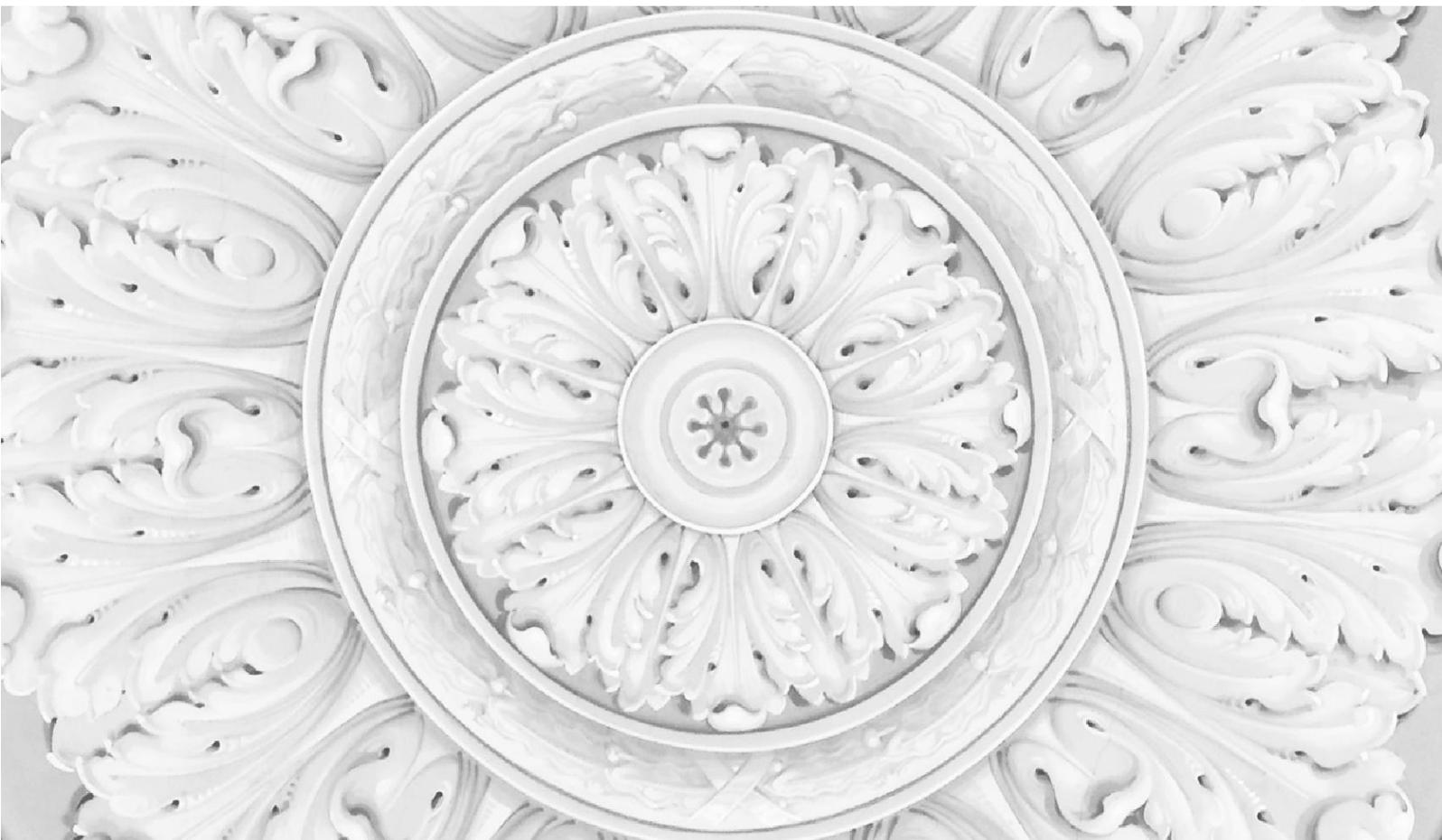




Банк России

Центральный банк Российской Федерации



WORKING PAPER SERIES

Elena Deryugina
Alexey Ponomarenko
Anna Rozhkova

When are credit gap estimates
reliable?

No. 34 / July 2018

Elena Deryugina

The Bank of Russia, Research and Forecasting Department

E-mail: Deryuginaeb@cbr.ru

Alexey Ponomarenko

The Bank of Russia, Research and Forecasting Department

E-mail: Ponomarenkoaa@cbr.ru

Anna Rozhkova

The Bank of Russia, Research and Forecasting Department

E-mail: anna.rozhkova92@gmail.com

The views expressed in this paper are those of the authors and do not necessarily represent the position of the Bank of Russia. We are grateful to Bence M  r , Andrea Roventini, Sergei Seleznev, Laurent Weil and Francesco Zanetti for their helpful comments and suggestions.

  The Central Bank of the Russian Federation, 2018

Postal address 12, Neglinnaya Street, Moscow, 107016
Telephone +7 (495) 771 9100, +7 (495) 621 6465 (fax)
Website www.cbr.ru

All rights reserved. The views expressed in this paper are solely those of the authors and do not necessarily reflect the official position of the Bank of Russia. The Bank of Russia disclaims liability for the contents of the paper. Any reproduction of these materials is subject to authors' permission.

Summary

We evaluate the reliability of credit gap measures estimated over time samples of different lengths. We augment our empirical analysis (which turned out to be somewhat inconclusive) with Monte Carlo experiments. For this purpose we build an agent-based model that realistically reproduces credit cycles and use it to generate the artificial data set. We found that 12–15 years of available data is sufficient for the estimation of reliable credit gaps (i.e. the reliability of credit gap estimates will not improve substantially as more data are added to the sample).

Key words: credit gap, credit cycle, countercyclical capital buffer, agent-based models

JEL-classification: C63, E37, E44, E51.

Table of Contents

INTRODUCTION	5
1. THE AGENT-BASED MODEL	6
2. DESIGN OF EXPERIMENTS AND EMPIRICAL VALIDATION	13
3. CREDIT GAP ESTIMATES' RELIABILITY ANALYSIS	18
4. CONCLUSIONS	24
REFERENCES	26
APPENDIX	30

INTRODUCTION

The debate about the reliability of credit gap estimates intensified after Basel III introduced it as a measure of the credit cycle phase and a guide for setting countercyclical capital buffers (CCB) (BCBS 2010). Although the useful properties of this indicator are confirmed generally for a broad array of countries and over a long time span, which includes the most recent crisis, criticism of this choice appears in the literature, focusing on several areas (see Drehmann and Tsatsaronis 2014 for a summary). In this paper we will address one such criticism: the problem of the credit gap's practical measurement and the end-point problems.¹ Edge and Meisenzahl (2011) and Geršl and Seidler (2015) emphasized potential weaknesses of the statistical technique behind the construction of the gap indicator, in particular the reliability of end-of-sample estimates and the difficulties associated with the changes in the pace of financial deepening (see also Égert et al. 2007 for a discussion of the latter issue). These problems are likely to be particularly relevant for emerging market economies where short time samples impede reliable trend and cycle decomposition of credit-to-GDP ratio series. In fact, Deryugina and Ponomarenko (2017) and Geršl and Jašová (2017) find that in emerging markets gap-based measures cannot outperform simple growth-based credit indicators as stand-alone early warning indicators for credit cycle peaks and banking crises, respectively.

This assessment is not, however, without its caveats. The underlying analysis is usually conducted by means of an examination of the early warning properties of credit gaps estimated recursively of the expanding time sample (i.e. 'real-time' credit gap estimates). Accordingly, the performance of these estimates is the average usefulness of the gap measures calculated over the range of time samples of different lengths (starting from the shortest possible one to the longest currently available). The result of this exercise is obviously different from the expected usefulness of the indicator in the coming years if it is used as an early warning indicator from now on. Similarly, if the credit gap estimates were not found useful during the period of rapid changes in the pace of financial deepening it does not mean that they may not eventually become useful once this process stabilizes. Therefore, there are good reasons to try to separately evaluate the reliability of credit gap measures estimated over different length time samples and to examine the evolution of credit gap measures' reliability during and after the changes in the financial deepening process.² This is the goal of our paper and the main contribution to the literature.

¹ Accordingly, the discussion of whether credit gap measures are good predictors of banking crises and whether multivariate models can improve these predictions stays beyond the scope of this paper.

² Incidentally, the evaluation of the predictive ability of forecasting models over a wide range of window sizes is a common issue in time series econometrics literature (Pesaran and Timmerman 2007, Inoue and Rossi 2012, Inoue et al. 2017).

Admittedly, this task is extremely data-demanding and is seriously impeded even for advanced economies. We therefore augment purely empirical analysis with Monte Carlo experiments. This approach allows us to generate a large number of artificial credit-to-GDP ratio series and examine the reliability of their trend-cycle decomposition under different circumstances (i.e. depending on the initial time sample availability and proximity to the change in the financial deepening process). The Monte Carlo approach is commonly applied for the analysis of output trend-cycle decomposition (see e.g. Nelson 1988, Basistha 2007, Gonzalez-Astudillo and Roberts 2016) and is not unprecedented for credit gap analysis (Drehmann and Tsatsaronis 2014). Our contribution to the literature in this regard is that instead of using a simple time series model as a data generator process, we employ a structural agent-based model that is arguably well suited for credit cycle modelling.

The rest of the paper is structured as follows. Section 2 describes the details of the agent-based model. Section 3 explains the set-up of the simulations and compares the properties of the artificial and actually observed series. Section 4 measures the reliability of artificial and actually observed credit gaps depending on the initial time sample used for estimation and proximity to the change in the financial deepening process. Section 5 concludes.

