increase in the observed CPI (all items less fresh food) exceeds 2 percent and stays above the target in a stable manner...” (Statement of Monetary Policy, September 21, 2017).

It is relevant to note, however, that the choice of the two percent target rate is somewhat arbitrary and, although the two-percent target rate has become the consensus target for most of the industrial-countries’ central banks, its theoretical foundations can be questioned especially if the two-percent target is set as a rigid quantitative target rather than as the mid-point of a wider target range. In this context it should be emphasized that the “whatever it takes” dictum stated by Mario Draghi (the ECB’s President), is not intended to apply to the medium-term inflation target. Specifically, on July 26, 2012, in his breakthrough speech, Mario Draghi stated that: “Within our mandate, the ECB is ready to do whatever it takes to preserve the euro. And believe me, it will be enough”. Thus, this justifiably famous
statement addresses the challenge of preserving the euro (within the ECB’s mandate) rather than sticking blindly with the inflation target.

Recently, there have also been proposals that the Federal Reserve should raise its inflation target above two-percent. The logic of the proposal rests on the fact that real interest rates have declined to very low levels and that with the existing two-percent inflation targets, the resultant levels of nominal rates (the sum of the real rate of interest and expected inflation) are too low. Hence, so the argument goes, a higher level of inflation targets would permit higher nominal rates of interest even though the real rates of interest are at historically low levels.

This policy recommendation, however, is subject to several reservations. Before considering to raise the inflation target, we should have a better understanding as to the reasons for the decline in the real rate of interest. Specifically, among the factors responsible for the low real rates are demographic factors, uncertainty, a lower level of productivity, and the like. Raising the inflation target would imply that the exceedingly low level of the real rate of interest is here to stay; the higher inflation target in fact would validate the low real rate. I believe that this verdict is premature. Some sources of the uncertainty are policy induced and should be removed; furthermore, the low level of productivity should also not be taken as a given, and policy efforts should be directed towards raising productivity. These efforts should typically be affected through the implementation of structural policies that remove distortions and increase the flexibility of the economic system, through improved infrastructure, education, and tax incentives that promote an innovative culture.

In addition to these points of principle, there is a practical issue. Most of the major central banks such as the Fed, the ECB, the Bank of Japan, and the Bank of England, have been struggling for a while to raise their corresponding inflation rates from a near zero levels towards their inflation targets of about two-percent per year. What good would it make to raise the inflation target above its current level of two-percent when the level of actual inflation is stuck below two-percent?

Furthermore, there are still many countries especially in emerging economies, which suffer from high inflation. For such countries, the challenge is to lower inflation towards its target. The credibility of the inflation targeting strategy, which is the main strategy adopted by this group of countries, would be seriously eroded by changing the inflation target. If the industrial countries were to raise the inflation target as proposed, it would damage the efforts of the emerging markets who are still struggling to adhere to their inflation targeting strategy in an effort to achieve price stability.

1.6. Conclusions on Central Banking

I conclude this section on central banking and monetary policy by noting that in spite of the changing circumstances and the new challenges faced by policy makers, it is important to maintain the fundamental principles of central banking.
The medium term focus on stability, both price stability and financial stability, are and should remain the key objectives of the central bank. In order to achieve these objectives, the central bank must be granted operational independence (autonomy). Furthermore, a strong financial system is fundamental for the maintenance of stability, and such stability is the precondition for the attainment of sustainable growth. In fact, the best way that the central bank can contribute to sustainable growth is through delivering price stability and financial stability. A strong banking system is key for such stability as well as for the effective transmission of monetary policy. High capital ratios, low leverage, and high liquidity, characterize a strong banking system. In order to discharge the central bank's responsibilities, it must have the authority and the tools to bring about a strong financial system. Hence, in most cases it would be desirable that the responsibility for bank supervision rests in the hands of the central bank. Furthermore, the central bank should also be granted the authority and the instruments to secure the smooth operation of the payment system. This is especially the case during financial turmoil.

