BOFIT Discussion Papers
6 • 2019

Alexander Kostrov and Mikhail Mamonov

The formation of hidden negative capital in banking: A product mismatch hypothesis

THE BANK OF FINLAND
INSTITUTE FOR ECONOMIES
IN TRANSITION

BOFIT Discussion Papers
Editor-in-Chief Zuzana Fungáčová

BOFIT Discussion Papers 6/2019
3.4.2019

Alexander Kostrov and Mikhail Mamonov: The formation of hidden negative capital in banking: A product mismatch hypothesis

ISBN 978-952-323-268-6, online
ISSN 1456-5889, online

Suomen Pankki
Helsinki 2019

The Formation of Hidden Negative Capital in Banking: A Product Mismatch Hypothesis

Alexander Kostrov[*], Mikhail Mamonov[†]

This draft: April 2019

Abstract

This paper investigates the phenomenon of hidden negative capital (HNC) associated with bank failures and introduces a product mismatch hypothesis to explain the formation of HNC. Given that troubled banks tend to hide negative capital in financial statements from regulators to keep their licenses, we attempt to capture this gambling behavior by evaluating product mismatches reflecting disproportions between the allocation of bank assets and the sources of funding. We manually collect unique data on HNC and test our hypothesis using U.S. and Russian banking statistics for the 2004–2017 period (external validity argument). To manage the sample selection concerns, we apply the Heckman selection approach. Our results clearly indicate that product mismatch matters and works similarly in both U.S. and Russian banking systems. Specifically, an increase in mismatch has two effects: it leads to a higher probability that a bank's capital is negative and raises the conditional size of the bank's HNC. Further, we demonstrate that the mismatch effect is heterogeneous with respect to bank size being at least partially consistent with the informational asymmetry view. Our results may facilitate improvements in the prudential regulation of banking activities in other countries that share similar features with either the U.S. or Russian banking systems.

Keywords: Bank failure, Hidden negative capital, Product mismatch, Misreporting, Heckman selection model.

JEL: G21 G33 C34

[*]Chair of Mathematics and Statistics, University of St. Gallen, Bodanstrasse 6, 9000 St Gallen, Switzerland, +41 71 224 2475, alexander.kostrov@unisg.ch, ORCID ID: https://orcid.org/0000-0002-1872-7887.

[†]Corresponding author. CERGE-EI, a joint workplace of Charles University and the Economics Institute of the Czech Academy of Sciences, Politickych veznu 7, 111 21 Prague, Czech Republic, +420 776 071 490, mikhail.mamonov@cerge-ei.cz. ORCID ID: https://orcid.org/0000-0002-4256-4838.

1 Introduction

Bank capital has to satisfy the official regulatory requirements in all times. However, during periods of financial turbulence, the bank capital adequacy ratio can not only fall below the regulatory thresholds but, depending on the size of shocks, even further below zero thus rendering bank capital negative. In the latter case, troubled banks have huge incentives to falsify their financial statements by reporting artificially sufficient positive levels of capital to keep their licences, thus engaging in fraud and insider abuse (James, 1991; Kang et al., 2015; Cole and White, 2017).[1] These banks essentially operate with hidden negative capital (hereinafter, HNC) which is likely to bear certain — and possibly very large — losses to society. Therefore, it is of crucial importance to keep track of HNC and be able to prevent its further formation in the banking system. In this paper, we investigate the process of HNC formation in banks that are in financial trouble. We develop an empirical setting that allows us to distinguish between banks that operate with true positive capital and banks that operate with falsified positive capital, i.e. with HNC.

HNC can be treated as a consequence of either incorrect business decisions of bank managers (bad luck in the spirit of Berger and DeYoung, 1997) or banks' illegal activities and falsifications. Previous research has identified several signals that are associated with negative capital in banks (James, 1991; Kang et al., 2015): a lack of capital[2], higher portion of non-performing loans, and income earned but not collected, among others (see Section 2 for further details). However, negative capital can also be associated with either product, risk, or liquidity mismatches, though, to the best of our knowledge, neither has been examined in this context. In this paper, in investigating the process of HNC formation, we focus specifically on *product mismatch*, leaving other types of mismatch (liquidity, risk, currency, and others) for future work. Product mismatch, as we explain it further, can be viewed as a more broader concept which encompasses other types of mismatch. To start the work, we need to show empirically whether product mismatch matters for HNC in general and, if so, identify the concrete channels through which it affects HNC in future research. We treat product mismatch as a mismatch between the sources of funding and the allocations of assets. For instance, household deposits may appear to be more expensive compared to corporate deposits (the liabilities side) and, normally, credit to households is likely to be charged with a higher interest rate than

[1] By overreporting its capital, a troubled bank may hope that the situation will later improve thus allowing the bank to turn back to fair capital reporting. This can be possible if, during the financial turbulence, either the too-many-to-fail effect materializes (Acharya and Yorulmazer, 2007; Brown and Dinç, 2011) or the central banks decide to postpone the costs of bank failures to future due to, for instance, high monetary and/or non-monetary costs associated with the closure of particular troubled banks in the current period (the regulatory forbearance effect, see, e.g., Kang et al., 2015).

[2] One can imagine a situation in which a bank reports sufficient positive capital but the ratio of capital to risk-weighted assets (i) is just slightly above the regulatory thresholds and (ii) remains there for a relatively long time.

credit to firms (the assets side). Thus, proper matching of bank assets and liabilities with respect to the types of clients may be crucial for banks' profitability and, eventually, for stable bank performance in the long run. We anticipate that product mismatch plays an additional role in the process of HNC formation alongside the forces already analyzed in the literature.

There are many reasons why product mismatch may occur. A typical example is so-called pocket banks: Owners of non-financial businesses create a bank to finance their projects at an interest rate lower than the market would imply. Since corporate depositors pay greater attention to the stability of the banks they choose than private depositors do[3], pocket banks usually rely on household funds. Thus, mismatching follows immediately. Apart from the pocket bank creation, other driving forces that may push banks to pursue a mismatching strategy include, among others, a toughening of competition in respective markets for banking services and dumping strategies. Abstracting from the exact reason for a mismatch (since it is rather hard to identify at the bank level), in this study we are specifically interested in considering those banks that rely primarily on (more expensive) household deposits and, at the same time, on (less profitable) loans to non-financial firms. Therefore, in our empirical setting we test whether the product mismatch leads to banks' capital depletion resulting in the emergence of HNC (average treatment effect) and whether there are any differences in the formation of HNC between larger and smaller financial institutions (heterogeneous treatment effect) consistent with the informational asymmetry view.[4]

Examining the formation of HNC in banking and determining whether product mismatch explains the formation are important and non-trivial issues because banks with HNC can survive for a relatively long time. Banks can exist with HNC for at least two reasons. First, the theory of bank capital states that a bank can operate with HNC as long as it is able to maintain the confidence of its creditors. Second, banking systems are also subject to excessive regulatory forbearance (Wheelock and Wilson, 2000; Kang et al., 2015; Cole and White, 2017), meaning that regulators are unable to process license revocation faster than they do. This is because of either (i) informational asymmetry and possible political pressure (Brown and Dinç, 2005; Kang et al., 2015), (ii) a government budget deficit when the banking system is weak (Brown and Dinç, 2011), or (iii) the risk of missing the opportunity to sell failing banks to healthier banks (Bennett and Unal, 2014; Granja et al., 2017). However, a bank cannot operate with HNC forever: A recent study by Berger and Bouwman (2013) shows that bank capital matters for retaining

[3]Among the latter, private deposit holders are likely to be less aware of banking problems due to informational asymmetry (Diamond and Rajan, 2000) and the rational inattention argument.

[4]Larger banks have more options to diversify their assets and liabilities than smaller banks because of greater confidence and transparency (see, for instance, Kashyap and Stein, 2000), meaning that the effect of mismatch may be lower compared to smaller banks or may even not exist. Larger banks may be more willing to eliminate the possibility of HNC formation than smaller banks because they have higher costs of license withdrawal (wastage of greater goodwill, larger branching networks, etc).

market shares and for sustaining the stability of banking services provision (for smaller banks this holds across all phases of business cycles; for larger banks during periods of crisis). Therefore, either the bank's capital must become positive again (after the shock dies out) or the bank will eventually lose the market and face losses, thus providing a strong signal of its problems to the central bank. Such concerns are reinforced by the credit cycles mechanism of Kiyotaki and Moore (1997): Having experienced even a small temporary financial shock, affected banks and their borrowers face a large intertemporal deterioration of their respective net worth through tightened credit constraints. Thus, these banks may simply lose their market shares and the ability to generate profit before the shock starts to die out and the banks' net worth starts to recover. Overall, we can anticipate that a self-disappearing HNC is unlikely to be a very frequent event in practice.

In addition, bank failures *per se* are associated with certain costs to society, either monetary or non-monetary (James, 1991; Kang et al., 2015; Cole and White, 2017). In the wake of systemic banking crises, particularly after the Great Recession, these costs, as well as the question of banks' survival, have generated increased concerns of policymakers and academics, because monetary costs are primarily costs to the government budget and the deposit insurance system.[5] Non-monetary costs originate through decreased availability of bank credit (Kang et al., 2015) and take the form of a reduced economic activity in particular regions (Aschcraft, 2005) and at the firm level (Chodorow-Reich, 2014; Gropp et al., 2018).

It is challenging to gather bank-level data on HNC even for a small number of countries. Indeed, Bankscope, the most common source for cross-country banking studies, does not provide this information. Financial regulators have to bear certain reputational costs of publishing these data because the existence of HNC *per se* may question the validity of the central bank's prudential policy (theoretically, valid prudential regulation should exclude HNC) and, in addition, may undermine the confidence in operating banks (some of these banks may be those with not yet revealed HNC). Fortunately, we have discovered two notable exceptions: the Federal Deposit Insurance Corporation (FDIC) in the United States and the Central Bank of Russia in the Russian Federation do disclose the size of HNC in failed banks, starting from the 1980s and 2007, respectively. A comparison of HNC formation in these two countries is interesting and informative of the external validity of our empirical setting because the U.S. banking system is market-based and global, whereas the Russian system is bank-based and local. We thus seek to identify whether there are similar underlying forces — including product mismatch — in these two very different banking systems. A positive answer could allow one to generalize our

[5]Though government support of banks is deeply unpopular, it may be necessary for a faster macroeconomic recovery. For instance, a recent cross-country study by Homar and van Wijnbergen (2017) concludes that timely bank recapitalization undertaken by governments leads to a significant reduction in the duration of macroeconomic recessions and decreases the costs associated with regulatory forbearance.

finding.[6]

The U.S. banking system is subject to a rather large problem of bank misreporting and, hence, HNC formation, as was shown by James (1991) for the banking crisis of the 1980s, and more recently by Kang et al. (2015), Cole and White (2017) and Balla et al. (2015) for the banking crisis of the late 2000s. Specifically, Balla et al. (2015) estimate that among all U.S. banks with licenses revoked during the 2007 to 2013 period as many as 403 financial institutions had HNC: They reported a pre-failure capital-to-assets ratio of +1.5% on average, whereas after respective failures the FDIC refined this figure to −24%. Regarding the Russian banking system, we reveal that it lost about 550 banks during the 2007 to 2017 period (half of the pre-2007 quantity of operating banks), and that those banks failing with HNC had an average capital-to-assets ratio of +17% prior to failure and −51% after. This rough comparison shows that, despite apparent differences, the situations in both banking systems may be qualitatively similar, though quantitatively more dramatic in the Russian case.