1. THE AGENT-BASED MODEL

Reproducing endogenous credit cycles is not an easy task for standard macroeconomic models. As pointed out by King (1994) and von Peter (2005), the representative agent assumption seriously impedes the modelling of two intermittent phases of debt build-up and subsequent credit busts. In fact, in his seminal work Minsky (1982) points out that it is the increased number of heavily indebted agents that sets off the debt-deflation mechanism while the aggregate debt-to-income ratio may only inform us about the probability of such an event. The heterogeneity of agents appears to be an essential for a model's capability to capture credit cycle developments.

We therefore employ an agent-based approach, which we think is the most appropriate strategy for our objectives.³ The merits of agent-based models (ABMs) are discussed in detail by Caiani et al. (2016), Fagiolo and Roventini (2017) and Haldane and Turrell (2018). Among them are several that are particularly relevant to our case. Firstly, we find it easier to set up a stock flow-consistent model that can be useful in analysing credit developments following the agent-based approach. Secondly, by employing an agent-based model, we can calculate the aggregate

³ See e.g. Hosszú and Mérő (2017), Popoyan et al. (2017) and Krug (2018) for the examples of agent-based models' application to macroprudential issues and credit cycle analysis. Chiarella and Di Guilmi (2017) employ an agent-based model specifically to conduct policy experiments in the context of debt deflation.

indicators from individual transactions and balance sheets without having to rely on the representative agent concept.

Our model's specification is mostly based on Ashraf et al. (2017) and Popoyan et al. (2017) with the following main differences. Firstly, we omit the comprehensive banking sector from our model and do not model fiscal policy. Secondly, our economy has one consumption and one investment good, as in Dosi et al. (2010). This is different from Ashraf et al. (2017) and Popoyan et al. (2017), who assume the existence of multiple consumption goods and also use them to define the input for fixed capital formation. Accordingly, the agents' labour types are also different. Finally, we also add several simplified elements from Dosi et al. (2010) to model endogenous productivity growth. In the rest of this section we describe the model in further detail.

In the model, there are N agents.

They have three skills that may be learned or forgotten. The first (entrepreneurial) skill allows an agent to own a shop. The other two skills allow an agent to engage in consumption or capital goods production (i.e. to set up or be employed by a shop in the respective industry).

The agents own deposits (shop owners have two separate – personal and corporate – accounts) and may take out loans. The deposits are remunerated at the policy interest rate less deposits spread ($IR_t - DS$).

The agents have partner shops from whom they buy goods in each of the markets. Each period with probability p^c they observe the prices in one of the other shops and change if those are lower than in their current shop.

The agents gauge their trend income $Y_{n,t}^p$ depending on their current income $Y_{n,t}$, which is wages or dividends received in the current period for workers or shop owners accordingly, or zero for unemployed individuals:

$$I_{n,t}^T = \lambda_a I_{n,t-1}^T + (1 - \lambda_a) I_{n,t}$$

All trend variables in this model are determined using such law of motion.

There is one consumption and one capital good in the economy. In the model, time steps correspond to months. In every time period t , the sequence of events runs as follows:

1. Agents attempt to improve the current labour productivity of the existing shops and set up new ones;
2. Loans are extended;
3. Workers are hired and wages are paid;
4. Capital goods are produced and sold;
5. Consumer goods are produced and sold;
6. Shop owners set prices and wages; shops enter and exit the market;
7. Agents determine which shops they will buy products from and loans are repaid.

At the end of each time step, aggregate variables (e.g. GDP, credit, prices, etc.) are computed, summing the corresponding microeconomic variables. Let us now turn to a more detailed description of these events.

1.1 Shop entry

In the spirit of Schumpeterian growth model of Dosi et al. (2010) shop owners may try to improve the labour productivity of their enterprise, and employees with entrepreneurial skill may try to set up a new shop for the consumer or capital goods industry (in the case where they have the needed skill).⁴ With probability p^e they randomly draw the labour productivity value:

$$lp_n^* = \max [0, \sim N(lp_t^{max}, \lambda_e lp_t^{max})]$$

where lp_t^{max} is currently the best vintage of labour productivity in the industry.

Agents proceed by calculating the expected cash flows (CF_n) from the project in excess of their current income. Namely, agents that are presently shop owners compare the cash flows with those expected, given their current labour productivity:

$$CF_n = (1 / lp_n - 1 / lp_n^*) Y_{n,t}^{trg} W_t$$

Workers compare the expected cash flows with their present permanent income:

$$CF_n = (P_t - W_t / lp_n^* - \lambda_c P_t^{cap}) Y_{n,t}^{trg} - I_{n,t}^p$$

where P_t is the current average price of the good (consumption or capital) that is planned for production, W_t is the current average wage in the industry under consideration, $\lambda_c P_t^{cap}$ is the cost of the input of capital goods that are needed for production and $Y_{n,t}^{trg}$ is the target output. Shop owners use the actual current target, and workers draw the random value from

$$Y_n^{trg} = \max [0, \sim N(1 / (J + 1), \lambda_z)] * YT_t$$

where J is the number of shops currently in the industry and YT_t is the trend output in the industry. The evolution of the trend⁵ output is gauged from the current total output of the industry Y_t as:

$$YT_t = \lambda_a YT_{t-1} + (1 - \lambda_a) Y_t$$

The present value of cash flows over h years is calculated using the real interest rate rr as the discount factor.

$$CFR_n = CF_n \frac{1 - (\frac{1}{1+rr})^h}{1 - (\frac{1}{1+rr})}$$

⁴ This approach may be regarded as a simplified version of Dosi et al.'s (2010) model as we do not model innovation and imitation separately, the volume of capital stock is fixed and firms cannot own different vintages of capital.

⁵ The evolution of all the trend variables in our model is determined using this law of motion.

The interest rate is transformed into real terms using the realized trend inflation of the average mark-up in the industry under consideration. The mark-up is defined as the price of the goods less the average labour and capital goods input costs.

To set up the project, an agent must buy Ω of capital goods (Ω is lower in the case where an agent is upgrading an existing shop). Accordingly, the cost of the project is the price of the required capital goods. An agent will decide to implement the project if the present values of the expected cash flows are larger (with a certain safety margin α_n that represents the agent-specific attitude towards risk taking that is determined at the model initialization stage) than the costs.

$$\frac{CFR_n}{\Omega P_t^{cap}} > \alpha_n$$

Once the decision is taken, the agent will calculate the funds needed to run the project. These are constituted by the set-up costs and the funds to cover the expected monthly labour and capital input costs (given the target output). If the amount of deposits owned by the agent is insufficient and his/her current debt service ratio⁶ ($DSR_{n,t}$) is lower than threshold value DSR_t^* , the agent will attempt to obtain a loan in the next phase (they will try to borrow the insufficient amount times the factor Φ). Subsequently, the agent will attempt to buy capital goods on the market. If he/she is successful, the agent will implement the project at the end of the month. If not, he/she will check the profitability of the project again in the next month and continue to accumulate the required capital goods. If the project is no longer profitable, the agent will abandon the idea and discard this vintage of lp_n^* .

1.2 Credit market, financial deepening and monetary policy

Agents with $DSR_{n,t}$ lower than DSR_t^* obtain loans with probability p_t^L , which represents current state of financial development. The agents attempt to get a loan after the shop entry and the price- and wage-setting phases.

Loan extension creates money, and the amount of deposits is increased accordingly. The maturity of loans is set to M years and the key interest rate plus credit spread ($IR_t + CS$).

Agents make principal and interest payments on their loans (these transactions destroy deposits). If they do not have sufficient funds, they miss the payment. If they miss the payment D times they default and if the defaulted borrower is currently a shop owner they exit the market. The loan is destroyed.

⁶ The debt service ratio is calculated as the sum of the principal and interest payments on all of an agent's loans expected this month as the ratio to permanent income.

In order to represent the financial deepening process (i.e. the increasing reliance on credit and its wider availability) DSR_t^* and p_t^L are determined by a random walk process with positive mean innovations:

$$DSR_t^* = DSR_{t-1}^* + v_t$$

$$p_t^L = p_{t-1}^L + \mu_t$$

The threshold debt service ratio and probability of getting a loan cannot be higher than DSR^{*max} and p^{Lmax} , respectively.

The policy interest rate is set by the central bank in response to the deviation of trend consumer good prices inflation (π_t) and unemployment (u_t) from the respective targets (policy interest rate cannot be negative):

$$IR_t = IR^* + \sigma^I (\pi_t - \pi^{trg}) - \sigma^U (u_t - u^{trg})$$

1.3 Labour market

At the start of this phase, shops will with probability p^F randomly fire unwanted workers if the current number of employees is higher than the target employment level.

Workers and unemployed individuals will ask a random shop (with an employment level lower than its target and from the industry for which the agent has the skill) for a wage offer. The agent will become employed by this shop if the offer is higher than the agent's current wage (unemployed agents accept the offer automatically).

The employed workers receive wages.

In order to introduce job market frictions, we utilize the system of job skills, which is different from Ashraf et al. (2017) and Popoyan et al. (2017). Agents learn and forget industrial and entrepreneurial skills with probabilities p^{s1} and p^{s2} ($p^{s1} > p^{s2}$). An agent will not forget the skill required to work in his/her current industry but may forget the other industrial skill with probability p^{s1} and learn it with probability p^{s2} .

An unemployed agent will learn the skill in the industry with a higher average wage with probability p^{s1} and forget it with probability p^{s2} . He/she will forget the other industrial skill with probability p^{s1} and learn it with probability p^{s2} . If an agent does not own a shop, they will forget the entrepreneurial skill with probability p^{s1} and learn it with probability p^{s2} .