These objectives represent a very heavy list of tasks and it should be clear that the central banks could not do it alone. There should be a full cooperation from the other branches of government that will secure fiscal responsibility, open trade, structural policies, and the like. Without such cooperation, monetary policy will be overburdened, the central bank will be the only game in town, and economic performance will be below potential.

Crisis are a typical phenomenon of the modern economic and financial systems. The role of policy is to design effective mechanisms for crisis prevention, crisis management, and crisis resolution. Occasionally, the appropriate policy response to a crisis entails a departure from typical norms of policy. These are the occasions that the credibility of the policy strategy plays a critical role in the restoration of stability. For such credibility to be maintained it is important that such departures from the norm be viewed as a temporary detour rather than a permanently new paradigm. The medium term perspective of the conduct of policy provides the compass that ensures that the long-term objectives are achieved.

2. Globalization, China, Trade and Protectionism

The Global Financial Crisis has also brought about skepticism about the benefits from globalization, a revival of calls for protectionism and a growing prevalence of populism in the economic and the political spheres. These sentiments have also been stimulated by, and reflected in, the relations with, and the attitude towards, China. The growing weight of China in the global economy illustrates the benefits from trade, the dangers from protectionism, and the role of globalization. Over the past 20 years China’s economic growth has been spectacular, averaging about 10 percent per year. More recently, its annual growth rate has moderated to a more sustainable rate of about 6.5 percent. This exceptional performance of the Chinese
China's opening to international trade has been a critical factor underlying its exceptional economic performance. It also brought about the rapid alleviation of poverty of tens of millions of Chinese citizens.

In addition to its dominant position in world output, China has also become an indispensible player in world trade, thereby reflecting the shift in the economic center of gravity towards Asia. During 2017 the volume of trade between China and the rest of Asia has been about twice as large the corresponding volume of trade between China and the US plus Europe taken together. Furthermore, China has become the most important trading partner of most of the major industrial countries in the world. For example, during 2017, as illustrated in Figures 8a and 8b, approximately one-quarter of US and European exports are shipped to China.

These facts imply that the degree of interdependence within the world economy is very high. As a result, a protectionist induced disruption to the interrelation between China and the rest of the world would be extremely costly.

An illustration of the strong relationship between international trade and economic growth is provided by reference to Figure 9, which depicts the evolution of the volume of world trade during the past 30 years. It shows that except for 2009 the volume of world trade has expanded in each and every year. During 2009 — the worse year of the global financial crisis — the volume of world trade shrunk by 10.5 percent. As shown in Figures 1 and 2, 2009 was also the year characterized by
the worse economic performance of the world economy: global GDP was stagnant while, at the same time, the level of GDP of the industrial countries shrunk by 3.4 percent.

Protectionism would reverse the great benefits that have been brought about through international trade; the damage to the interdependent world economy would be immense.

One of the main lessons from the recent global financial crisis is the importance of ensuring a robust financial system with a special focus on the health of the banking system. This is one of the major challenges that still remains and that must be addressed with great urgency by the Chinese authorities. In this regard it is also
important that the activities of the Chinese shadow banking be illuminated, taken out of the shadow, and become fully transparent. Furthermore, it is important that the non-performing loans that are still prevalent on the balance sheets of the Chinese financial institutions are handled properly.

Notwithstanding these considerations, however, I am still optimistic about the role that China plays and can play in the global economic system. The recent initiative of China in launching the Asian Infrastructure Investment Bank (AIIB) is a positive and encouraging development. It is encouraging that most of the major countries have joined this new multilateral bank, and one can only hope that the United States and Japan will also join this initiative before too long. Generally speaking, China should be granted its appropriate place in the formal international architecture and its quota shares in the multilateral organizations should be commensurate with its economic size.