Our study contributes to the literature on bank failures in several respects.

First, we hand-collect unique data on HNC and provide the first cross-country evidence on HNC formation. Specifically, we formulate and estimate the same regression model of HNC for each of the two countries. Since we can observe HNC only in those banks that have already failed, the estimation procedure is subject to the sample selection concerns. Thus, we employ the Heckman selection approach (Heckman, 1979) for our regression analysis.

Second, we introduce product mismatch into the Heckman selection framework, i.e. in the models of (i) the probability of a bank failure with HNC (selection equation) and (ii) the conditional size of HNC (outcome equation) alongside the standard set of explanatory variables applied in previous research (i.e., equity-to-assets ratio, bank profitability, non-performing loans ratio, assets growth rate, and liquidity ratios). The official statistics on the products' deposit interest rates and the returns on loans for both the United States and Russia — even at the aggregate level — clearly support the idea of a reduced interest rate margin of the product mismatch strategy (Table 1). This explains why we further employ the product mismatch at the bank level and test whether funding with household deposits combined with granting corporate loans is associated with both a higher probability that HNC exists and with a larger size of HNC.

Third, for both countries, we analyze the heterogeneity of the product mismatch effect on HNC. Specifically, we split our samples into two asset classes, smaller and larger banks, and consider the transmission of mismatch in each class. This allows us to study whether and how the heterogeneity of the mismatch effect differs between U.S. and Russian banks.

In essence, our estimation results suggest that the Heckman selection approach is

[6] Generalization of results in this respect has become increasingly important since the Great Recession revealed a substantial depletion of bank capital in many countries around the globe (McKinsey, 2010).

Table 1: Data on banking interest rates in the United States and Russia

	United States	Russia
Panel 1: Deposit rates (annual), %		
Households	1.2	7.8
Firms	0.5	7.3
Panel 2: Returns on loans (annual), %		
Households	8.5	16.4
Firms	2.5	10.6

Note: The table contains averaged values for the United States and Russia during the 2007–2016 period.

effective in describing and forecasting the HNC formation for both U.S. and Russian banks. We show that there is indeed a common pattern in the HNC formation in these two countries and demonstrate that the mismatch variable is a valid determinant of HNC for both banking sectors: It increases both the probability of bank failures with HNC and the conditional size of HNC. Finally, the mismatch effect is heterogeneous with respect to the size of assets and this heterogeneity works differently in the two countries. In the United States, the transmission of mismatch takes place through smaller banks only and not through larger competitors, consistent with the informational asymmetry view. Somewhat differently, in Russia, mismatching increases the probability of HNC formation in both smaller and larger banks, but — and similarly to the U.S. case — mismatching rises the conditional size of HNC for the smaller banks only. Thus, the results for Russia are partially consistent with the informational asymmetry view.

The remainder of this paper is organized as follows. In Section 2, we briefly describe the literature on negative capital and discuss its relation to research on bank failures. Our empirical strategy is introduced in Section 3, which also contains the data description. The estimation results for both countries are presented in Section 4. Section 5 confirms the robustness of our findings. The final section provides concluding remarks.

2 Literature Review

In this section, we first relate the research on bank failures with that on negative capital in banking (Section 2.1) and then proceed to the description of the determinants of negative capital identified in previous studies (Section 2.2).

2.1 General remarks: Bank failures and hidden negative capital

The literature on financial stability has not paid much attention to the problem of banks' HNC so far and, thus, to the determinants of HNC formation. Previous studies have tended to focus on bank failures *per se*, motivating their analysis by the arguments that bank failures destruct relationship lending, reduce the total supply of credit to the economy (Aschcraft, 2005), and increase GDP losses associated with them (Boyd, Kwak and Smith, 2005). However, a systematic attempt to predict both bank failures and the size of negative capital has not been undertaken.

In most cases, the focus of previous studies on bank failures has been to achieve an accurate prediction of the episodes of banking license withdrawals given available data. The research on bank failures is much richer than that on HNC. Moreover, the latter covers only the U.S. banking system, to our knowledge, while the former encompasses many countries around the globe.[7]

2.2 The determinants of negative capital

In the influential paper by James (1991), negative capital is measured as the book value of assets (at the moment of closure) minus the value of assets in the FDIC or acquirer receivership. The study proposes an *ad hoc* linear regression of the FDIC losses on failing banks with 11 explanatory variables such as book value of equity, core deposits, and several types of non-performing assets. By applying OLS, James (1991) obtains the surprising result that the pre-failure equity capital, i.e. the one reported by a failing bank in its balance sheet, was positively associated with the ex-post size of negative capital revealed by the FDIC after the bank failed. In attempts to clarify this finding, he attributes it to fraud and insider abuse that are more prevalent in the better-capitalized banks among those that failed. From the technical standpoint, the proposed regression model considers only failure cases, which may raise concerns due to possible sample selection bias that is likely to have an upward influence on the estimated relationships.

The research following James (1991) has developed the analysis of negative capital in many directions, e.g. the examination of regulatory forbearance effects (Brown and Dinç, 2011; Kang et al., 2015; Cole and White, 2017, etc.), understanding the real effects of failing banks (Bennett and Unal, 2014; Granja et al., 2017), analyzing the differences

[7]Indeed, one can find numerous studies on the probability of bank failures dealing with U.S. banks (Cole and White, 2012; DeYoung and Torna, 2013; Cleary and Hebb, 2016; Audrino et al., forthcoming), to name the most recent ones, banks in developing and emerging economies (Karminsky and Kostrov, 2017; Brown and Dinç, 2011; Arena, 2008; Mannasoo and Mayes, 2010; Fungacova and Weill, 2013, among others), and even EU banks (Poghosyan and Cihak, 2011; Betz et al., 2014). Conversely, the literature on HNC is biased towards the United States and appears during and just after the crisis periods in the late 1980s to the early 1990s (Bovenzi and Murton, 1988; James, 1991; Osterberg and Thomson, 1995) and after the Great Recession (Shaeck, 2008; Bennett and Unal, 2014; Kang et al., 2015; Balla et al., 2015; Cole and White, 2017; Granja et al., 2017).

between high-cost and low-cost failures (Shaeck, 2008), and addressing the sample selection bias concerns (Balla et al., 2015). While considering all these studies in detail is beyond the scope of this paper, we have to acknowledge that they exploit essentially the same determinants of negative capital that can be extracted from U.S. banks' financial reporting.

From studies on regulatory forbearance, we can borrow a number of determinants of negative capital beyond those used by James (1991). Osterberg and Thomson (1995) bring the off-balance-sheet variables into the analysis and demonstrate significant additional effects of loan commitments, letters of credit, and derivative securities. Specifically, the authors find that loan commitments and letters of credit decrease negative capital, which is consistent with the market discipline view. Derivative securities also have the negative effect as expected since this item is likely to be used to hedge the on-balance-sheet risk. Osterberg and Thomson (1995) agree with James (1991) that fraud is a significant reason for bank failure with negative capital; however, unlike in James (1991), they provide strong evidence of the negative, rather than positive, effect of the pre-failure capital adequacy on the ex-post negative capital.

Further, research on the acquirers of failing banks shows that more comprehensive regulatory disclosure requirements lead to lower resolution costs of failed banks and, by reducing the informational asymmetry, increase the bidders' incentives on acquiring banks (Granja et al., 2017). Thus, we can anticipate that in banking systems with stricter disclosure requirements the ex-post revealed HNC will be smaller than in informationally opaque systems. Further, Bennett and Unal (2014) find that failed banks with more branches tend to be acquired rather than liquidated.

Research on the differences between high- and low-cost failures brings another set of relevant predictors for HNC. Shaeck (2008) emphasizes that regulators are mainly concerned with expensive failures and that factors driving high- and low-cost failures may be different. Using the sample of U.S. banks that failed in 1984-2003, though not addressing the sample selection concerns, Shaeck (2008) shows that a higher reliance on Fed funds is associated with less costly bank failures. Conversely, the usage of brokered deposits, poor asset quality, uncollected income, and a weak macroeconomic environment increase the cost of bank failures. Note that Shaeck (2008) is the first study to incorporate liability structure into the empirical models of negative capital, a fact that we will also exploit in our estimations.

Finally, Balla et al. (2015) make a first attempt to address the sample selection concern by analyzing both failed and surviving banks. The sample selection bias appears since we can observe negative capital only in those banks that already failed, while some of the existing banks have *de facto* failed but continue operating. The authors apply the Heckman selection model to identify the determinants of negative capital. Specifically, they estimate the selection equation, i.e. the probability that negative capital exists, and the outcome equation, i.e. the size of negative capital conditional on being selected. Importantly, Balla et al. (2015) show that the correlation between the selection and outcome regression errors is statistically significant, thus confirming the existence of the sample selection bias. The list of the determinants of negative capital is essentially the same as in previous research.

3 Model and data

In this section, we describe the main steps of our empirical strategy, introduce the baseline product mismatch hypothesis (H1) and the heterogeneous mismatch hypothesis (H2), and then discuss the data we use on U.S. and Russian banks.

3.1 Empirical strategy: Heckman selection models

We employ the Heckman selection approach (Heckman, 1979) to test the product mismatch hypothesis and thus to uncover the role of banks' mismatching behavior in the formation of HNC at the bank level.

The Heckman selection model of HNC formation consists of selection and outcome equations that describe the probability and the conditional size of HNC, respectively. Both equations contain the set of basic bank-specific determinants stemming from the literature (*Basic*), our proposed product mismatch variable (*Mismatch*). In addition, the selection equation employs the bank size (*Size*) as an identification variable. The bank size is an appropriate identifying variable because, in all outcome regressions that follow, we consider the relative, not absolute, size of HNC as the dependent variable; that is, we normalize the absolute value of HNC with the total liabilities for each bank in our U.S. and Russian samples. The point is that the bank size variable does affect HNC at the selection stage (through the usual too-big-to-fail argument, O'Hara and Shaw, 1990) but does not affect (relative) HNC at the selection stage of the Heckman model. Indeed, in both samples there is no clear pattern of the relationship between relative HNC and bank size (no statistical correlation): There are large failed banks with small relative HNC and small failed banks with large HNC, and vice versa.