1.4 Production

Goods are produced using the following production function:

$$Y_{n,t} = \min [(L_{n,t} + 1) l p_n, CI_{n,t} / \lambda_c]$$

where $L_{n,t}$ is the number of the shop's employees and $CI_{n,t}$ is the amount of capital goods purchased by the shop for input in the previous month. This is slightly different from Ashraf et al. (2017) and Popoyan et al. (2017) where firms convert purchased labour into goods and do not need capital good as an input.

The output is accumulated in shops' inventories.

Wages are paid to the workers in this phase by transferring deposits accordingly from shop owners to employees.

1.5 Goods market

There are two markets for capital and consumption goods. Capital goods are bought by shops for input and by agents wishing to set up a new shop or upgrade their existing one. Consumption goods are bought by all agents. Similarly to Ashraf et al. (2017) and Popoyan et al. (2017) the spending on consumption goods of an agent is determined as follows:

$$C_{n,t} = \beta_n^I I_{n,t}^T + \beta_n^W D_{n,t} - \beta_n^D DSR_{n,t} I_{n,t}^T + \varepsilon_{n,t} I_{n,t}^T$$

$D_{n,t}$ is financial wealth (personal deposits owned by the agent); β_n^I , β_n^W and β_n^D are agent-specific elasticities that link spending with income, wealth and debt burden; and $\varepsilon_{n,t}$ is a random component of the demand. Agents' spending cannot be negative or exceed their stock of deposits.

Agents buy goods from the associated shop, but, if it does not have sufficient inventories, they will randomly choose another one. Transactions are settled by transferring deposits from buyers to sellers and decreasing sellers' inventories accordingly.

1.6 Price and wage setting

After trading takes place, shops adjust their prices and wages and set targets for output, capital goods input and employment. In our model the firms make the adjustments basing them on the deviation of inventories to output ratios (both firm-specific and global) from their natural level (which is predetermined). Competitors' prices also influence an individual firm's price setting. This is slightly different from Ashraf et al. (2017) and Popoyan et al. (2017) where firms set their output target based on the previous sales.

New output targets $Y_{n,t}^{trg}$ are set based on the current firm-specific inventories to the target output ratio ($\Lambda_{n,t}^{FS}$) and the global (i.e. industry's overall) inventories to the trend output ratio (Λ_t^G) relative to the target level (Λ^*). Firms also react to the gap between current the industry's actual output and industry's trend output ($\frac{Y_{t-1}}{Y_{t-1}^T}$):

$$Y_{n,t}^{trg} = Y_{n,t-1}^{trg} (1 + \xi^{Y1} (\Lambda_{n,t}^{FS} - \Lambda^*)) (1 + \xi^{Y2} (\Lambda_t^G - \Lambda^*)) (1 - \xi^{Y3} (\frac{Y_{t-1}^T}{Y_{t-1}^T} - 1))$$

The target employment $L_{n,t}^{trg}$ and capital goods input $CI_{n,t}^{trg}$ are determined as:

$$L_{n,t}^{trg} = Y_{n,t}^{trg} / lp_n - 1$$

$$CI_{n,t}^{trg} = \lambda_c Y_{n,t}^{trg}$$

Wages $W_{n,t}$ are set based on the new labour demand and the gap between current wage offer and the trend average wage in the industry:

$$W_{n,t} = W_{n,t-1} (1 + \xi^{W1} (L_{n,t}^{trg} - L_{n,t}) / L_{n,t}^{trg}) (1 - \xi^{W2} (W_{n,t-1} - W_{t-1}^T) / W_{n,t-1})$$

Prices $P_{n,t}$ are set based on the current firm-specific inventories to the target output ratio ($\Lambda_{n,t}^{FS}$) and the global (i.e. industry's overall) inventories to the trend output ratio (Λ_t^G) relative to the target level (Λ^*). Firms also react to the gap between current price and the trend average price in the industry:

$$P_{n,t} = P_{n,t-1} (1 + \xi^{P1} (\Lambda_{n,t}^{FS} - \Lambda^*)) (1 + \xi^{P2} (\Lambda_t^G - \Lambda^*)) (1 - \xi^{P3} (P_{n,t-1} - P_{t-1}^T) / P_{n,t-1})$$

None of these variables may be negative. Price setting cannot result in negative mark-ups. Labour demand cannot be larger than L^{\max} and output target cannot be larger than $L^{\max} * lp_n$.

Shop owners will leave the expected amount needed to operate their shops (that is, to cover planned labour and capital input costs and debt service) in the corporate deposit account. The rest is transferred to the personal account and is considered as a dividend payment that affects the permanent income. Note that this amount may be negative, representing losses and requiring a transfer from the personal to the corporate account. If the funds in the personal account are insufficient, the shop owner will try to get a current loan (they will try to borrow the insufficient amount times the factor Φ) and will exit the market if they fail. The shop owner will also exit the market if their trend income or output target becomes zero.

At this stage agents who have bought enough capital goods upgrade their shops and set up new ones.

1.7 Market exit

A shop owner who exits the market becomes unemployed, sells all the remaining inventories at a discount Θ and transfers all the remaining funds to his/her personal account. Random Z shops from the respective industry will buy the $1/Z$ fraction of the inventories given that they have enough funds. If there are not enough shops with sufficient funds the excess inventories are wasted.

1.8 Calibration

The parameters of the model are presented in Table A1 in the Appendix. The criterion for the choice of the parameters was the ability of the model to generate series of credit-to-GDP ratios with realistic properties. These properties are discussed in more detail in Section 2.

2. DESIGN OF EXPERIMENTS AND EMPIRICAL VALIDATION

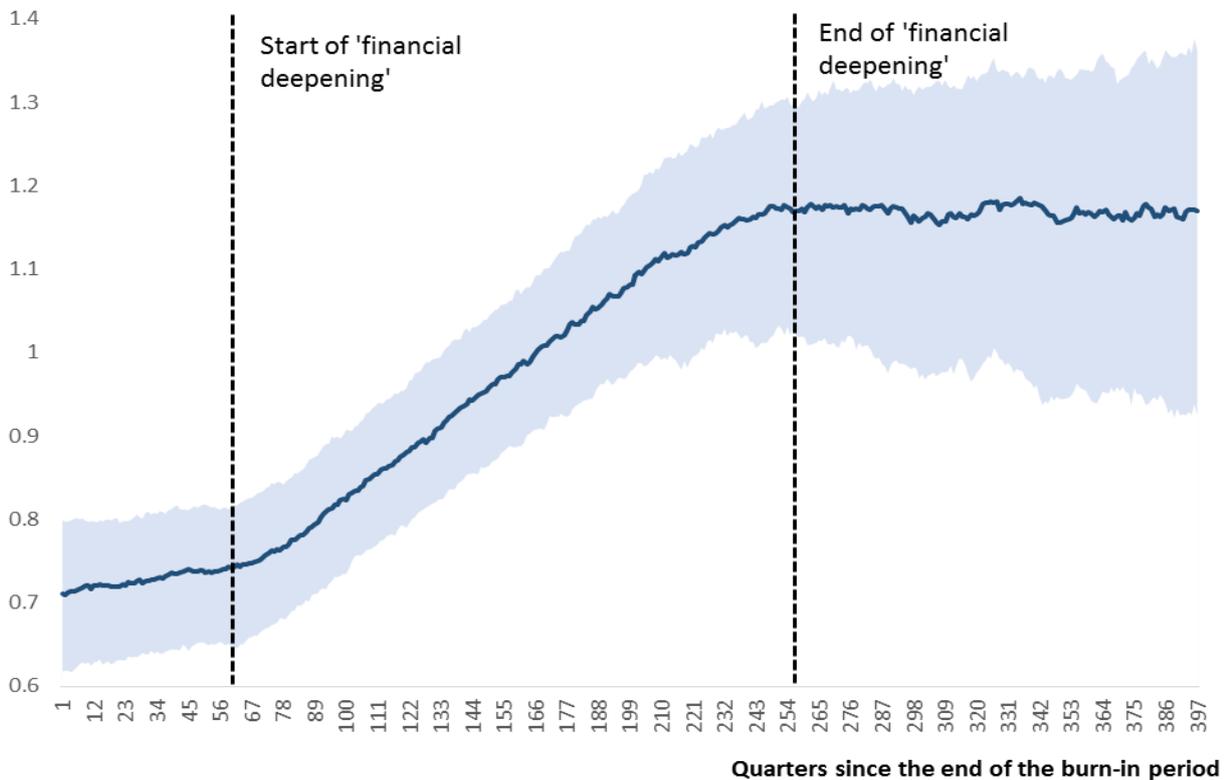
We proceed by obtaining the set of artificial credit, nominal and real GDP series by running the ABM simulations. We initialize the model (i.e. generate agents and their characteristics), simulate the model and start to collect the data after 150 years of the burn-in period.⁷ After 15 years the financial deepening begins (as described in Section 1.2) and triggers the drift in credit-to-GDPs. This drift comes to an end in approximately 50 years after it begins when credit availability and threshold debt service ratio (DSR) reach their maximum values. The credit-to-GDP ratios increase by approximately 1.5 times during this period (Figure 1).⁸ We conduct 1000 independent simulation runs and treat the observations between the start and the end of financial deepening as the main sample of artificial data for our analysis.⁹ We proceed with empirical validation of our model and examine the properties of the obtained series.

⁷ Note that although the model generates monthly observations, we aggregate them into quarterly data.

⁸ The absolute values of credit-to-GDPs ratio vary depending on the model's initialization but generally range from 30% to 60%.