In view of these the developments the recent popularity of protectionism is worrisome. It needs to be confronted with massive educational efforts highlighting the key issues and the ways to address them. In fact, most people who support protectionism would clearly prefer not to give up the gains from trade. They justify their protectionist stance by noting that, in many cases, opening the economy to international trade may inflict hardship on some segments of the population. In this regard it is important to emphasize that the appropriate way to address this challenge is through fiscal measures. Such measures include trade-adjustment assistance, retraining programs, and appropriately budgeted safety nets designed to support the weakest segments of society. Generally, hardships that are associated with the opening to international trade do not arise directly from trade but rather from the failure of governments to enact the appropriate fiscal measures. It will be tragic to forego the gains from trade, just because government are unable or unwilling to implement the appropriate fiscal policy measures that simultaneously secure the gains from trade, while at the same time, reduce the hardship that may accompany trade.

3. Populism

In addition to the dangerous popularity of protectionism, the global financial crisis has also resulted in the emergence of populist sentiments that are manifested in the economic sphere as well as in the political sphere. In the economic sphere, trend changes in the distribution of income have been a source for the growth of populism. Since the 1980s, in the industrial countries, the relative share of labor in GDP has been on the decline. The decline of the share of labor has been correlated with the rise in income inequality. Those at the higher end of the income distribution are also the owners of capital, so when the relative share of capital in GDP rises, it is associated with a rise in income inequality — but that, again, has little to do with trade or with technological advance. This implies that we need to focus on the
question, why the relative share of labor has declined, rather than adopt a populist stance.

In the political sphere one of the reasons for the growth of populism has been the dramatic decline in the levels of trust in politicians, in the political system, and in the media. An increasing number of people feel that the politicians and the political system have failed to serve them well. A growing number of people feel that they have been left behind and, for the first time since many years, a growing number of families doubt that their children will be better off than they are. Furthermore, people have lost their trust in what they read or listen to in the media. Everyone has his own “truth”. The truth that people believe in is typically the one that is shared by the group of people that they are associated with and, typically, people associate themselves with others who hold similar beliefs and values. This is the way in which groups are formed in the social media world. Individuals who belong to a WhatsApp group or WeChat group or other groups within Facebook typically share the same beliefs and values. By talking to each other they define their collective “truth”. At the very same time, the same happens in other groups who define their collective truth. Under these circumstances, it is typical that political election results are not easy to forecast, since every group believes that, everyone thinks like them because these are the views that they encounter within their own group. It is no wonder that most of the recent forecasting of political election results failed dramatically.

In addition to the diminished trust in politicians and the media, there is also a diminished trust in experts. One of the unfortunate consequences of the global financial crisis has been the perception that “experts have failed”. As a result, being an expert is not viewed as an advantage but rather as a disadvantage. Likewise, having experienced persons in leadership positions is viewed as a liability rather than an asset. This has resulted in a growing popularity and prominence of inexperience political candidates and in the elections of “outsiders”. In the complex world of today, this phenomenon of the rise to prominence of outsiders and inexperienced leaders should be viewed with concern, especially as it has been accompanied by populism and anti trade sentiments of protectionism.

4. Conclusion

The recent global financial crisis has resulted in a new creative set of economic policies. Both, fiscal and monetary policies have departed from their conventional course. Faced by a colossal crisis, budget deficits were raised to levels that normally would have been viewed as excessive, especially in view of the already high levels of leverage, which prevailed in the major industrialized countries. At the same time monetary policies responded to the crisis by adopting an extremely expansionary stance which, under normal circumstances, would have been regarded as excessive. Central banks’ interest rates were reduced to zero (and below)
whereas real interest rates became negative. The justification for the unconventional course of the policy response was based on the argument that “unusual challenges must be met by unusual policies”. Furthermore, the implicit assumption was that the departure from the norms of macroeconomic policies would be temporary, and that following a relatively short detour into the unconventional territory, the course of macroeconomic policies would return to its conventional path. In practice this detour has lasted longer than expected. Now that the process of normalization has started in the United States and is likely to be followed (albeit in some delay) in Europe, it would be important that policy makers emphasize that the unconventional set of economic policies were just a detour from the longstanding convention rather than representing a new paradigm. This experience should be recorded in history as reflecting a period during which new and important policy chapters were drafted. These chapters should be added to the corpus of knowledge of macroeconomic theory and policy. The new chapters contain important lessons that should definitely not be forgotten once the crisis is over. They should be added to, but not replace, the old textbooks. These old textbooks should definitely not be thrown away as they summarize lessons from the experience from many decades past.