The resultant specification of the Heckman selection model that we use for U.S. and for Russian banks takes the following form:

$$D_{i,t} = \alpha_1 + \sum_{i=1}^{n} \beta_1 \mathrm{BASIC}_{i,t-k} + \gamma_1 \mathrm{Mismatch}_{i,t-k} + \delta_1 \mathrm{Size}_{i,t-k} + \epsilon_{1i,t}, \quad (1)$$

$$\frac{\mathrm{HNC}_{i,t}}{\mathrm{Liabilities}_{i,t}} = \alpha_2 + \sum_{i=1}^{n} \beta_2 \mathrm{BASIC}_{i,t-k} + \gamma_2 \mathrm{Mismatch}_{i,t-k} + \epsilon_{2i,t}, \quad (2)$$

where $\mathrm{HNC}_{i,t}$ is the absolute size of hidden negative capital of bank i at time t uncovered by the regulator (zero for operating banks). $\text{Liabilities}_{i,t}$ are the value of total liabilities officially reported in financial statements. $D_{i,t}$ is the probability that a bank has negative capital (a latent variable); in estimation, it takes value "1" if a bank fails with HNC in period t and "0" otherwise. $\text{Size}_{i,t}$ is the natural logarithm of bank total assets, the identifying variable. α, β, γ, and δ are the coefficients to be estimated. ϵ_1 and ϵ_2 are the error terms in the selection equation (1) and the outcome equation (2), respectively. We fix k to be 1 quarter in our main regression analysis and report the results for larger forecasting horizons in the Appendices.

The set of $\text{Basic}_{i,t}$ predictors includes:

- Capital — the ratio of bank equity capital to total assets.
- NPL — the ratio of non-performing loans to total loans.
- Liquidity — the share of state bonds and cash holdings in total assets.
- ROA — the return on assets.
- TA growth — the annual growth rate of total assets.

Finally, our main variable of interest, Mismatch_{it}, reflects the product mismatch in a bank's i business model: funding with (relatively expensive) deposits from households and granting (relative cheaper) loans to non-financial firms. We define mismatch for bank i at time t as the product of the household deposits to liabilities ratio ($\text{DepH}_{i,t}$) and corporate loans to assets ratio ($\text{LnsF}_{i,t}$):

$$\text{Mismatch}_{i,t} = \text{LnsF}_{i,t} \times \text{DepH}_{i,t}, \tag{3}$$

that is, we say that the product mismatch in bank i balance sheet increases if either $\text{LnsF}_{i,t}$ rises (holding $\text{DepH}_{i,t}$ fixed), $\text{DepH}_{i,t}$ rises (holding $\text{LnsF}_{i,t}$ fixed), both components rise simultaneously, or one of the components rises by more than the other falls.

H1: *The product mismatch hypothesis* states that the probability of HNC formation and (conditional on being formed) the size of HNC are higher for banks with a greater product mismatch. Thus, H1 is accepted if (i) both $\gamma_1 > 0$ and $\gamma_2 > 0$ and (ii) γ_1 and γ_2 are statistically significant, meaning that larger product mismatch increases both the probability and the conditional size of HNC in bank i. The hypothesis is partially accepted if at least one of the two coefficients is positive and statistically significant. In all other cases the hypothesis is rejected.

Importantly, in our definition (3), product mismatch rarely takes zero values. Exceptions are those banks with either $\text{LnsF}_{i,t} = 0$ (i.e., full specialization on lending to non-financial firms) or $\text{DepH}_{i,t} = 0$ (i.e., full specialization on households deposits). We acknowledge that, up to some degree, product mismatch may not be dangerous for sustaining bank stability. However, we estimate equations (1) and (2) as a sequence of cross-sectional regressions at each date t and thus we effectively trace the differences in the mismatch effect *across* banks, not *within* banks; therefore, we assume that bank i is going to be less stable than bank j if bank i is characterized with higher values of mismatch, even if the values of mismatch for both banks are relatively small (say, below the banking system's average value at the respective date). At the same time, we — at least, partially — account for this concern by further formulating a modification to our main hypothesis — the heterogeneous mismatch hypothesis.

H2: *The heterogeneous mismatch hypothesis* states that larger banks should be less affected by the product mismatches than smaller competitors because they have more opportunities to diversify their assets and liabilities (the information asymmetry view). To test this hypothesis, we then divide each of our samples of U.S. and Russian banks into two parts, respectively, by the size of banks' assets: 5% of banks with the greatest value of assets are classified as large, the remaining 95% of banks are assigned to be small. Thus, the definition of large banks coincides with that used, for instance, in Kashyap and Stein (2000), and our definition of small banks encompasses their small and medium banks. We do not distinguish between small and medium banks because we are interested in capturing the difference between large and all the other banks, which is sufficient at this stage of analysis. We augment the baseline version of the Heckman selection model (1)–(2) with the cross-products between size dummies (denoted $\text{Large}_{i,t}$ and $\text{Small}_{i,t}$) and the product mismatch variable:

$$\begin{aligned} D_{i,t} = \alpha_1 + \sum_{i=1}^{n} \beta_i \text{BASIC}_{i,t-k} + \gamma_1 \text{Mismatch}_{i,t-k} + \theta_{11} \text{Mismatch}_{i,t-k} \text{Large}_{i,t-k} \\ + \theta_{12} \text{Mismatch}_{i,t-k} \text{Small}_{i,t-k} + \delta_1 \text{Size}_{i,t-k} + \epsilon_{1i,t}, \end{aligned} \tag{4}$$

$$\begin{aligned} \frac{\text{HNC}_{i,t}}{\text{Liabilities}_{i,t}} = \alpha_2 + \sum_{i=1}^{n} \beta_i \text{BASIC}_{i,t-k} + \gamma_2 \text{Mismatch}_{i,t-k} + \theta_{21} \text{Mismatch}_{i,t-k} \text{Large}_{i,t-k} \\ + \theta_{22} \text{Mismatch}_{i,t-k} \text{Small}_{i,t-k} + \epsilon_{2i,t}. \end{aligned} \tag{5}$$

We do not include the dummy variable for either large banks or small banks itself in the outcome equation (5) because it would contradict with the idea of bank size being the identifying variable of the Heckman selection model of bank HNC. In addition, in Appendix A we show that the correlation between relative HNC and the dummy variable for large banks is negligible.

The heterogeneous mismatch hypothesis is accepted if (i) $\theta_{j1} < \theta_{j2}$ and (ii) both θ_{j1} and θ_{j2} are statistically significant for any $j = 1, 2$.[8] In other words, the effect of mismatch

[8] In particular, one possible scenario is that θ_{j1} is statistically zero and $\theta_{j2} > 0$ statistically for any $j = 1, 2$.

on large banks must be lower than that on the small banks (informational asymmetry view). Further, the heterogeneous hypothesis is partially accepted if the same effects work in only one of the two equations $j = 1$ or $j = 2$.

We estimate each pair of selection and outcome equations (1)–(2) or (4)–(5) simultaneously by the maximum likelihood (efficient two-step estimates appear in the robustness check, see Section 5). The sample selection bias is said to be present in the data and captured by the model if $Corr(\epsilon_1, \epsilon_2) \neq 0$ statistically, i.e. there is (are) common unobserved force(s) that may affect both selection and outcome equations even after controlling for all available observable characteristics. We test it by applying the likelihood ratio test with the null hypothesis of no correlation between the selection and outcome regression errors.

3.2 Data

In this section, we describe our bank-level data on HNC and on the determinants of HNC for the U.S. banking sector and then for the Russian banking sector.

We start with the notion of data cleaning. We apply the same cleaning procedure to both U.S. and Russian data samples to address, in the same manner, possible outliers among failed and operating banks. Specifically, we remove (i) observations outside the 5^{th} and 95^{th} percentiles for the return-on-assets (ROA), (ii) observations outside the 1^{st} and 95^{th} percentiles for the equity ratio (Capital), (iii) 5% of observation with the highest growth rate of total assets (TA growth), as suggested by Kang et al. (2015) to manage the M&A cases. In the following sections, we present the data on the cleaned samples only.

3.2.1 Data on U.S. banks

Within the period from 2004 to 2016Q2, we identified 525 bank failure cases with corresponding negative capital documented in the FDIC's Failed Bank list. After the data cleaning procedure we are left with 504 cases. In Table 2, we present the descriptive statistics on the relative size of HNC in these cases.

As can be inferred from Table 2, HNC is larger when it is measured as a portion of remaining total assets: Total liabilities exceed total assets when capital is negative. Most banks disclose near zero capital shortly before they fail. In general, we observe a very large variation of the relative size of HNC: from near zero values to as much as three quarters of remaining assets and more than 40% of remaining liabilities (with the average value equals one quarter of assets, or almost 20% of liabilities).

Further, we collect the bank-level data on the determinants of HNC from the Call Reports on FDIC-insured banks for the same period as for the HNC (i.e. from 2004 to 2016Q2). These reports are disseminated by the FDIC in quarterly sets of files with de-

Table 2: Descriptive statistics for 504 failed U.S. banks:
The relative size of hidden negative capital (HNC), *2004Q1–2016Q2*

Size of HNC	Obs	Mean	SD	Min	Max	Percentiles		
						$Q_{25\%}$	$Q_{50\%}$	$Q_{75\%}$
fraction of remaining assets	504	0.24	0.13	0.00	0.75	0.14	0.23	0.34
fraction of remaining liabilities	504	0.19	0.09	0.00	0.43	0.12	0.19	0.25

Note: By definition, HNC takes only negative values. For the sake of simplicity, we transform them to positive values and then divide by either the size of assets or liabilities remaining after the closure of a bank by the FDIC.

tailed banking statistics. Our sample thus includes 50 quarters of data for 9,936 unique U.S. banks before the data cleaning procedure. After removing outliers, we have approximately 343–382 thousand bank-quarter observations. In Table 3 we compare the descriptive statistics for failed and operating banks. There are two panels in the table: the first contains the determinants of HNC that were used in previous research (basic set) and the second contains the components of our mismatch variable (additional set).

Table 3: Descriptive statistics for U.S. banks:
Failed *(2004–2016Q2)* vs. Operating *(2016Q2)*

Indicator	Failed banks					Operating banks				
	Obs	Mean	SD	$Q_{1\%}$	$Q_{99\%}$	Obs	Mean	SD	$Q_{1\%}$	$Q_{99\%}$
Panel 1: Basic set										
Capital	504	0.01	0.03	–0.09	0.09	1867	0.11	0.02	0.07	0.20
NPL	504	0.16	0.10	0.00	0.44	1867	0.01	0.02	0.00	0.07
Liquidity	504	0.07	0.07	0.00	0.31	1867	0.11	0.09	0.00	0.38
TA growth	504	–0.12	0.16	–0.46	0.50	1867	0.05	0.07	–0.11	0.23
ROA, %	504	–6.57	6.40	–26.6	1.03	1867	0.95	0.43	0.07	2.00
Size	504	12.4	1.39	9.67	16.36	1867	13.0	1.48	10.5	18.4
Panel 2: Additional set										
LnsF	504	0.08	0.07	0.00	0.34	1867	0.09	0.06	0.00	0.28
DepH	504	0.09	0.13	0.00	0.61	1867	0.04	0.05	0.00	0.20

Note: LnsF is the ratio of all commercial and industry loans to total assets, DepH is the ratio of brokered deposits to total assets.

Comparing failed and operating U.S. banks, one can conclude that, first, the average size of banks in both groups is almost the same and, second, the differences in the values of the determinants of HNC are consistent with expectations. Specifically, failed banks possessed less capital, reported higher NPLs, had lower liquidity, grew much slower, and were subject to substantial losses. Notably, while there are no visible differences between operating and failed banks in terms of lending to non-financial firms, we reveal such differences between them in terms of attracting deposits from households. That is, failed banks attracted these deposits three times more intensively on average than the operating banks; moreover, in the upper tail of the distribution, the differences become even more pronounced.