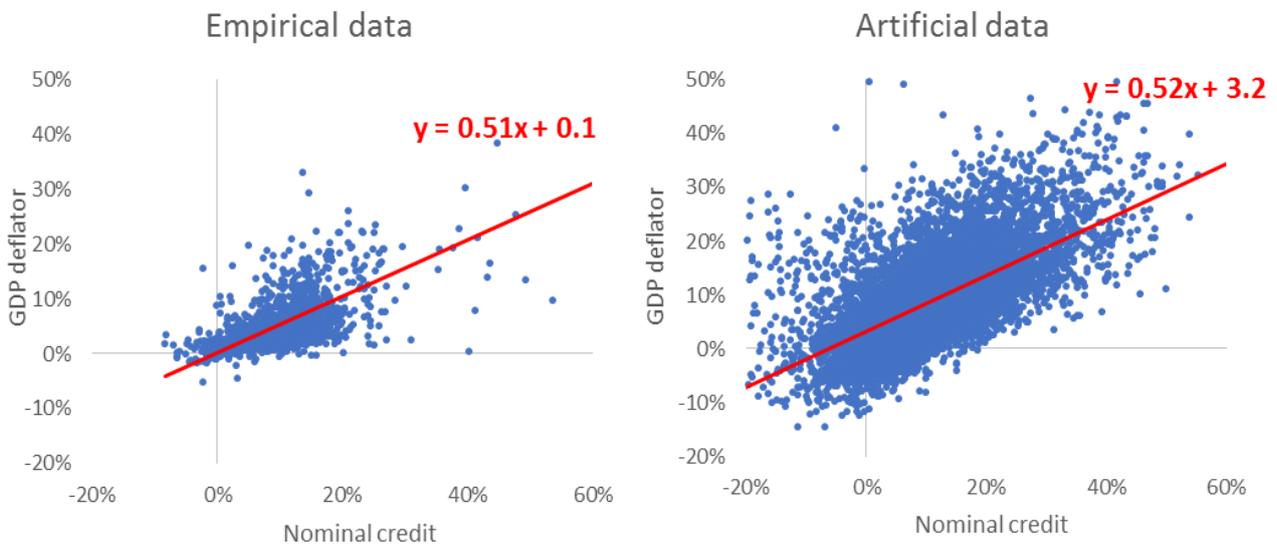
⁹ We discard some of the most extreme realizations from our sample. Namely, we discard the simulations where the output of either consumption of capital goods dropped to zero at some point and the simulation where the price level deflation occurred for 20 consecutive years. The overall proportion of discarded observations amounted to less than 5% of total.

Figure 1. Distribution of artificial credit-to-GDPs (ratio to simulation-specific mean, the confidence band is based on 16th and 84th percentiles)



As outlined in Fagiolo et al. (2007) and Fagiolo and Roventini (2017) one of the approaches to the ABM validation implies examination of how the patterns in the model's simulated macroeconomic coincide with real macro data. The choice of which stylized facts to reproduce largely depends on the objective of the model. In our case this choice is quite straightforward: we develop our model to generate realistic fluctuations of credit-to-GDP ratio and, therefore need to make sure that joint dynamics in credit, output and prices is in line with empirical regularities.

We use the cross-section of 27 countries to validate the model (see Table A2 in the Appendix for the details of the empirical data set). Firstly, we examine simple correlations of annual growth rates of GDP and GDP deflators with real and nominal credit, respectively (Figures 2–3). The results indicate that the model is able to replicate the co-movement of output and price level with credit expansion.

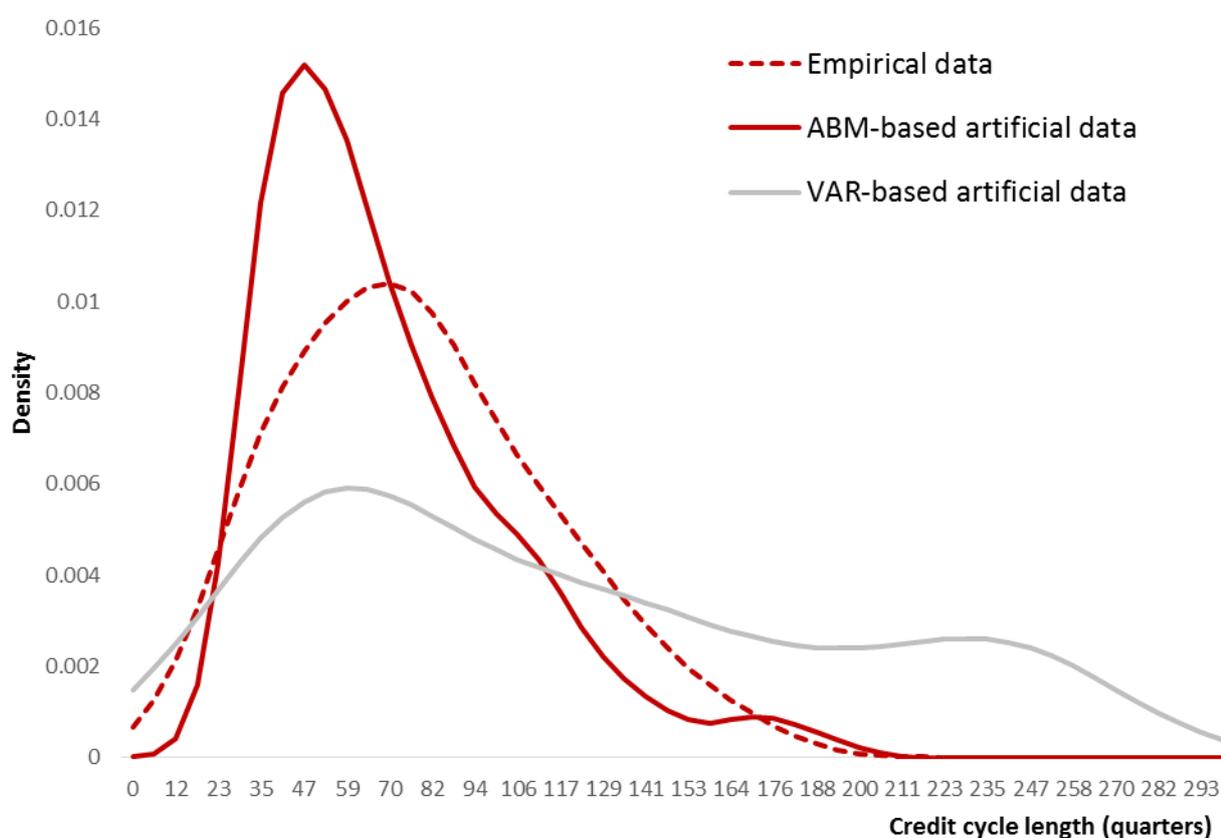
Figure 2. Correlation between annual growth of GDP and real credit.**Figure 3.** Correlation between annual growth of GDP deflator and nominal credit.

Admittedly, this type of correlation may be easily captured by any time series model. The important advantage of the ABM is that it is also able to replicate the emerging cyclical pattern that is observed empirically in the fluctuations of credit-to-GDP ratios. We illustrate this using several indicators.

Firstly, we measure the lengths of the credit cycles in artificial and in empirical data. Following Drehmann et al. (2012) we define credit cycle peaks as local maxima of a credit-to-GDP ratio over a nine-quarter window and set the minimum cycle length equal to 30 quarters. We measure the lengths between the peaks and plot its kernel density in Figure 4. For comparison,

we estimate a panel VAR model using our empirical data set and generate another set of artificial credit-to-GDP ratio series.¹⁰ Notably, the credit cycles generated by the ABM usually have a length of 12 years, which is generally consistent with empirical data.¹¹ By contrast, the VAR model cannot reproduce these regularities and generate fluctuations of credit-to-GDP ratio that do not exhibit distinct cyclical regularities. These results emphasize the importance of the agent-based approach, which is able to replicate the emergence of cyclical patterns from random underlying shocks.¹²

Figure 4. Kernel density of credit cycle lengths.



¹⁰ Namely, we estimate a tri-variate panel VAR model comprising four lags of seasonally adjusted credit, GDP and GDP deflator series from our empirical data set. Based on Cholesky decomposition, we estimate the covariance matrix for orthogonalized innovations. Using this set of parameters and estimates of the shocks' variance we generate 200 sets of 50 years long series and calculate respective credit-to-GDP ratios.

¹¹ See Table A3 in the Appendix for other characteristics of actual and artificial credit cycles.

¹² Admittedly, there are an increasing number of DSGE models developed for credit cycle analysis (see among others Kiyotaki and Moore 1997, Iacoviello 2005, Jakab and Kumhof 2015, Mumtaz and Zanetti 2016 and Gerba and Żochowski 2017 for some recent examples on the integration of financial frictions in structural models of the business cycle). Nevertheless, we believe that in terms of generating endogenous credit cycles ABM appear to be a more practical tool.

Finally, we evaluate the properties credit gap measures in the artificial and real data sets. Arguably, in a world with distinct credit cycles the positive deviations of credit-to-GDP ratio from the trend will be a good predictor for credit cycles peaks if they are reliable (i.e. real-time estimates are close to the *ex post* values). In order to examine this property, we estimate credit gap indicators by applying one-sided Hodrick–Prescott filter ($\lambda=400000$) to the log of credit-to-GDP ratio recursively over the expanding window (with the minimum size of 12 quarters). We expect the credit gap to start issuing the warning signal 12 quarters before the credit cycle peak (defined as described earlier) and exclude four observations after the peak from the analysis. We formally evaluate the credit gap’s early warning properties in terms of the area under the ROC curve (AUC): a statistical methodology that captures the trade-off between true positives and false positives for the full range of policymakers’ preferences (see Drehmann and Juselius 2014 for details). The results for empirical and two artificial data sets are reported in Table 1. They indicate that the early warning properties of credit gap estimates generated by the ABM are generally comparable with those observed in reality.¹³ Meanwhile, the usefulness of credit gaps generated by the VAR is very low (presumably due to the lack of pattern in credit-to-GDP ratio fluctuations).