The occasion of the 10th anniversary of the Global Financial Crisis provides an excellent opportunity to look back, reflect, and appreciate the critical role that central banks can play in an economy that undergoes fundamental changes. It also provides an opportunity to appreciate the benefits from globalization, from keeping open markets, and from maintaining free international trade. At the same time, this reflection should lead us to recognize the dangers and damages that result from inward looking protectionism and from populism. Let’s remember: the fundamental principles that over the past several decades have provided the basis for the conventional set of macroeconomic policies are still highly relevant and robust; they stood the test of time and space, and are likely to serve us well also in the future.
Reflections on Reflections from the Russia’s Mirror

Commentary to Dr. Jacob Frenkel’s Article¹

by Alexander Morozov, Bank of Russia

Dr. Frenkel’s well-thought-out reflections on central banking, globalization and populism concern the global economy, developed countries in particular. However, his thoughts have important implications for Russia as well. They are summarized in this short note.

The perception of central banks as “the only game in town” common in developed countries is quite popular in Russia as well, though arguably less than it had been a couple of years ago. At the same time, the reasons for such perception in Russia are somewhat different. First, historically, the central bank was considered as a lender of first, not last resort to the real sector along with the financial sector. Second, public opinion in Russia does not clearly differentiate between monetary and fiscal policies, and between the Central bank and commercial banks.

Third, like in other major economies, monetary policy in Russia is often seen as an obvious and simple way to quickly support economic growth and institutions, both financial and non-financial, without macroeconomic costs or costs to taxpayers. In contrast, the use of fiscal and particularly structural policies has always been much more challenging, including from the political economy perspective.

Dr. Frenkel summarizes his views on central banking as follows: “The best way that the central bank can contribute to sustainable growth is through delivering price stability and financial stability. A strong banking system is key for such stability as well as for the effective transmission of monetary policy”. Nothing better can be said. The Bank of Russia is fully up-to-date in its mandate and activities in regards to the newly emerged characteristics and functions of central banking. It pays significant attention to clear communication with the financial industry and markets, consistency in policy actions and regulatory principles. In bearing responsibility for the soundness of the financial industry underpinning the stability of the national economy, the Bank of Russia has been cleaning the banking system and other segments of the financial sector.

Dr. Frenkel offers important comments on risks for financial stability stemming from the delay in policy normalization in developed countries. These comments are also relevant for Russia, but from a different angle. Although policy normalization in Russia implies policy easing (i.e. the reduction of policy rates) rather than policy

¹ See p. 108.
tightening, the risks for financial stability may arise from too speedy policy normalization as a possible overreaction to low headline inflation observed in the last several months. Why? The so-called “missing inflation” (below-the-target inflation accompanying strong real business cycle indicators) appears to be a global phenomenon, embracing Russia among other countries. Thus, attempts to offset structurally low price growth in some goods by pushing up price growth in other goods and services with the help of monetary policy could put medium-term financial and price stability at risk.

Dr. Frenkel argues for inflation target to be interpreted as the mid-point of a wider target range in developed countries. He argues even stronger in favour of a vague and less time-bound interpretation of inflation targets in emerging markets, especially those that have to disinflate. Russia adopted its inflation targeting policy as a strategy for disinflation, so his arguments fully apply to Russia, as does the argument that in the medium term there are more upside risks than downside ones because of unanchored inflation expectations.