3.2.2 Data on Russian banks

We compile several sources of statistics on Russian banks for the 2007–2017 period. Monthly balance sheets (Form 101) are publicly disclosed through the official web-site of the Central Bank of Russia.[9] Since 2007, the information on the actual value of assets, liabilities and resultant HNC of failed banks has been publishing in the press releases of the Central Bank of Russia (in the so-called *Vestnik Banka Rossii*). We hand-collect this information for each case for the same 2007–2017 period. Moreover, in some cases the size of HNC is re-evaluated (usually increased) by the Central Bank of Russia. To address this concern, we double-check the data *for each* failed bank in our database in the press-releases that follow the one in which the data were first published; in addition, we re-check the information in media sources.[10]

After the data cleaning procedure, the sample size of Russian operating banks (i.e. those still keeping the license and reporting the Forms) is 925 at the beginning of the sample period and 531 at the end. These banks cover approximately 95% of the total banking system assets. In total, the number of bank-month observations varies between 126-142 thousand. The sample size of failed banks is 359.

In Table 4, we report the descriptive statistics on the HNC of failed Russian banks averaged across the sample period.

Table 4: Descriptive statistics for 359 failed Russian banks:
The relative size of hidden negative capital (HNC), *2007M7–2017M12*

Size of HNC	Obs	Mean	SD	Min	Max	Percentiles		
						$Q_{25\%}$	$Q_{50\%}$	$Q_{75\%}$
fraction of remaining assets	359	3.63	8.84	0.01	105	0.33	0.98	3.27
fraction of remaining liabilities	359	0.31	0.14	0.01	0.50	0.20	0.33	0.43

From Table 4, we can infer that the relative size of the HNC of failed Russian banks is substantially large, being equal to more than one half of their remaining liabilities or, equivalently, more than four times larger than their remaining assets. Comparing these figures to their counterparts in the sample of U.S. failed banks (Table 2), we conclude that

[9] http://cbr.ru/credit/forms/
[10] Business news streaming at https://www.rbc.ru/, one of the largest media sources in Russia).

Table 5: Descriptive statistics for Russian banks:
Failed *(2007M7–2017M12)* vs. Operating *(2017M12)*

Indicator	Failed banks					Operating banks				
	Obs	Mean	SD	$Q_{1\%}$	$Q_{99\%}$	Obs	Mean	SD	$Q_{1\%}$	$Q_{99\%}$
Panel 1: Basic set										
Capital	359	0.18	0.13	0.01	0.60	440	0.21	0.14	0.02	0.68
NPL	359	0.05	0.08	0.00	0.28	440	0.11	0.17	0.00	0.99
Liquidity	359	0.02	0.06	0.00	0.28	440	0.03	0.05	0.00	0.24
TAgrowth	359	0.25	0.80	–0.57	3.97	440	0.02	0.16	–0.39	0.34
ROA, %	359	–0.40	4.13	–16.0	9.53	440	0.25	0.65	–1.42	1.66
Size	359	1.37	1.59	–1.88	5.71	440	2.23	1.96	–0.96	8.02
Panel 2: Additional set										
LnsF	359	0.44	0.24	0.00	0.91	440	0.28	0.18	0.00	0.77
DepH	359	0.39	0.25	0.00	0.82	440	0.32	0.22	0.00	0.72

Note: LnsF – total loans to firms over total assets, DepH – total deposits of households over total assets.

the strength of the banking problems associated with HNC is much stronger in Russia compared to U.S.

Similarly to the U.S. case, we further compare the descriptive statistics on the determinants of HNC for operating versus failed Russian banks. The results appear in Table 5. Several outcomes emerge from the descriptive statistics. First, the failed banks are smaller on average in terms of assets than the operating banks (in contrast to what we have observed in the case of U.S. banks); however, within the failed banks, differences in size are substantial, ranging from very small banks to very large banks. Second, failed banks were much more aggressive in terms of asset growth rate: the average growth rate of assets is four times higher compared to that of the operating banks. Third, the reported quality of total loans to firms and households is surprisingly higher and less volatile for the failed banks compared to the operating banks (again, the opposite was true for U.S. banks).[11] Fourth, in terms of liquidity, failed banks are unsurprisingly worse than the operating banks. Fifth, returns and capital reported are also lower for the failed banks compared to the operating banks. Finally, failed banks report relative losses, though not as large as U.S. failed banks, while operating banks are profitable on average.

The variables we use to test the product mismatch hypothesis also deliver notable results. We document that the failed banks are much more prone to attracting (relatively expensive) deposits from households and granting (relatively cheap) loans to firms than the operating banks. These differences in the behavior of failed and operating banks are in line with those reported for U.S. banks, which is important from the standpoint of the

[11] This delivers clear evidence of balance sheet falsifications and misreporting by Russian failed banks.

external validity of our results.

4 Estimation results

In this section, we present the main estimation results for the Heckman selection model of bank HNC formation (1)–(2) placing the emphasis on the relative importance of product mismatch and estimating the model separately for U.S. and Russian banks. Specifically, we start with testing the (homogeneous) mismatch hypothesis H1 through the Heckman selection framework for U.S. banks (Section 4.1) and then for Russian banks (Section 4.2). Further, we discuss the bank size channel of the mismatch effect on HNC formation: we compare the transmission of the mismatch effect through this channel in U.S. and Russian banks using an augmented specification of the Heckman selection model (4)–(5) (the heterogeneous mismatch hypothesis H2, Section 4.3). Finally, we analyze the results of the out-of-sample forecasting exercises for both banking systems aimed at examining the intertemporal importance of product mismatch in predicting the size of HNC formation (Section 4.4). For the sake of simplicity, we disclose the detailed estimation results mainly for the last periods (2016Q2 for U.S. banks and 2017M12 for Russian banks).

4.1 Hidden negative capital and product mismatch: The case of U.S. banks

The estimation results of the Heckman selection model for U.S. banks are reflected in Table 6. The first two columns contain the benchmark version of the model, as in the previous research on the negative capital of U.S. failed banks. The second pair of columns add the product mismatch variable to the set of HNC determinants. In both models, the selection equation enters the first column and the outcome equation enters the second column. All regressions are estimated on the bank-level data for 2016Q2 (still operating banks) and the historical failure events accumulated from 2004 to 2016Q2 (failed banks) using a one-quarter time lag for all explanatory variables (deeper lags appear in the robustness section).

Benchmark model (without mismatch). The estimation results for the benchmark specification of the Heckman selection model are clearly in line with the previous findings on the cost of U.S. banking failures. That is, first, the equity capital reported prior to failure tends to decrease both the probability of selection and the size of HNC conditional on being selected, in line with, e.g., Kang et al. (2015), Cole and White (2017) and empirical literature stressing the importance of bank capital (Berger and Bouwman, 2013). Second, the opposite holds for non-performing loans (NPLs): the lower the quality of banking loans (reflected in the balance sheets), the higher both the probability and the size of HNC, which again supports the previous findings. Third, liquidity is negatively

Table 6: Heckman selection models for U.S. banks

	Benchmark Model		Final Model	
	Selection	Outcome	Selection	Outcome
Mismatch			74.78*** (22.32)	0.599*** (0.208)
Capital	–61.12*** (11.04)	–0.472*** (0.147)	–62.30*** (10.00)	–0.456*** (0.151)
NPL	32.20*** (5.32)	0.145*** (0.039)	34.49*** (5.36)	0.140*** (0.039)
Liquidity	–0.291 (1.730)	–0.058 (0.052)	–1.758 (2.114)	–0.066 (0.052)
ROA	–1.707*** (0.346)	–0.001** (0.001)	–1.990*** (0.335)	–0.001 (0.001)
TA growth	8.889*** (1.844)	0.082*** (0.024)	8.634*** (1.671)	0.085*** (0.024)
Size	–0.566*** (0.140)		–0.731*** (0.151)	
Constant	9.589*** (2.264)	0.175*** (0.011)	11.58*** (2.36)	0.175*** (0.011)
N obs.	2371	2371	2371	2371
N censored	1867	1867	1867	1867
N observed	504	504	504	504
ρ	–0.342* (0.187)		–0.707*** (0.238)	
Log Likelihood	522.7		532.8	
Convergence	Yes		Yes	

Note: Dependent variable is the probability of HNC formation (in the Selection equation) or the size of HNC conditional on being selected (in the Outcome equation). All explanatory variables are taken with a one-quarter lag. Selection and outcome equations are estimated simultaneously using maximum likelihood (ML). ρ is the correlation between regression errors in selection and outcome equations.

***, **, * indicate that a coefficient is significant at the 1%, 5%, 10% level, respectively. Standard errors are clustered at the bank level and appear in the brackets under the estimated coefficients.

associated with both the probability of selection and the conditional size of HNC, but the effects are not significant in both cases. Fourth, the profitability of banking assets (ROA) decreases both the probability of selection and the conditional size of HNC. This indicates that there is an additional effect of ROA even after controlling for capital.

This may suggest that if profits are higher but capital is lower than expected, higher profits can partially substitute the negative effect of lower capital (owners support banks as long as they receive positive dividends). Fifth, the growth of total assets increases both the probability of selection and the conditional size of HNC, meaning that excessive growth may be harmful for bank stability in future (similar to the findings of Shaeck, 2008). Finally, a bank's size (log of assets) has a negative effect on the probability of being selected, thus supporting the too-big-to-fail view regarding U.S. banks (O'Hara and Shaw, 1990). In our setting, this finding implies that the larger the size of a failed bank was, the more difficult it was for the FDIC to reveal its HNC (larger banks have more instruments to better falsify their accounts). In addition, the size variable does a good job in the identification of the Heckman selection model: when added to the list of the explanatory variables in the outcome equation, it is indeed insignificant at any conventional levels (not reported).

Final model (with mismatch). We further augment the benchmark version of the Heckman selection model by adding the product mismatch variable into the set of explanatory variables. We have two main concerns regarding the augmented specification of the model: first, whether it supports the product mismatch hypothesis (H1) and, second, how much the estimated coefficients on the other variables change compared to the benchmark model. If the product mismatch possesses its own effect on the formation of HNC (i.e., different from the effects already accounted for in the model), then the effect must be significant and the effects of the other variables must be quantitatively similar to the benchmark. As our estimation results show, this is what we indeed observe in the data: we find strong support for our product mismatch hypothesis H1 while we still obtain the same effects from the benchmark control variables as above. That is, product mismatch matters: the higher the mismatch[12], the higher both the probability of selection and the conditional size of HNC. Regarding the benchmark variables, the only change in the estimated effects pertains to ROA: while it is still significant at the selection stage, it is no longer significant at the outcome stage. This could suggest that the effect of ROA on HNC formation is transmitted through our product mismatch variable, at least partially. This makes sense because banks first decide on the structure of their assets and liabilities (choose the degree of mismatch) and then face the consequences of their decisions (higher or lower ROA).