Table 1. Credit gap AUC in empirical and artificial data sets

<i>Dataset</i>	<i>AUC</i>
Empirical	0.63
ABM-based	0.7
VAR-based	0.54

We, therefore, conclude that our ABM realistically¹⁴ reproduces the fluctuations of credit-to-GDP ratio and proceed with a more detailed analysis of credit gap estimates’ reliability.

¹³ Admittedly, the usefulness of credit gaps in the ABM world is somewhat higher. This may indicate that our data generator process is more homogeneous (both in time and across cross-sections) than in the real world. In addition the series of some of the economies in the empirical data set are significantly shorter than 50 years available in the ABM-based data set (which as we will see later may affect the reliability of credit gap estimates). Note, however, that we do not intend to use the ABM to evaluate the indicator’s usefulness in absolute but to examine the evolution of its properties in different circumstances.

¹⁴ More accurately, the model is able to reproduce credit cycles in line with historical regularities. Obviously, this means that the results of credit gap reliability analysis are generally applicable if the intensity of credit developments remains similar to the one observed previously.

3. CREDIT GAP ESTIMATES' RELIABILITY ANALYSIS

We employ two measures of the credit gap estimates' reliability.

Firstly, we measure the magnitude of the indicators' typical revision by comparing the real-time estimate (i.e. the estimate of the credit gap obtained in a given period using only the data available in that period) with the *ex post* value (we use the estimate obtained when 20 more years of data are added to the sample). This is a standard measure of reliability (see e.g. Edge and Meisenzahl 2011) but it may be difficult to judge to what extent the practical value of an indicator deteriorates due to such revisions.

Therefore, secondly, we measure the usefulness of credit gap estimates for real-time identification of credit cycle phases (i.e. prediction of turning points) in terms of AUC, as described in Section 2. Essentially, we calculate the (formally quantifiable) usefulness of real-time gap estimates in distinguishing between cyclical and trend developments that are identified *ex post*. Arguably, this is an appropriate definition of reliability.

The usefulness of credit gap estimates is known to be heavily dependent on the data available for their calculation.¹⁵ A detailed exploration of this relationship is an important part of our analysis. We conduct our estimates separately for the subsets of credit gap measures calculated based on the time samples with different lengths. Namely, instead of using an expanding time sample (as in Section 2) we estimate credit gaps using a fixed-length rolling time samples. We start with the minimum time-sample length of 12 quarters and then increase the size by one-quarter at a step (up to the maximum of 180 quarters), obtaining the credit gap estimates and the measures of their reliability (i.e. revision magnitude and AUC) for each of these steps.

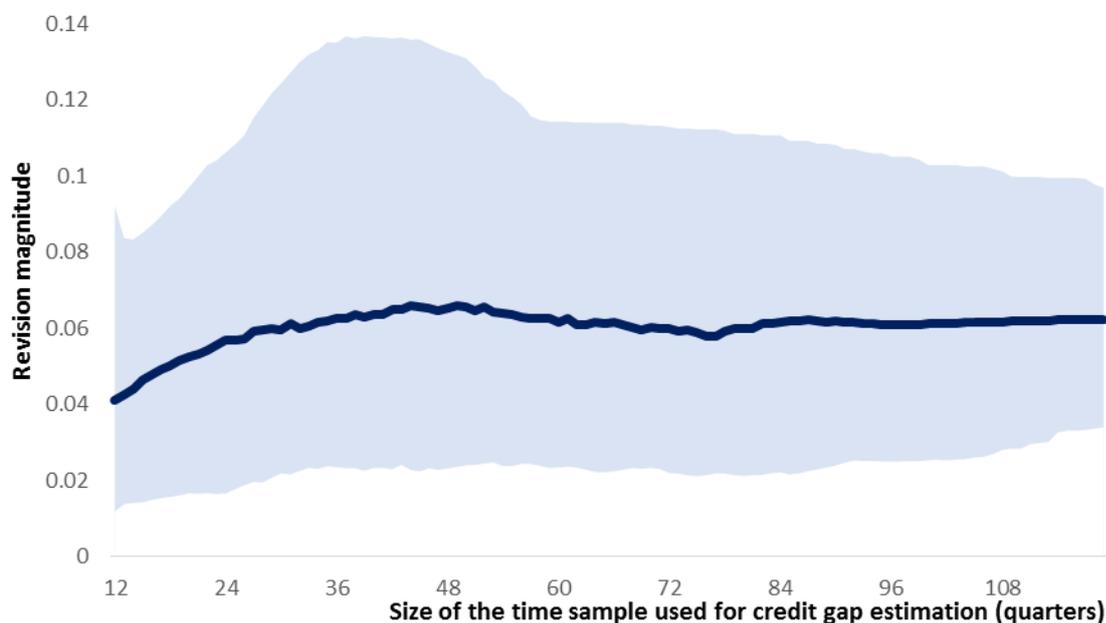
3.1 Empirical data set

We begin by conducting the evaluation on the empirical data set. The typical revisions for the subsets of credit gaps are presented in Figure 5. As the time sample increases there is an initial rise¹⁶ in the revision magnitude but subsequently it stabilizes around 6 percentage points. This finding is counterintuitive as it suggests that the reliability of the credit gap estimate is not sensitive to the length of the time sample used for calculation.

¹⁵ BCBS (2010) states that a country should have a long time series (at least 20 years of quarterly data) to employ credit gaps for countercyclical capital buffer implementation. Geršl and Jašová (2017) report the difference in the usefulness of credit gaps estimated using the time samples of various lengths.

¹⁶ This pattern may be associated with the fact that very short samples of credit-to-GDP ratio fluctuations can be better fitted with the trend and the gap measures therefore tend to have smaller variance.

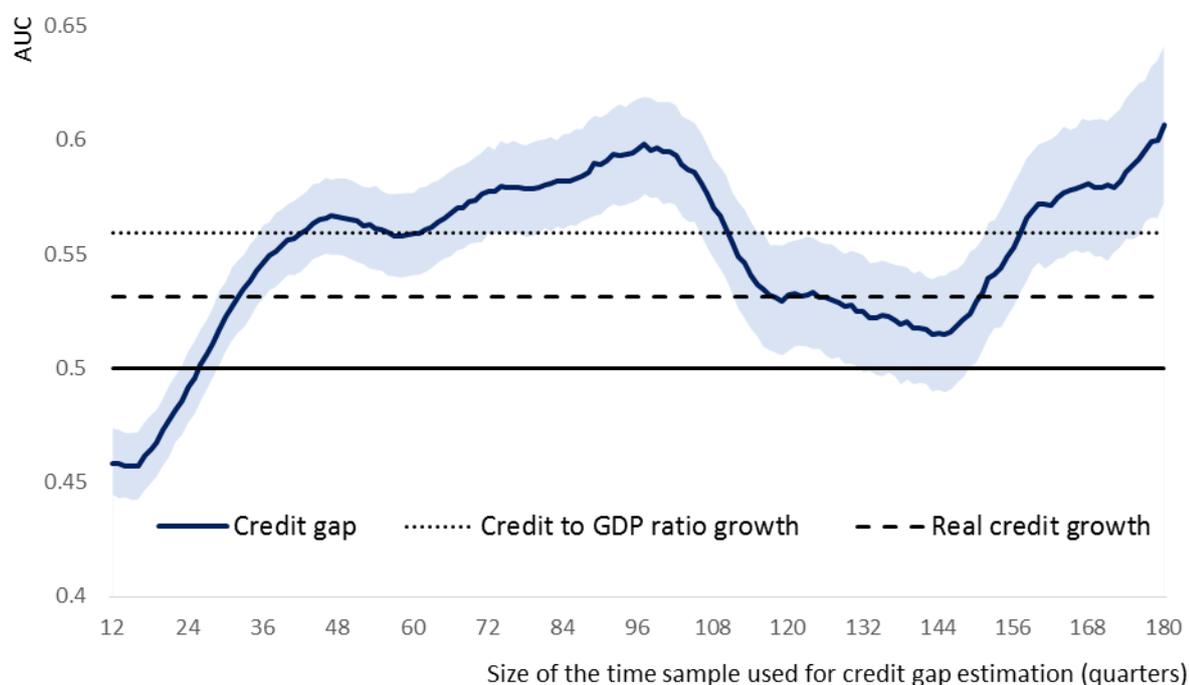
Figure 5. Typical revision of credit gaps (median, 16th and 84th percentiles).



The AUCs of credit gaps estimated based on time samples of various lengths are presented in Figure 6. The results indicate that the performance of the indicators increases rapidly with each quarter added until 12 years of observations are available. As the time samples' length increases further the AUCs continue to rise at a somewhat slower pace. The indicators' performance seems to be at its highest when 25-year-long time samples are used. We may compare these results with some of the benchmarks. Firstly, the credit gaps' AUCs may be expected to be informative (i.e. larger than 0.5) once 5 years of data is available. Credit gaps are expected to be better predictors of credit cycle peaks than the alternative indicators¹⁷ (real credit growth and credit-to-GDP ratio growth) once the time sample length exceeds 10 and 20 years accordingly. Notably, there are also unexpected fluctuations of AUCs for the time samples from 25 to 40 years.

¹⁷ Note that as growth-based measures do not require de-trending, their calculation (and therefore their AUC) does not depend on the time sample length.