The Global Financial Crisis, as Dr. Frenkel explained, had challenged both economic theory and experts who “failed to predict it”, following traditional textbooks. The resulting mistrust in traditional economic theory and experts echoed in Russia in two ways. First, unconventional monetary policies that major central banks had had to run were often interpreted as a “new normal” in central banking globally, regardless economic stance. It was taken completely out of the context of the key central bank mandates. The Bank of Russia had to argue and prove that conventional monetary policy was alive and right medication for the Russian economy. Second, the mistrust in experts gave support to pseudoexperts with little knowledge of economic theory, but with strong ambitions to invent and propagate their own ideas, which were often derived from poor empirical data and primitive economic analysis. Fortunately, such views have not been translated into policy decisions, yet they got entrenched among the general public. It will take a massive and prolong educational campaign to get economic basics right.

The normalization of monetary policy in developed countries does not in itself carry high risks provided it occurs gradually and does not affect the global economy’s growth rate, does not slow it and does not bring any surprises to financial markets. If, however, inflation starts accelerating, requiring faster policy tightening, this may hit assets hard, including in emerging markets, via capital outflows. In previous years, monetary policy tightening in developed countries or announcement about the launch of tightening tended to cause a rise in volatility in global markets. This is, for instance, what happened in 2013 after the Fed’s announcement that it would start winding down quantitative easing. The conventional wisdom is that now monetary policy tightening in the US will have a milder impact on emerging economies today because, first, it is happening against the background of sustainable economic growth and, second, emerging markets themselves now have more shock
absorbers to alleviate these risks. Namely, these include macroeconomic stability, floating exchange rate and reserves.

Compared with other emerging markets, Russia looks better protected for two reasons. First, financial deleveraging has occurred: external debt dwindled from $732.8 bln in mid-2014 to $529.1 bln at the start of 2018, or by 28%. It dwindled even more dramatically in the banking sector – by half, from $208.9 bln to $104.5 bln, over the same period. The need to refinance debt, although relevant, is not problematic. Second, thanks to the interest rate difference, Russia can afford to pursue a policy opposite to that of the Fed by cutting interest rates, whereas the US is hiking them. This will be feasible until the Bank of Russia’s key rate reaches a neutral level of 6–7%. That the rate is higher now makes Russian assets more attractive than those of other countries, damping the effect of the Fed’s policy normalization. In addition, over the recent years, the Bank of Russia has developed an array of instruments to support financial stability. These include the provision of foreign exchange liquidity via repos and loans where there is a dollar shortage. This mitigates the perception of potential negative effects by market participants.

Finally, a few words on protectionism induced by changes in international trade and technological innovations. Dr. Frenkel makes the point that the use of fiscal measures (trade-adjustment assistance, retraining programs, social safety net adjustment, etc.) is the appropriate way to address this challenge rather than protectionism. This goes against common beliefs of policy makers in most countries. Thus, it is important to fully understand the point and act accordingly, with appropriate policy tools. In Russia decreased energy prices prompted economic rebalancing, fostering export diversification and integration into global value chains. Therefore, it is in Russia’s interests that global markets remain open.
XXVII International Financial Congress

THE FINANCIAL SYSTEM: STABILITY FOR GROWTH

ORGANIZER: BANK OF RUSSIA

ST. PETERSBURG
JUNE 6–8
2018

A detailed programme is available on our official website, ibcongress.com, where you can also register as a Congress participant.
CONTENTS

Neutral Interest Rates in CEEMEA — Moving in Tandem with Global Factors
Clemens Grafe, Sara Grut, Lorenzo Rigon

Impact of Banking Supervision on Banking System Structure: Conclusion from Agent-Based Modelling
Alexey Ponomarenko, Andrey Sinyakov

Hidden Negative Capital in Banking Before and After the Senior Management Change at the Bank of Russia
Mikhail Mamonov

The Folk Theorem of Decreasing Effectiveness of Monetary Policy: What Do the Data Say?
Ugo Panizza, Charles Wyplosz

Reflections on Central Banking, Protectionism and Globalization
Jacob A. Frenkel

Reflections on Reflections from the Russia’s Mirror
(Commentary to Dr. Jacob Frenkel’s Article)
Alexander Morozov