Further, to assess the magnitude of the implied economic effect, consider an increase in the mismatch variable by its 1 sample standard deviation. This corresponds, for example, to an increase in each component of the mismatch variable (i.e., the ratio of household deposits to liabilities and the ratio of corporate credit to assets; see Equation (3)) by 4–5 percentage points. The final model suggests that the resulting increase in the size of HNC is 1.2 percentage points of the failed banks' total liabilities, which constitutes 13% of the

[12] The composition of an increase in the mismatch is discussed in Section 3.1, see Equation (3).

standard deviation of HNC and is thus meaningful.[13] We then repeat the computation of the economic effects: we generate the size of failed banks' HNC in the four quartiles of the banks' distribution by the product mismatch variable. The computation results clearly show that the largest response of HNC to changes in mismatch occurs in the 4^{th} quartile (Fig. 1). That is, the expected value of the size of HNC, conditional on being selected, increases from 18.5% in the 3^{rd} to 22.6% in the 4^{th} quartiles (the mean prediction for the size of HNC is 18.6% of failed banks' total liabilities).

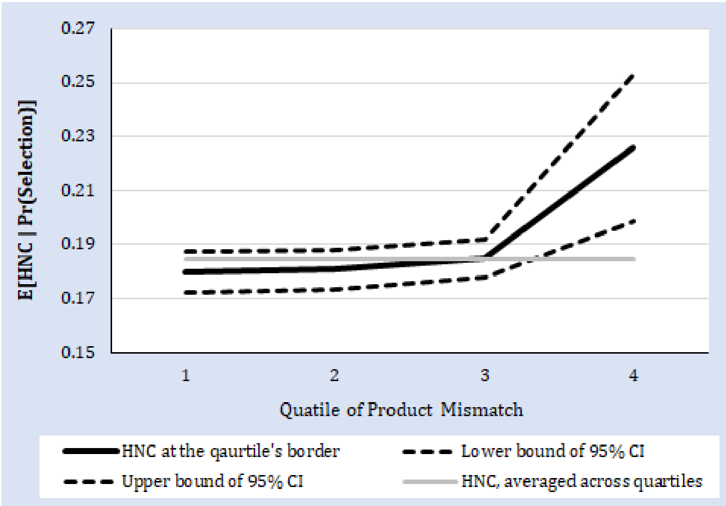


Figure 1: Average predicted size of HNC in the quartiles of the product mismatch distribution (*U.S. failed banks*)

Note: The values of HNC are computed based on the outcome equation of the Final Model, see Table 6).

Notably, the use of the mismatch variable caused a doubling of the correlation between the selection and outcome regression errors (in absolute terms); moreover, the precision of the correlation also improved substantially. This implies that the mismatch variable is crucially important in the identification of sample selection bias. From the technical standpoint, the use of both the benchmark and the final versions of the Heckman selection model for U.S. banks is justified by the presence of a statistically significant correlation between the regression errors in selection and outcome equations.

We further compare the predicted values of the probability and the size of HNC for the whole sample of U.S. banks and for the subsamples of failed and operating banks (Table 7; computations are based on the Final Model from Table 6). Several conclusions emerge. First, the model distinguishes between failed and operating banks well: the average probability of selection is about 98.5% for the former compared to 0.4% for the

[13]The effect is computed for the sample of failed banks using the marginal effects routine after the estimation of the Heckman selection model (the *margins* build-in routine in Stata). First, we generate the average value of HNC at the means of each explanatory variable, including the product mismatch, and then we generated the average value of HNC, having increased only the mismatch variable's mean (0.007) by its one standard deviation (0.017).

latter. Importantly, the model predicts very large variation in both subsamples: from 1% to 100% for failed and from 0% to 95% for operating banks. The latter also implies that the sample of operating banks contains candidates for being failed banks with HNC not yet revealed by the FDIC as of 2016Q2. Second, the model captures the average size of HNC in the subsample of failed banks, 18.5%, which almost coincides with the actual value, 18.6% (good in-sample mean prediction). However, the maximal predicted value is 11 percentage points lower than the actually observed (32% compared to 43%), which is still valuable but should be improved in future research (by using, for instance, the quantile regressions as in Shaeck, 2008). The range of the predicted size of HNC in the subsample of operating banks ranges from 0% to as much as 14%, which is below the actual mean, thus indicating that the major financial problems of U.S. banks may have already been solved as of 2016Q2.

Table 7: Predictions of the probability (selection) and the conditional size (outcome) of banks' HNC: The case of U.S. banks

	Obs	Mean	SD	Min	Max
Panel 1: The whole sample					
Selection (predicted)	2.371	0.213	0.405	0.00	1.00
Outcome (predicted)	2.371	0.040	0.077	0.00	0.32
Panel 2: Failed banks					
Selection (predicted)	504	0.985	0.101	0.01	1.00
Outcome (predicted)	504	0.185	0.036	0.00	0.32
Panel 3: Operating banks					
Selection (predicted)	1,867	0.004	0.036	0.00	0.95
Outcome (predicted)	1,867	0.000	0.005	0.00	0.14

Note: The values of HNC are computed based on the outcome equation of the Final Model, see Table 6).

In general, we conclude that product mismatch in U.S. financial institutions may provide a valid signal of deteriorating banking stability and could be monitored by the FDIC to detect fragile banks in advance.

4.2 Hidden negative capital and homogeneous product mismatch: The case of Russian banks

Now we analyze to what extent the empirical results for Russian banks are similar to or different from those obtained for U.S. banks. The estimation results from the Heckman selection model appear in Table 8. The structure of this table fully mimics Table 6. All models are estimated on the bank-level data for 2017M12 (operating banks) and the historical failure events accumulated from 2007 to 2017M12. All explanatory variables are included with a three-month lag, as in the previous section.

Table 8: Heckman selection models for Russian banks

	Benchmark Model		Final Model	
	Selection	Outcome	Selection	Outcome
Mismatch			2.315*** (0.463)	0.103* (0.061)
Capital	–2.802*** (0.459)	–0.064 (0.060)	–1.857*** (0.489)	–0.014 (0.065)
NPL	–1.945*** (0.491)	–0.261*** (0.093)	–1.478*** (0.480)	–0.231** (0.091)
Liquidity	–2.183** (0.945)	–0.069 (0.122)	–1.734* (0.944)	–0.041 (0.121)
ROA	–4.542** (1.800)	–0.242 (0.193)	–4.840*** (1.774)	–0.258 (0.195)
TA growth	0.861*** (0.187)	0.022** (0.011)	0.828*** (0.182)	0.020* (0.011)
Size	–0.321*** (0.036)		–0.308*** (0.019)	
Constant	1.100*** (0.148)	0.301*** (0.019)	0.549*** (0.182)	0.277*** (0.028)
N obs.	799	799	799	799
N censored	440	440	440	440
N observed	359	359	359	359
ρ	0.31**		0.28*	
Log Likelihood	241.1		227.4	
Convergence	Yes		Yes	

Note: Dependent variable is the probability of HNC formation (in the Selection equation) or the size of HNC conditional on being selected (in the Outcome equation). All explanatory variables are taken with a three-month lag. Selection and outcome equations are estimated simultaneously using maximum likelihood (ML). ρ is the correlation between regression errors in selection and outcome equations.

***, **, * indicate that a coefficient is significant at the 1%, 5%, 10% level, respectively. Standard errors are clustered at the bank level and appear in the brackets under the estimated coefficients.

The estimation results obtained for the Russian banks share many similar features with those for U.S. banks. Most importantly, Heckman selection models are identified and the mismatch hypothesis cannot be rejected. That is, the correlation between the selection and outcome regression errors is significant, indicating the presence of sample selection bias in the data. Further, greater mismatch leads to a higher probability of selection and increases the conditional size of banks' HNC, exactly as in the case of U.S. banks. Finally, adding the mismatch variable does not change qualitatively the effects from the other control variables (compare the Benchmark against the Final models in the table).

Regarding the effects of control variables, many of the effects for Russian banks work in a similar manner as we described for U.S. banks, though some notable differences also emerge. In particular, the effects of capital, ROA, the growth of assets, liquidity, and the bank size — that is, 5 out of 6 control variables, apart from the product mismatch — are qualitatively very similar to those in U.S. banks. However, Russian banks' capital and ROA significantly affect only the selection process, while the same effects for U.S. banks are significant at both selection and outcome stage. In addition, Russian banks' liquidity significantly affects the selection process, as opposed to what is revealed for U.S. banks. Importantly, as in the case of U.S. banks above, we find support for the too-big-to-fail view for Russian banks using the bank size variable. This implies that, when increasing their size, Russian banks are more successful in hiding their problems from the Central Bank of Russia, as are larger U.S. banks in their relationships with the FDIC.

The only substantial difference between U.S. and Russian banks is that the NPL variable, being significant in both cases, appears with an expected positive sign for U.S. banks in the selection and the outcome equations, whereas its effect is rendered negative for Russian banks in both equations, which is counterintuitive at first sight. However, this could be the case in which, when already engaged in fraud, a relatively greater portion of truthfully reported NPL in a bank's balance sheet means that a corresponding part of the losses is also accounted for and thus is not a surprise for the Central Bank of Russia when closing a failed bank. In Russia, many cases of bank closure were a surprise because banks, prior to failure, reported lower NPL ratios than that reported by operating banks (recall the descriptive statistics in Table 5).

For the computation of the magnitudes of implied economic effects, consider again, as in the case of U.S. banks, an increase in the mismatch variable by 1 sample standard deviation (or, equivalently, by about 14–15 percentage points[14] in both components of the mismatch variable). Then, the Final Model for Russian banks from Table 8 predicts that the resulting increase in the size of HNC is 4 percentage points of the failed banks' total liabilities, which is equivalent to 30% of the standard deviation of HNC. This effect is thus more than two times larger than those obtained for the U.S. (13%). This is rather expected because classical financial intermediation plays a more important role in Russia than in the U.S. Further, as in the case of U.S. banks, we compute the predicted values

[14]Note that these figures are approximately three times larger than those for U.S. banks, see Section 4.1.

of the size of HNC in each of the four quartiles of the Russian failed banks' distribution by the product mismatch variable (see Fig. 2). The computation results show that, unlike U.S. banks in which a substantial increase of the predicted value of HNC occurs only in the 4^{th} quartile, for Russian banks we observe meaningful increases in each of the quartiles. For instance, in the 1^{st} quartile the predicted size of HNC equals 15% of the failed banks' total liabilities, which is 3 percentage points below the average, while in the 4^{th} quartile this value rises to as much as 28.2%. Note that this latter figure is expectedly larger in the case of Russian banks (again, due to the argument of their higher reliance on classical financial intermediation) than what we obtain for U.S. banks in the respective 4^{th} quartile (22.6%), but not substantially (only 5.6 percentage points).