Figure 6. AUCs of credit gaps and alternative indicators (with 90% bootstrapped confidence band).



Overall, we need to admit that the results of empirical analysis (in particular for long time samples) do not seem to be fully credible. One possible explanation (in addition to a blatantly insufficient number of observations) is the heterogeneity across sub-samples used to evaluate credit gaps' properties. Note that while the properties of credit gap estimates over the short time samples are tested over the whole period under consideration (i.e. using the observations from 1960s to 2010s),¹⁸ the reliability of credit gaps estimated using, for example, 20-year-long time samples may only be tested starting from the 1980s. The likely heterogeneity in the intensity of credit cycle developments over the last decades may distort the results of this exercise. These considerations underline the necessity to cross-check our findings using the artificial data set.

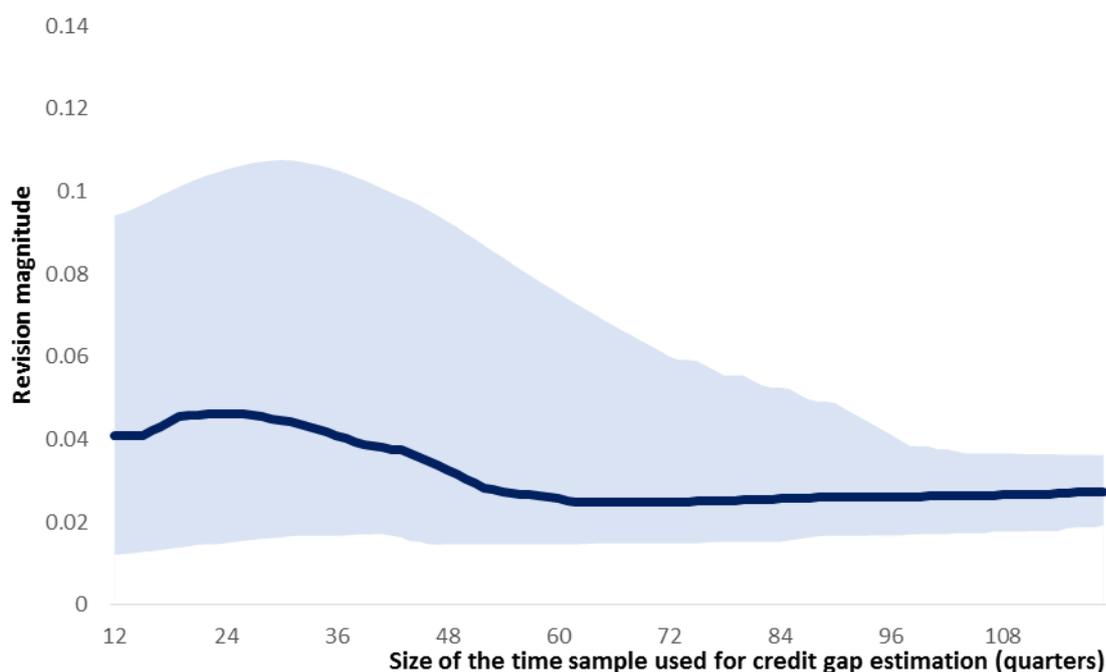
3.2 Artificial data set (without structural breaks)

We proceed by conducting the same exercise for the artificial series generated by the ABM. Here we analyse the 50-year-long periods between the start and the end of financial deepening in 1000 independent cross-sections.

¹⁸ In fact, as the available time samples are shorter for some of the countries in our data set we have fewer cross-sections for evaluation of the reliability of the respective subsets of credit gaps.

The results appear to be much more distinct compared to those obtained from the empirical data set. We found that the typical credit gaps' revision magnitude is higher for the short time samples (the typical revision is about 5 percentage points which is similar to the empirical results) but starts to decline as the available data increases (Figure 7). It stabilizes around 2.5 percentage points once the time sample reaches 15 years.

Figure 7. Typical revision of credit gaps (median, 16th and 84th percentiles).

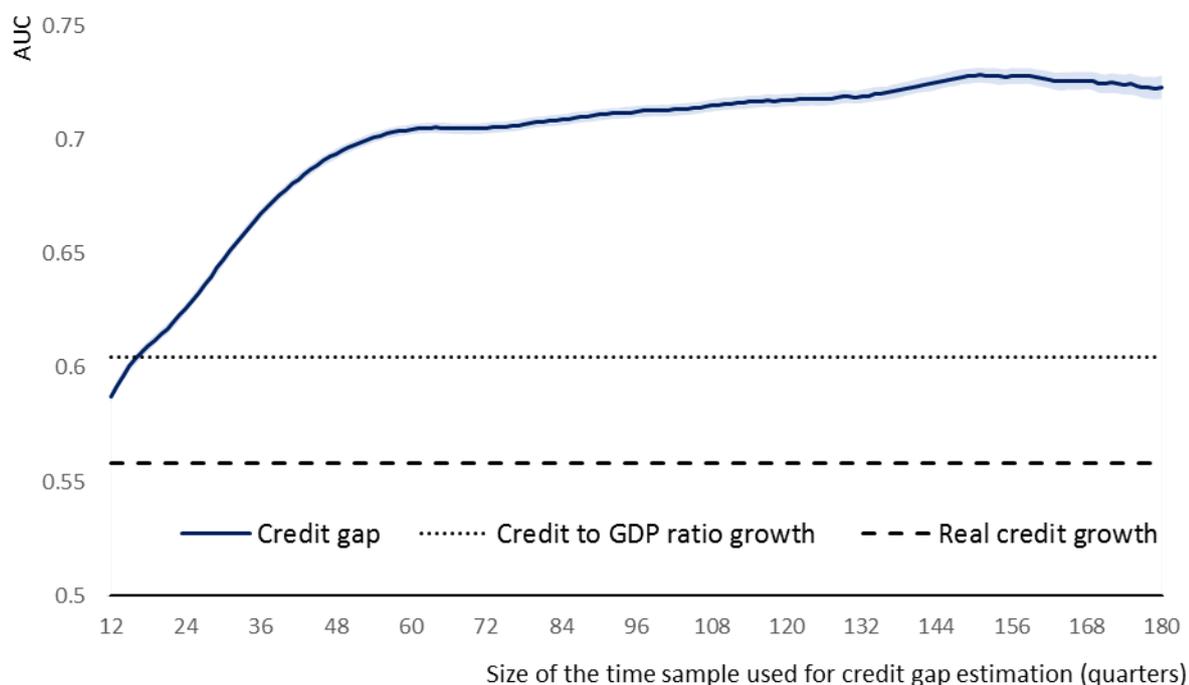


The AUC analysis confirms that the rapid improvement of credit gap estimates' reliability takes place while the length of the time sample expands up to 12–15 years (Figure 8).¹⁹ Further lengthening of the time sample only leads to moderate improvement.²⁰ In the artificial world the credit gaps estimated using a 5-year-long time sample are expected to outperform the alternative indicators.

¹⁹ Let us once again emphasize the importance of measuring the credit gap reliability for different time sample lengths. Note, that the standard approach to measuring the predictive power of credit gap indicators (i.e. using a recursively expanding time sample for its estimation) conducted for a country with 10-year-long data would imply averaging the performance of indicators estimated over the range of time samples (in our case from 3 to 10 years). The resulting AUC estimate (between 0.58 and 0.68) would be considerably lower than the expected AUC of credit gaps estimated 'from this moment on' (0.68).

²⁰ Note that (as may be deduced from Figure 4) the 15-year-long periods are likely to be sufficiently long to contain both contractionary and expansionary phases of credit cycle of the same length, which appears to be crucial for reliable credit gap estimation.

Figure 8. AUCs of credit gaps conditional on the used time sample length and alternative indicators (with 90% bootstrapped confidence band).



3.3 Artificial data set (with structural breaks)

We proceed by examining to what extent instability in the dynamics of credit-to-GDP ratio may lead to deterioration of credit gap estimates' reliability. For this purpose we conduct the following exercise.

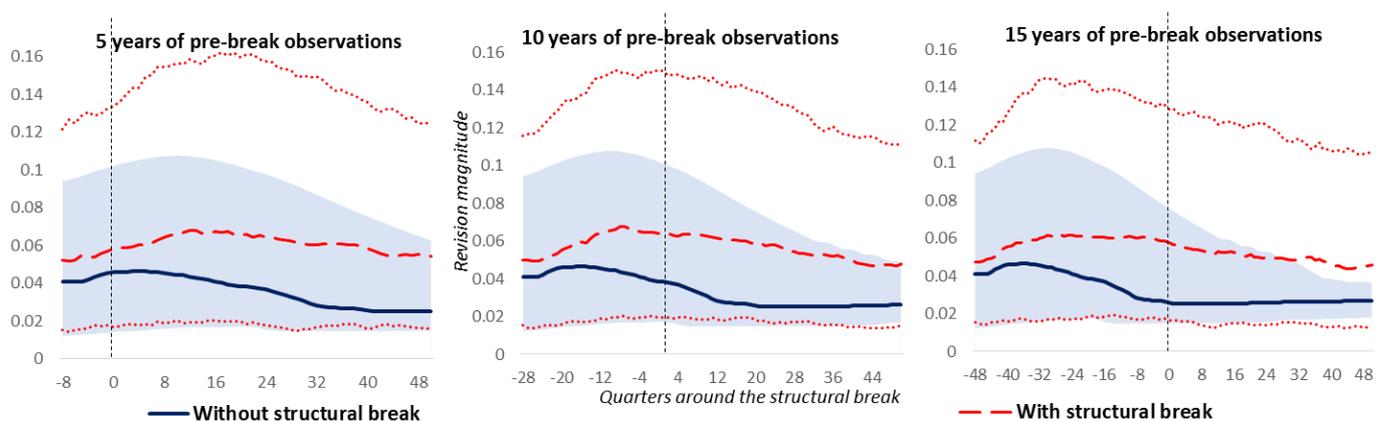
We label the periods of the start and the end financial deepening (as discussed in Section 2) as structural breaks.²¹ We estimate credit gaps over the expanding time sample that starts before the structural break. Namely, we conduct this analysis over three alternative time samples that contain 15, 10 and 5 years of pre-break observations as well as the benchmark sample that does not contain pre-break data. Accordingly, each of 1000 simulations provides us with two time samples containing structural breaks: the first one begins before and expands into the period of financial deepening, the second one starts shortly before the end of financial deepening. The results are pooled from both time samples and across all cross-sections for the estimation of the calculation of overall reliability indicators. We evaluate the reliability separately for the subsets of

²¹ We believe that these periods represent the structural breaks of a considerable magnitude: the transition between positive drift and flat dynamics of credit-to-GDP ratio have been rarely observed in practice.