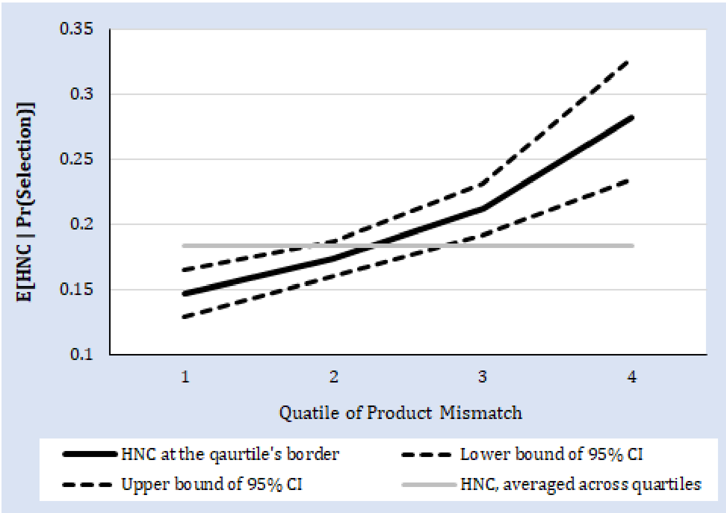


Figure 2: Average predicted size of HNC in the quartiles of the product mismatch distribution (*Russian failed banks*)

Note: The values of HNC are computed based on the outcome equation of the Final Model, see Table 8).

Finally, we discuss the in-sample predictive power of the Final Model for Russian banks, similarly to our analysis in the previous section for U.S. banks. The following conclusions materialize (Table 9; computations are based on the Final Model from Table 8). First, the model for Russian banks is also able to distinguish between failed and operating banks, but possibly not as accurately as is the case for U.S. banks. That is, the average probability of HNC formation in the subsample of failed banks is 58.4%, which is 40 percentage points lower than the corresponding number for U.S. failed banks; the average probability of HNC formation in the subsample of operating banks is 34.6%, which is more than 30 percentage points larger compared to the respective number for U.S. banks. This finding implies that the differences between failed and operating banks in Russia are rather small — much smaller than those in the U.S., meaning that many banks still operating in Russia could be those with HNC not yet revealed by the Central Bank of Russia as of 2017M12, in sharp contrast to the much favourable situation with banking stability in U.S. Among the similar features, we observe that both Final Models for the U.S. and Russian banks predict that the probability of HNC formation ranges between almost zero and 86–100% for both failed and operating banks. Second, regarding the size of HNC, again in contrast to the case of U.S. banks, the Final Model for Russian banks underpredicts the size of HNC. For instance, the mean prediction is 18.4% for failed banks, which is about 12 percentage points lower than the actually observed value (recall the descriptive statistics in Table 5). The maximal value, 40%, is also underpredicted by about 10 percentage points compared to the actual value. Thus, these figures need to be improved in future research, but this would require accounting for more specific features of the Russian banking system, thus deviating from the unified model's composition that we use here to compare the U.S. and Russia. Predictions for operating banks could thus also be underestimated: the Final Model for Russian banks delivers a mean prediction for the size of HNC of 11%, which is 3 percentage points lower than that for U.S. banks. We treat this number with caution.

Table 9: Predictions of the probability (selection) and the conditional size (outcome) of banks' HNC: The case of Russian banks

	Obs	Mean	SD	Min	Max
Panel 1: The whole sample					
Selection (predicted)	799	0.453	0.241	0.00	1.00
Outcome (predicted)	799	0.141	0.079	0.00	0.40
Panel 2: Failed banks					
Selection (predicted)	359	0.584	0.224	0.01	1.00
Outcome (predicted)	359	0.184	0.077	0.00	0.40
Panel 3: Operating banks					
Selection (predicted)	440	0.346	0.197	0.00	0.86
Outcome (predicted)	440	0.106	0.062	0.00	0.26

Note: The values of HNC are computed based on the outcome equation of the Final Model, see Table 8).

4.3 Heterogeneous mismatch effect: Small vs. large banks

Having established the existence of the product mismatch effect, we now turn to the analysis of the possible heterogeneity of this effect for banks in different size clusters. The estimation results appear in Table 10. Panel 1 of the table contains the part of the selection equation with the mismatch variable and its product with size dummies, large and small, while Panel 2 contains the coefficients on respective variables from the outcome equation. The first two columns of the table describe the heterogeneous mismatch effects for U.S. banks and the last two columns for Russian banks. For each of the two countries, the first column reflects the average mismatch effect reported above and the second column brings the estimates of the heterogeneous mismatch effect, thus enabling the comparison with the previous estimation results and between U.S. and Russian banks.

Table 10: Heckman selection models: Heterogeneous mismatch effect

	U.S. banks: 2016Q2		Russian banks: 2017M12	
	with Mismatch	+ Size clusters	with Mismatch	+ Size clusters
Panel 1: Selection equation				
Mismatch	78.78***		2.315***	
	(22.32)		(0.463)	
Mismatch×Small		73.93***		2.278***
		(23.00)		(0.464)
Mismatch×Large		–278.81		5.191**
		(196.23)		(2.098)
Size	–0.731***	–0.656***	–0.308***	–0.328***
	(0.151)	(0.165)	(0.037)	(0.040)
Panel 2: Outcome equation				
Mismatch	0.599***		0.103*	
	(0.208)		(0.061)	
Mismatch×Small		0.593***		0.101*
		(0.208)		(0.061)
Mismatch×Large		4.286		0.189
		(2.800)		(0.276)
N obs.	2371	2371	799	799
N censored	1867	1867	440	440
N observed	504	504	359	359
ρ	–0.707***	–0.632***	0.283*	0.277*
Log Likelihood	–532.8	–534.2	–227.4	–226.4

Note: Dependent variable is the probability of HNC formation (in the Selection equation) or the size of HNC conditional on being selected (in the Outcome equation). All explanatory variables are taken with a one-quarter lag for U.S. banks and a three-month lag for Russian banks. Selection and outcome equations are estimated simultaneously using maximum likelihood (ML). ρ is the correlation between regression errors in selection and outcome equations.

***, **, * indicate that a coefficient is significant at the 1%, 5%, 10% level, respectively. Standard errors are clustered at the bank level and appear in the brackets under the estimated coefficients.

The estimation results suggest that the heterogeneous mismatch hypothesis (H2) is relevant for both U.S. and Russian banks. That is, the mismatch effect is indeed heterogeneous with respect to bank size and it shares many similar features between U.S. and Russian banks.

First, the results of the selection equation for U.S. banks indicate that the average mismatch effect, revealed in the previous sections, is due to small banks only: the product of the mismatch variable and the dummy for small banks is positive and significant, whereas the same product with the dummy for large banks is not significant. We interpret this result as indicating the fact that the sample of U.S. banks covers the last decade when the major problems of U.S. banks were outside of the corporate loan market (see Chodorow-Reich (2014) for details). Additional confirmation of this conclusion comes from the outcome equation, which shows the absence of any significant influence of mismatch on the conditional size of HNC for large banks. Therefore, the effect is transmitted through the U.S. small banks. Conversely, for Russian banks, mismatch increases selection for both small and large banks. This, in turn, reflects the bank-based essence of the Russian banking system, in which larger financial institutions switch to other non-traditional banking services to a much lesser extent than those in the U.S. In this sense, the differences obtained in the estimation results are quite expected.

Second, the results for the outcome equation show that, in the case of U.S. banks, the product mismatch plays a role in determining the conditional size of HNC only for small U.S. banks, as we have already noted above. This is consistent with the informational asymmetry view (smaller banks have fewer opportunities to substitute the chosen types of assets and liabilities for others compared to larger competitors). Surprisingly, product mismatch plays a similar role in explaining the conditional size of HNC in the Russian banking system: It increases the size of HNC in small banks only. For both countries, we observe that the effect on small banks is almost the same as the average effect, implying that large banks are immune to the product mismatch.

Finally, we are interested in the extent to which the in-sample predictions of the size of HNC are different when we include or exclude the mismatch variable and how it differs between U.S. and Russian banks. Given the estimated models described above, we predict the modeled values of HNC with and without the heterogeneous mismatch effects and then plot the densities for both banking systems. The results appear in Appendix B and suggest that the role of the product mismatch in the in-sample prediction is visible for both countries, and for Russian banks it is much more substantial than for U.S. banks. Indeed, the in-sample predictions change visibly when we add the mismatches in regressions for Russian banks, while they change only slightly for U.S. banks. This is another reflection of the bank-based vs. market-based banking system features.

4.4 Rolling window regression and out-of-sample forecasting

In this section, using a rolling window regression, we analyze the accuracy of forecasting the size of HNC for bank failures in the 2011–2017 period in the United States and Russia. The out-of-sample forecasting horizon is fixed to 1 quarter (3 months): At time t, an estimation sample for each country includes historical cases of HNC registered before time t plus observations for operating banks in the current period t. The estimation results are used to predict the size of HNC in banks for h period ahead (at time $t + h$). This estimation procedure allows us to mitigate the class-imbalance problem in the data and guarantees no looking into the future at the time of making a forecast. Rolling window regression is applied for direct multi-step prediction of the HNC size in failed banks after 2011 in the two countries. In every step, the parameters of the model are re-estimated.

The main coefficients of our interest are γ_1 and γ_2 in the Heckman selection (1) and outcome (2) equations, respectively, verifying the mismatch hypothesis, and ρ, i.e. a correlation between errors of selection and outcome equations, justifying the use of the Heckman selection model. We report the key results on every step of forecasting (Figures 3 and 4). As we can see, for the United States, the product mismatch is permanently significant and sample selection remains in the data throughout the forecasting steps (correlation ρ is statistically significant). It is notable that the only period in which sample selection is not discovered is the end of 2013 to the beginning of 2014. This could be attributed to the appointment of Jerome Powell as the Chair of the Federal Reserve and the respective transition period. For Russia, sample selection arises in the data soon after the appointment of Elvira Nabiullina as the head of the Central Bank of Russia in mid-2013. She toughened supervision and started to close fragile banks at a much faster rate than before. The mismatch coefficient is positive and significant in the selection and outcome equations, which confirms our mismatch hypothesis.

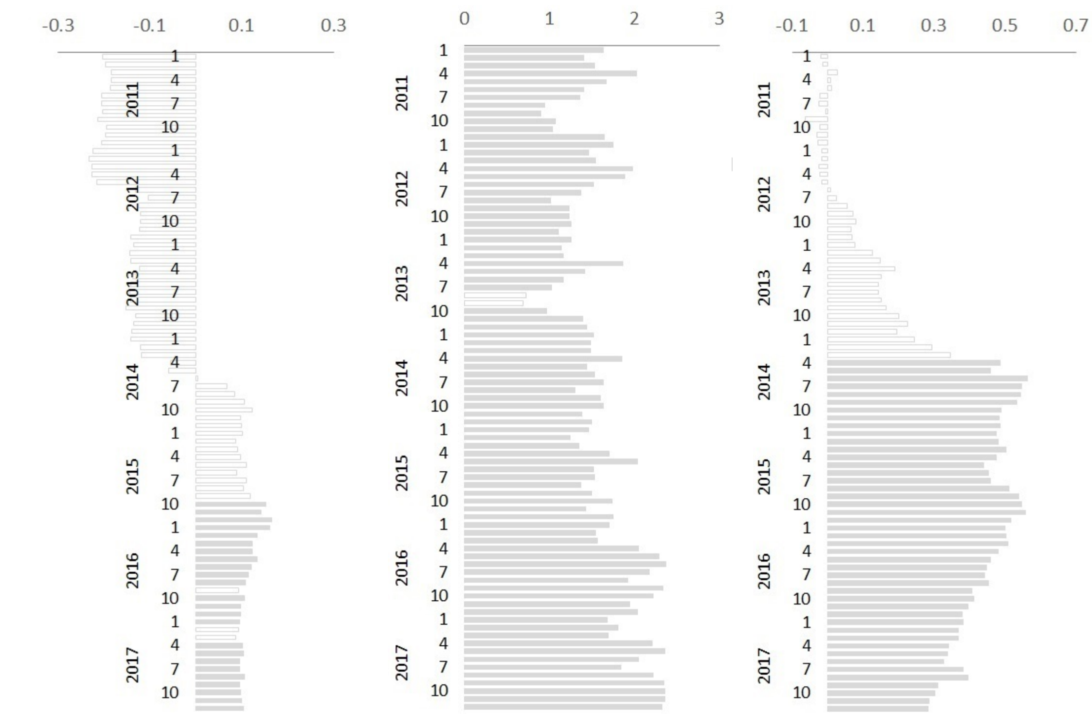
Table 11 shows the prediction results for HNC size in failed banks for the United States and Russia. We report the standard measures of forecasting accuracy, namely, mean absolute error (MAE), and mean squared forecast error (MSFE). MAE characterizes the mean error in the prediction of the ratio of HNC to liabilities (as we define the size of HNC in equation 2). Although similar in spirit, MSFE is more sensitive to large errors. We compare the performance of the suggested Heckman model with a simple OLS regression that ignores the selection problem in the data. First, the accuracy of forecasting is higher (i.e., errors are lower) for the U.S. case. This is consistent with our expectations and reflects higher transparency in the U.S. banking sector in comparison to the Russian sector. Second, the Heckman selection model generates more precise predictions of HNC in both countries, which confirms the importance of sample selection in the data. Finally, t-test for the difference in mean errors implies the statistically significant predominance of the Heckman selection model.

(a) Mismatch in selection equation

(b) Mismatch in outcome selection

(c) Correlation ρ

Figure 3: Heckman selection model for HNC in the United States in 2011–2016Q2 (22 quarterly steps): Key results. Coefficients significant at 10% significance level are solid grey.



(a) Mismatch in selection equation

(b) Mismatch in outcome selection

(c) Correlation ρ

Figure 4: Heckman selection model for HNC in Russia in 2011–2017 (84 monthly steps): Key results. Coefficients significant at 10% significance level are solid grey.

Table 11: Out-of-sample forecasting for the size of HNC in the United States and Russia

	The United States			Russia		
	Heckman	OLS	p-val.	Heckman	OLS	p-val.
MAE	0.054	0.072	0.000	0.146	0.237	0.000
MSFE	0.0046	0.0082	0.000	0.0306	0.0748	0.000

Note: MAE – mean absolute error, MSFE – mean squared forecast error, p-value is reported for a t-test (difference in means).

The United States: 347 failures with HNC in 2011Q1–2016Q2; Russia: 302 failures in 2011M1–2017M12.

5 Sensitivity analysis

We conducted several robustness checks to understand whether possible misspecification errors might affect our main findings. First, we focus on the existence of a common pattern in HNC formation in the United States and Russia and then on the role of mismatch in the process of HNC formation.

In Appendix C, we switched from the maximum likelihood (ML) to the original two-step estimator for the Heckman selection model (1)-(2).

In Appendix D, we varied the time lags of our explanatory variables – from 1 quarter in the basic specification for the United States and Russia to 2, 3 and 4 quarters. This was needed to trace the time evolution of the effects embodied in the HNC determinants. The concern we address here is that the largest economic impacts could be observed when using deeper lags of explanatory variables than the one we chose for our basic regressions.

Finally, we augment our analysis of the product mismatch heterogeneity, presented in Section 4.3, with a measure of the financial health of other banks in the system (System Capital), an extra explanatory variable suggested by Brown and Dinç (2011) to control for the too-many-to-fail effect. By doing so, we eliminate the risk that the mismatch effects of small and large banks might only capture the situations in which the banking system is weak. The idea is that when operating banks become weaker, the about-to-fail banks may easily follow their mismatching strategy knowing that financial regulators are less likely to revoke their licenses. System Capital$_{i,t}$ is computed as an average of the capital adequacy ratio in the banking sector at time t without the contribution of bank i. The estimation results are laid out in Appendix E. The results suggest that there is a significant too-many-to-fail effect in both banking systems; that is, the lower the capital of a banking system without bank i, the higher the probability of selection and the conditional size of HNC of bank i.

In all cases, our main results regarding the similar underlying forces of HNC formation in the U.S. and Russia, and the importance of mismatch effects, remain qualitatively unchanged. In addition, all estimations revealed the presence of sample selection bias in the data.

6 Concluding remarks

In this paper, we compare the formation of hidden negative capital (HNC) in the United States and Russia and empirically demonstrate that the underlying forces of HNC formation are similar in these two very different banking systems. To do so, we hand-collect unique data on the negative capital of failed banks (the negative difference between banks' assets and liabilities) revealed by financial regulators.

An obstacle on the way to modelling HNC formation is sample selection bias: We observe HNC only in failed banks. Thus, we apply the Heckman selection approach (Heckman, 1979), in which we use a bank size variable as the identifying variable in the selection equation. We argue that the size variable is valid in our case because we consider the relative, not absolute, size of HNC as a dependent variable in the outcome equation. Following the literature, we estimate a parsimonious model specification for both banking systems and use it as a reference model in our analysis.

We are primarily focused on the role of product mismatch in the formation of HNC and put forward a mismatch hypothesis: HNC formation is intensified in banks that specialize in borrowing funds from households and granting credit to non-financial firms. To test this hypothesis, the analysis is augmented with a mismatch variable. Our estimation results clearly indicate that the mismatching behavior of banks matters. On average, mismatch works similarly in both banking systems: as we hypothesized, greater mismatch leads to higher probability of selection and increases the size of HNC, conditional on being selected. Our key finding survived a number of robustness checks, and the model provides accurate out-of-sample predictions.

We further study whether the mismatch effect on HNC formation is heterogeneous with respect to bank size: the effect can be different in large banks (with more diversified assets and liabilities) compared to that in small banks. Indeed, we empirically confirm that the mismatch effect is heterogeneous; however, we discover that it works somewhat differently in U.S. and Russian banks. In the United States, the mismatch effect materializes only in small banks at both selection and outcome stages of the Heckman approach. In the Russian Federation, the effect is observed in both small and large banks at the selection stage, in contrast to that in the U.S. case, but, similar to U.S. banks, at the outcome stage the effect is significant for small Russian banks only. This points to the differences between the banking systems considered. Specifically, as the U.S. economy is transparent and market-based, larger banks are likely to expand their activities beyond traditional banking, while smaller banks have fewer opportunities to do so (consistent with the informational asymmetry view). Conversely, the Russian market is more opaque and bank-based, with traditional deposits and loans playing the key role in the banking system and banks' profits. Larger Russian banks dominate in both deposit and loan markets, implying that mismatches may seriously damage their activities in the event of managers' mistakes or intentional fraud.

Our conclusion that the same parsimonious model of HNC formation works well for two very different banking systems (external validity argument) opens avenues for future research on bank failures and HNC in other banking systems around the world. Another area for future research is the identification of banks that are still operating but are likely to be already hiding negative capital, as they have not yet been detected by financial regulators and thus are increasing potential losses to society.

Acknowledgements

We are grateful to Francesco Audrino, Zuzana Fungacova, Vasily Korovkin, Grayson Krueger, Ion Lapteacru, Nicolas Mittag, Simona Nistor, Steven Ongena, Juan-Pablo Ortega, Laura Solanko, Laurent Weill, and the participants of the BOFIT Research Seminar on March 8, 2016 (Helsinki), the 6th CInSt Banking workshop (Moscow, 2016), 2016 RCEA Money-Macro-Finance Workshop (Rimini), 36th International Symposium on Forecasting (Santander, 2016), 19th INFER Annual Conference, special session on bank risk-taking (Bordeaux, 2017), 34th International Symposium on Money, Banking and Finance (Nanterre, 2017), IFABS 2018 Porto Conference "Ten Years After the 2008 Financial Crisis: Where Are We Heading Now?", and 10th Biennial Conference of the Czech Economic Society (Prague, 2018) for their helpful comments and suggestions. The very first results regarding the Russian banks were achieved during the Mikhail Mamonov's research visit at BOFIT (March, 2016). Mikhail Mamonov acknowledges the financial support received from RFBR under the research project No. 16-36-60037.

References

Acharya, V.V. and Yorulmazer, T. (2007) "Too many to fail: An analysis of time-inconsistency in bank closure policies", *Journal of Financial Intermediation*, Vol. 16(1), pp. 1–31.

Audrino, F., Kostrov, A., and Ortega, J.-P. (forthcoming) "Predicting U.S. Bank Failures with MIDAS Logit Models", *Journal of Financial and Quantitative Analysis.*

Anzoategui, D., Martinez Peria, M.S., and Melecky, M. (2012) "Bank competition in Russia: An examination at different levels of aggregation", *Emerging Markets Review*, Vol. 13, pp. 42–57.

Arena, M. (2008) “Bank failures and bank fundamentals: A comparative analysis of Latin America and East Asia during the nineties using bank-level data”, *Journal of Banking and Finance*, Vol. 32, No. 2, pp. 299–310.

Aschcraft, A. (2005) “Are Banks Really Special? New Evidence from the FDIC-Induced Failure of Healthy Banks”, *American Economic Review*, Vol. 95, No. 5, pp. 1712–1730.

Balla, E., Prescott E.S., and Walter J. (2015) “Did PCA Fail? A comparison of bank failures and FDIC losses in the 1987-92 and 2007-13 periods”, FRB-Richmond working paper, WP15-05.

Barth, J.R., Bartholomew, P.F., and Bradley, M. (1990) “Determinants of thrift institution resolution costs”, *The Journal of Finance*, Vol. 45, pp. 731–754.

Bennett, R.L. and Unal, H. (2014) “The effects of resolution methods and industry stress on the loss on assets from bank failures”, *Journal of Financial Stability*, Vol. 15, pp. 18–31.

Berger, A.N., and Bouwman, C.H.S. (2013) “How does capital affect bank performance during financial crises?”, *Journal of Financial Economics*, Vol. 109, pp. 146–176.

Berger, A.N., and DeYoung, R. (1997) “Problem loans and cost efficiency in commercial banks”, *Journal of Banking and Finance*, Vol. 21, pp. 849–870.

Betz, F., Oprica, S., Peltonen T.A., and Sarlin, P. (2014) “Predicting distress in European banks”, *Journal of Banking and Finance*, Vol. 45, pp. 225–241.

Bovenzi, J.F. and Murton, A.J. (1988) “Resolution Costs of Bank Failures”, *FDIC Banking Review*, Vol. 1, pp. 1–13.

Boyd, J.H., Kwak, S. and Smith, B. (2005) “The Real Output Losses Associated with Modern Banking Crises”, *Journal of Money, Credit and Banking*, Vol. 37, No. 6, pp. 977–999.

Brown, C. and Dinç, I.S. (2005) “The Politics of Bank Failures: Evidence from Emerging Markets”, *The Quarterly Journal of Economics*, Vol. 120(4), pp. 1413-1444.