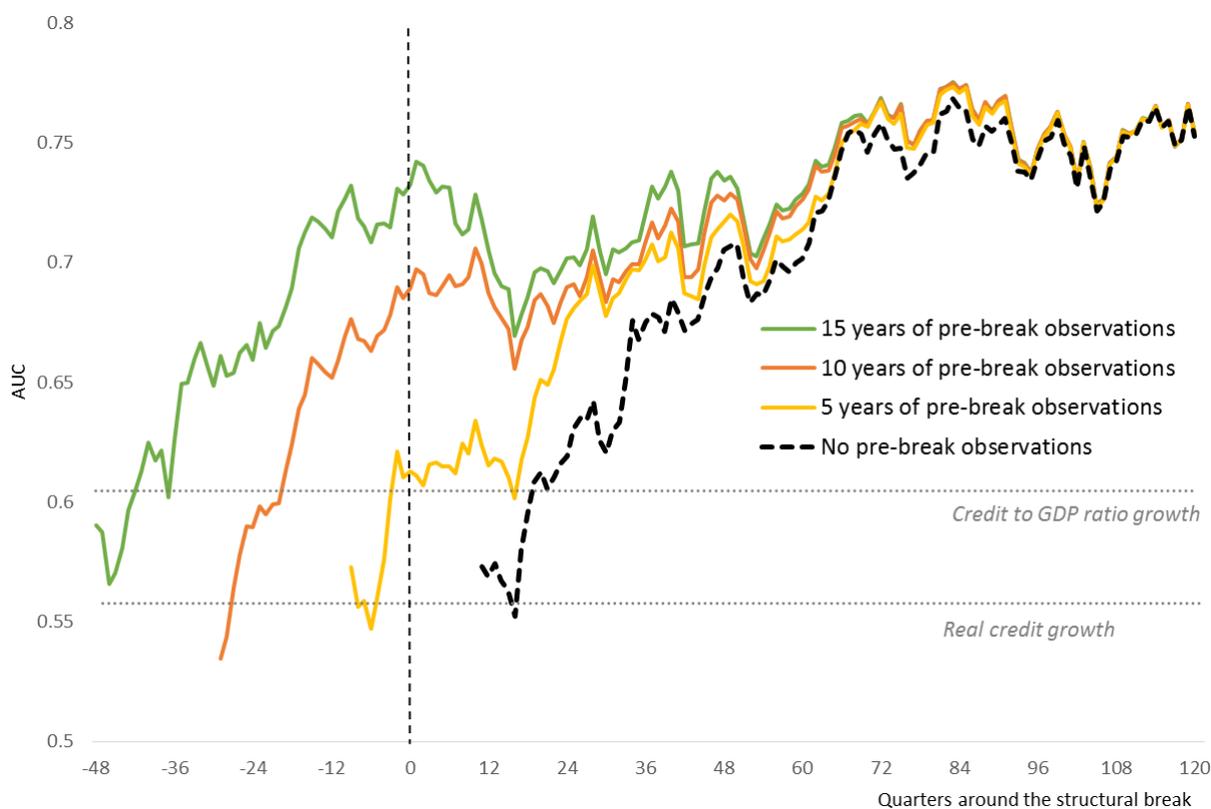
credit gap estimates depending on the vicinity to the structural break (as in the previous exercises we assume the minimum time sample for credit estimation is 12 quarters).

We present the typical revision of credit gap estimates over the time sample with structural breaks in Figure 9 (for comparison we also plot the typical revisions of credit gaps obtained over the homogeneous time sample with the respective length in Section 3.2). The presence of instability leads to a visible increase in typical revisions of both pre- and post-break estimates.

Figure 9. Typical revision of credit gaps in the vicinity of the structural break (median, 16th and 84th percentiles).



The AUCs across different subsets of credit gap estimates are presented in Figure 10. The reliability of credit gap estimates deteriorates after the structural break (although not dramatically). The subsequent improvement is slow and the highest levels of AUCs are obtained only after more than 15 years pass after the structural break. In fact, the reliability of credit gap estimates that are obtained using only 12 or more years of post-break observations is close (but never superior) to the usefulness of credit gaps estimated using longer time samples but with pre-break observations.

Figure 10. AUCs of credit gaps in the vicinity of the structural break.

4. CONCLUSIONS

The reliability of credit gap estimates is known to be heavily dependent on the length of data available for their calculation. Therefore, the problems of the credit gap's practical measurement and the end-point problems are likely to be particularly relevant for emerging market economies where short time samples impede reliable trend and cycle decomposition of credit-to-GDP ratio series. Arguably, application of the methods calibrated for advanced markets to emerging economies may be inappropriate.

It may also be misleading to equate the usefulness of the credit gap indicator observed over a given time sample with the expected usefulness in the future. The traditional analysis of the early warning properties of credit gap indicators implies its recursive estimation of the expanding time sample (i.e. in 'real-time'). Accordingly, the outcome of such an exercise is the average usefulness of the gap measures estimated over the range of time samples with different lengths (starting from the shortest possible one to the longest one currently available), which is obviously different from the expected usefulness of the indicator in the coming years if it is used as an early warning indicator from now on. Similarly, if the credit gap estimates were not found useful during the period of rapid changes in the pace of financial deepening it does not mean that they may not

eventually become useful once this process stabilizes. Therefore, there are good reasons to try to separately evaluate the reliability of credit gap measures estimated over time samples of different lengths and to examine the evolution of credit gap measures' reliability during and after the changes in the financial deepening process. This is the goal of our paper.

Admittedly, this task is extremely data-demanding and is seriously impeded even for advanced economies. We therefore augment our empirical analysis (which turned out to be somewhat inconclusive) with Monte Carlo experiments. For this purpose we build an agent-based model that realistically reproduces credit cycles and use it to generate the artificial data set.

We found that 12–15 years of available data is sufficient for the estimation of reliable credit gaps (i.e. the reliability of credit gap estimates will not improve substantially as more data are added to the sample). These results may be regarded as supportive for credit gap-based guidance for setting countercyclical capital buffers even in emerging markets, where the time samples of this length are generally available.

A pronounced change in the underlying pace of financial deepening leads to the deterioration of the credit gaps' reliability. Nevertheless, the credit gaps remain useful in real-time identification of credit cycle phases even in the presence of such structural breaks. Attempting to exclude the pre-break observations from the time sample may be unadvisable as their presence still improves the credit gaps' reliability (compared to those obtained using a shorter time sample).

REFERENCES

Ashraf, Q., Gershman, B. and Howitt, P. (2017). Banks, Market Organization, and Macroeconomic Performance: An Agent-Based Computational Analysis, *Journal of Economic Behavior & Organization*, 135: 143–180.

Basel Committee on Banking Supervision (BCBS). (2010). Guidance for National Authorities Operating the Countercyclical Capital Buffer, December.

Basistha, A. (2007). Trend-Cycle Correlation, Drift Break and the Estimation of Trend and Cycle in Canadian GDP (Corrélation tendance-cycle, discontinuité, et estimation de la tendance et du cycle dans le PIB canadien), *The Canadian Journal of Economics / Revue canadienne d'économique*, 40 (2): 584–606.

Caiani, A., Godin, A., Caverzasi, E., Gallegati, M., Kinsella, S. and Stiglitz, J. (2016). Agent-Based Stock Flow Consistent Macroeconomics: Towards a Benchmark Model, *Journal of Economic Dynamics & Control*, 69: 375–408.

Chiarella, C. and Di Guilmi C. (2017). Monetary Policy and Debt Deflation: Some Computational Experiments, *Macroeconomic Dynamics*, 21: 214–242.

Deryugina, E. and Ponomarenko, A. (2017). Real-Time Determination of Credit Cycle Phases in Emerging Markets. Bank of Russia Working Paper Series No. 17.

Dosi, G., Fagiolo, G. and Roventini, A. (2010). Schumpeter Meeting Keynes: A Policy-Friendly Model of Endogenous Growth and Business Cycles, *Journal of Economic Dynamics & Control*, 34: 1748–1767.

Drehmann, M., Borio, C. and Tsatsaronis, K. (2012). Characterising the Financial Cycle: Don't Lose Sight of the Medium Term! BIS Working Paper No. 380.

Drehmann, M. and Juselius, M. (2014). Evaluating Early Warning Indicators of Banking Crises: Satisfying Policy Requirements, *International Journal of Forecasting* 30, 759–780.

Drehmann, M. and Tsatsaronis, K. (2014). The Credit-to-GDP gap and Countercyclical Capital Buffers: Questions and Answers, *BIS Quarterly Review* (March), 55–73.

Edge, R. and Meisenzahl, R. (2011). The Unreliability of Credit-to-GDP Ratio Gaps in Real-Time: Implications for Countercyclical Capital Buffers, *International Journal of Central Banking*, 7 (4): 261–298.

Égert, B., Backé, P. and Zumer, T. (2007). Private-Sector Credit in Central and Eastern Europe: New (Over) Shooting Stars? *Comparative Economic Studies*, 49: 201–231.

Fagiolo, G., Moneta, A. and Windrum, P. (2007). A Critical Guide to Empirical Validation of Agent-Based Models in Economics: Methodologies, Procedures, and Open Problems, *Computational Economics*, 30 (3): 195–226.

Fagiolo, G. and Roventini, A. (2017). Macroeconomic Policy in DSGE and Agent-Based Models Redux: New Developments and Challenges Ahead, *Journal of Artificial Societies and Social Simulation*, 20 (1): 1–1.

Gerba, E. and Żochowski, D. (2017). Knightian Uncertainty and Credit Cycles, ECB Working Paper 2068.

Geršl, A. and Jašová, M. (2017). Credit-based Early Warning Indicators of Banking Crises in Emerging Markets, *Economic Systems*, <https://doi.org/10.1016/j.ecosys.2017.05.004>

Geršl, A. and Seidler, J. (2015). Countercyclical Capital Buffers and Credit-to-GDP Gaps: Simulation for Central, Eastern, and South-eastern Europe, *Eastern European Economics* 53 (6): 439–465.

Gonzalez-Astudillo, M. and Roberts, J.M. (2016). When can Trend-Cycle Decompositions be Trusted? Finance and Economics Discussion Series 2016-099. Washington: Board of Governors of the Federal Reserve System.

Haldane, A. and Turrell, A. (2018). An Interdisciplinary Model for Macroeconomics, *Oxford Review of Economic Policy*, 34 (1–2): 219–251.

Hosszú, Z. and Mérő, B. (2017). An Agent-based Keynesian Model with Credit Cycles and Countercyclical Capital Buffer, MNB Working Papers 5.

Iacoviello, M. (2005). House Prices, Borrowing Constraints and Monetary Policy in the Business Cycle. *American Economic Review*, 95: 739–764.

Inoue, A., Jin, L. and Rossi, B. (2017). Rolling Window Selection for Out-of-Sample Forecasting with Time-varying Parameters, *Journal of Econometrics*, 196: 55–67.

Inoue, A. and Rossi, B. (2012). Out-of-Sample Forecast Tests Robust to the Choice of Window Size, *Journal of Business & Economic Statistics*, 30 (2): 432–453.

Jakab, Z. and Kumhof, M. (2015). Banks are not Intermediaries of Loanable Funds – And why this Matters, Bank of England Working Papers 529.

King, M. (1994). Debt-Deflation: Theory and Evidence, *European Economic Review*, 38 (3–4): 419–445.

Kiyotaki, N. and Moore, J.H. (1997). Credit Cycles. *Journal of Political Economy* 105: 211–248.

Krug, S. (2018). The Interaction between Monetary and Macroprudential Policy: Should Central Banks ‘Lean Against The Wind’ to Foster Macro-financial Stability? *Economics: The Open-Access, Open-Assessment E-Journal*, 12 (2018–7): 1–69.

Minsky, H.P. (1982). Can “it” Happen Again? Essays on Instability and Finance, Armonk, N.Y.: M.E. Sharpe.

Mumtaz, A. and Zanetti, F. (2016). The Effect of Labour and Financial Frictions on Aggregate Fluctuations, *Macroeconomic Dynamics*, 20: 313–341.

Nelson, C.R. (1988). Spurious Trend and Cycle in the State Space Decomposition of a Time Series with a Unit Root, *Journal of Economic Dynamics and Control*, 12 (2–3): 475–488.

Pesaran, M.H. and Timmermann, A. (2007). Selection of Estimation Window in the Presence of Breaks, *Journal of Econometrics*, 137: 134–161.

Popoyan, L., Napoletano, M. and Roventini, A. (2017). Taming Macroeconomic Instability: Monetary and Macro Prudential Policy Interactions in an Agent-Based Model, *Journal of Economic Behavior & Organization*, 134: 117–140.

von Peter, G. (2005). Debt-deflation: Concepts and a Stylised model, BIS Working Papers 176.

APPENDIX

Table A1. Parameters of the model

<u>Description</u>	<u>Parameter</u>	<u>Value</u>
Market entry		
Productivity growth variation	λ_e	0.0015
Target market share variation	λ_z	0.01
Set-up cost	Ω (new shop)	40
	Ω (shop upgrade)	20
Horizon for project assessment	h	10
Risk taking	α_n	$\in (1,2)$
Shop targets adjustment		
Output (shop-specific inventories ratio)	ξ^{Y1}	0.1
Output (global inventories ratio)	ξ^{Y2}	0.1
Output (output gap)	ξ^{Y3}	0.3
Wage (labour demand)	ξ^{W1}	1
Wage (average wage)	ξ^{W2}	0.2
Price (shop-specific inventories ratio)	ξ^{P1}	0.0175
Price (global inventories ratio)	ξ^{P2} (consumption goods industry)	0.07
	ξ^{P2} (capital goods industry)	0.035
Price (average price)	ξ^{P3}	0.1
Target inventories-to-sales ratio	Λ^* (consumption goods industry)	1
	Λ^* (capital goods industry)	0.75
Consumption elasticities		
Income	β_n^I	$\max [0, \sim N(0.6, 0.05)]$
Wealth	β_n^W	$\max [0, \sim N(0.15, 0.01)]$
Debt	β_n^D	$\max [0, \sim N(0.2, 0.01)]$
Random	$\epsilon_{n,t}$	$\max [0, \sim N(0, 0.025)]$

Credit market

Threshold debt service ratio (starting and maximum values)	DSR* DSR* ^{max}	0.5 0.75
DSR threshold drift	v_t	$\sim N(0.005, 0.00025)$
Loan demand factor	Φ (investment loan)	1.1
	Φ (current expenses loan)	1.25
Loan maturity	M	10
Number of missed payments to default	D	9
Deposit spread	DS	3
Credit spread	CS	3

Monetary policy

Reaction to inflation	σ^I	0.6
Inflation target	π^{trg}	5
Reaction to unemployment	σ^U	0.5
Unemployment target	u^{trg}	5
Monetary policy rule intercept	IR^*	6

Other parameters

Trend inertia	λ_a	0.95
Maximum workers per shop	L^{max}	7
Required capital input	λ_c (consumption goods industry)	0.5
	λ_c (capital goods industry)	0
Fire sales discount	Θ	0.75
Number of shops that participate in fire sales	Z	10
Number of agents	N	300

Probabilities

Technology upgrade and new shop set-up	p^e (new shop)	0.5
	p^e (new shop)	1
Shop change	p^c	0.5
Job skills learning	p^{s1}	0.125
	p^{s2}	0.05
Credit availability	Investment loan	
	p_t^L (starting value)	0.4
	p^{Lmax} (maximum value)	0.65

	Current expenses loan	
	p_t^L (starting value)	0.132
	p_t^{Lmax} (maximum value)	0.215
Credit availability drift	μ_t (investment loan)	$\sim N(0.005, 0.00025)$
	μ_t (current expenses loan)	$\sim N(0.00135, 0.0000825)$
Firing unwanted employees	p^F	0.75

Table A2. Empirical data set

<i>Country</i>	<i>Time sample</i>	
	<i>From</i>	<i>To</i>
Australia	Q1 1960	Q3 2016
Austria	Q1 1960	Q4 2016
Belgium	Q3 1970	Q3 2016
Canada	Q1 1960	Q3 2016
Chile	Q1 1995	Q3 2016
Colombia	Q1 2000	Q3 2016
Denmark	Q1 1967	Q3 2016
Finland	Q3 1970	Q3 2016
France	Q3 1969	Q3 2016
Germany	Q1 1960	Q3 2016
Greece	Q1 1960	Q3 2016
Ireland	Q2 1976	Q3 2016
Israel	Q1 1995	Q3 2016
Italy	Q1 1960	Q3 2016
Japan	Q3 1964	Q3 2016
Korea	Q3 1962	Q3 2016
Mexico	Q3 1980	Q3 2016
Netherlands	Q4 1960	Q3 2016
New Zealand	Q2 1987	Q3 2016
Norway	Q1 1960	Q3 2016
Portugal	Q1 1960	Q3 2016
South Africa	Q4 1964	Q3 2016
Spain	Q4 1969	Q3 2016
Sweden	Q4 1960	Q3 2016
Switzerland	Q1 1960	Q3 2016
United Kingdom	Q4 1962	Q3 2016
United States	Q4 1951	Q3 2016

We use BIS database as the source for credit series (adjusted for breaks all sectors' credit to private non-financial sector). The availability of these data determines the composition of the data set. We use OECD database for GDP and price (GDP deflator if available and consumer prices otherwise) series. All data are seasonally adjusted using X-12 procedure.

Table A3. Characteristics of actual and artificial credit cycles (means of the distribution)

Data	Amplitude (%)		Duration (quarters)		
	<i>Expansion</i>	<i>Contraction</i>	<i>Expansion</i>	<i>Contraction</i>	<i>Cycle</i>
Empirical	79	-17	61	21	75
ABM-based	51	-33	61	16	70
VAR-based	520	-50	100	12	119

Amplitude measure indicates the percentage change from trough to peak (expansion) or peak to trough (contraction). The duration of the full cycle is measured fro