Brown, C. and Dinç, I.S. (2011) “Too many to fail? Evidence of regulatory forbearance when the banking sector is weak”, *The Review of Financial Studies*, Vol. 24, pp. 378–405.

Chernykh, L. and Cole, R. (2011) “Does deposit insurance improve financial intermediation? Evidence from the Russian experiment”, *Journal of Banking and Finance*, Vol. 35, pp. 388–402.

Chodorow-Reich, G. (2014) "The Employment Effects of Credit Market Disruptions: Firm-Level Evidence from the 2008–9 Financial Crisis", *Quarterly Journal of Economics*, Vol. 129(1), pp. 1–59.

Cleary, S. and Hebb, G. (2016) "An efficient and functional model for predicting bank distress: In and out of sample evidence", *Journal of Banking and Finance*, Vol. 64, pp. 101–111.

Cole, R.A. and White, L.J. (2012) "Déjà Vu All Over Again: The Causes of U.S. Commercial Bank Failures This Time Around", *Journal of Financial Services Research*, Vol. 42, pp. 5–29.

Cole, R. and White, L. (2017) "When Time Is Not on Our Side: The Costs of Regulatory Forbearance in the Closure of Insolvent Banks", *Journal of Banking and Finance*, Vol. 80(C), pp. 235–249.

Demirgüç-Kunt, A., Detragiache E., and Gupta, P. (2006) "Inside the crisis: An empirical analysis of banking systems in distress", *Journal of International Money and Finance*, Vol. 25, pp. 702–718.

DeYoung, R. and Torna, G. (2013) "Nontraditional banking activities and bank failures during the financial crisis", *Journal of Financial Intermediation*, Vol. 22, pp. 397–421.

Diamond, D.W. and Rajan, R.G. (2000) "A Theory of Bank Capital", *The Journal of Finance*, Vol. 55, No. 6, pp. 2431–2465.

Fungacova, Z. and Weill, L. (2013) "Does competition influence bank failures? Evidence from Russia", *Economics of Transition*, Vol. 21, pp. 301–322.

Fungacova, Z. and Poghosyan, T. (2011) "Determinants of bank interest margins in Russia: Does bank ownership matter?", *Economic Systems*, Vol. 35, pp. 481–495.

Granja, J., Matvos, G., and Seru, A. (2017) "Selling Failed Banks", *The Journal of Finance*, Vol. 72, No. 4, pp. 1723–1784.

Gropp, R., Mosk, T., Ongena, S. and Wix C. (2018). "Banks Response to Higher Capital Requirements: Evidence from a Quasi–Natural Experiment", *Review of Financial Studies*, hhy052.

João, G. (2013) "The relation between bank resolutions and information environment: Evidence from the Auctions for failed banks", *Journal of Accounting Research*, Vol. 51, pp. 1031–1070.

Heckman, J.J. (1979) "Sample selection bias as a specification error", *Econometrica*, Vol. 47, pp. 153–161.

Homar, T., van Wijnbergen, S.J.G. (2017). “Bank recapitalization and economic recovery after financial crises”, *Journal of Financial Intermediation*, Vol. 32(C), pp. 16-28.

James, C. (1991) “The losses realized in bank failures”, *The Journal of Finance*, Vol. 46, pp. 1223–1242.

Jimenez, G., Saurina, J. (2006) “Credit cycles, credit risk, and prudential regulation”, *International Journal of Central Banking*, Vol. 2, pp. 65–98.

Juurikkala, T., Karas, A., and Solanko, L. (2011) “The Role of Banks in Monetary Policy Transmission: Empirical Evidence from Russia”, *Review of International Economics*, Vol. 19, No. 1, pp. 109–121.

Kang, A., Lowery, R., and Wardlaw, M. (2015) “The cost of closing failed banks: A structural estimation of regulatory incentives”, *The Review of Financial Studies*, Vol. 28, No. 4, pp. 1060–1102.

Karas, A., Pyle, W., and Schoors, K. (2013) “Deposit Insurance, Banking Crises, and Market Discipline: Evidence from a Natural Experiment on Deposit Flows and Rates”, *Journal of Money, Credit and Banking*, Vol. 45, No. 1, pp. 179–200.

Karminsky A.M. and Kostrov A. (2017) “The back side of banking in Russia: forecasting bank failures with negative capital”, *International Journal of Computational Economics and Econometrics*, Vol. 7, No. 1-2, pp. 170–209.

Kashyap, A. and Stein, J. (2000). “What Do a Million Observations on Banks Have to Say about the Transmission of Monetary Policy?”, *The American Economic Review*, Vol. 90, pp. 407-428.

Kiyotaki N. and Moore J. (1997) “Credit Cycles”, *Journal of Political Economy*, Vol. 105, No. 2, pp. 211–248.

Maddala, G.S. (1983) “Limited-dependent and qualitative variables in econometrics”, Cambridge University Press, Cambridge.

Mannasoo, K. and Mayes, D.G. (2010) “Explaining bank distress in Eastern European transition countries”, *Journal of Banking and Finance*, Vol. 33, pp. 244–253.

McKinsey (2010) “Hidden in plain sight: The hunt for banking capital”, Article McKinsey Quarterly.

O’Hara, M. and Shaw, W. (1990) “Deposit Insurance and Wealth Effects: The Value of Being ”Too Big to Fail”, *The Journal of Finance*, Vol. 45(5), pp. 1587–1600.

Osterberg, W.P., and Thomson, J.B. (1995) “Underlying determinants of closed-bank resolution costs”, Chapter in *The Causes and Costs of Depository Institution Failures*, pp. 75–92.

Poghosyan, T., and Cihak M. (2011) “Determinants of bank distress in Europe: Evidence from a new dataset”, *Journal of Financial Services Research*, Vol. 40, pp. 163–184.

Shaeck, K. (2008) “Bank liability structure, FDIC loss, and time to failure: A quantile regression approach”, *Journal of Financial Services Research*, Vol. 33, pp. 163–179.

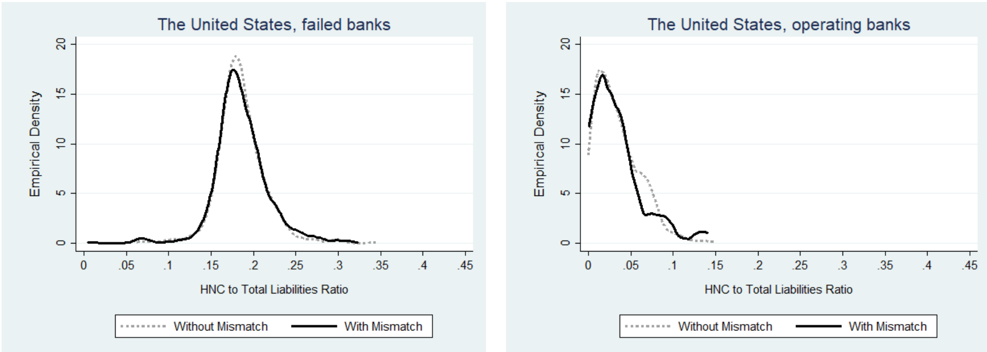
Wheelock, D. and Wilson, P. (2000) “Why do banks disappear? The determinants of U.S. bank failures and acquisitions”, *The Review of Economics and Statistics*, Vol. 82, pp. 127–138.

Appendix A. Correlations between HNC and bank size

Table I: Sample correlations between banks' HNC to total liabilities ratio and bank size

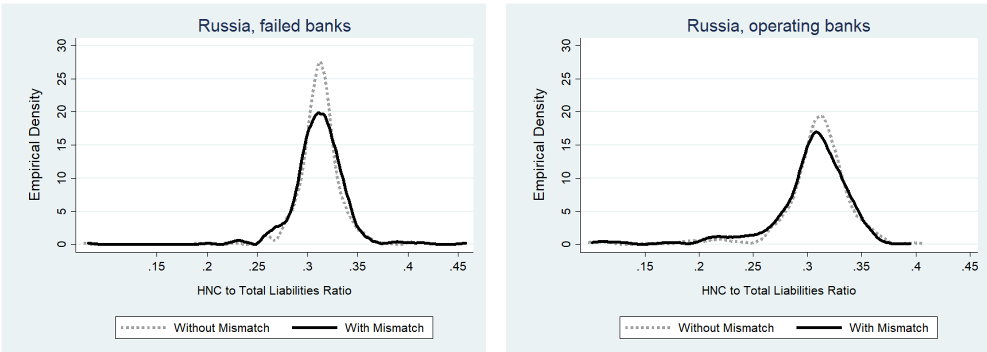
Panel 1: The United States			
HNC	1.000		
Bank size	–0.140	1.000	
Dummy for large banks	0.004	0.644	1.000
Panel 2: Russia			
HNC	1.000		
Bank size	0.131	1.000	
Dummy for large banks	–0.042	0.619	1.000

Appendix B. HNC in-sample predictions: The role of heterogeneous mismatches



(a) Failed banks (b) Operating banks

Figure 5: Density of HNC predictions with and without mismatch: U.S. banks



(a) Failed banks (b) Operating banks

Figure 6: Density of HNC predictions with and without mismatch: Russian banks

Appendix C. Robustness to the estimator: Heckman's efficient 2-step method

Table II: Heckman selection models: An alternative estimator

	U.S. banks: 2016Q2		Russian banks: 2017M12	
	Basic: ML	2-step	Basic: ML	2-step
Mismatch: selection equation	74.78*** (22.32)	62.04** (25.05)	2.315*** (0.463)	2.327*** (0.460)
Mismatch: outcome equation	0.599*** (0.208)	0.604*** (0.209)	0.103* (0.061)	0.119* (0.065)
N obs.	2371	2371	799	799
N censored	1867	1867	440	440
N observed	504	504	359	359
ρ	–0.707***	–0.565(-)	0.28*	0.378(-)
inv. Mills ratio	–	–0.045*** (0.017)	–	0.052* (0.028)

Note: Dependent variable is the probability of HNC formation (in the Selection equation) or the size of HNC conditional on being selected (in the Outcome equation). All explanatory variables are taken with 1 quarter lag for U.S. banks and 3 months for Russian banks. Selection and outcome equations are estimated either simultaneously using maximum likelihood (ML, the baseline version) or separately using Heckman's two-step efficient estimator. ρ is the correlation between regression errors in selection and outcome equations.

***, **, * indicate that a coefficient is significant at the 1%, 5%, 10% level, respectively. Standard errors are clustered at the bank level and appear in the brackets under the estimated coefficients.

Appendix D. Robustness to the forecasting horizon

Table III: Heckman selection models for U.S. banks:
Different lag structures of regressors

	Forecasting horizon			
	Basic: 1Q	2Q	3Q	4Q
Mismatch: selection equation	74.78*** (22.32)	77.45*** (22.54)	68.53*** (22.01)	56.36*** (20.87)
Mismatch: output equation	0.599*** (0.208)	0.586*** (0.210)	0.596*** (0.211)	0.613*** (0.210)
N obs.	2371	2370	2394	2382
N censored	1867	1866	1891	1880
N observed	504	504	503	502
ρ	–0.707*** (0.238)	–0.656*** (0.236)	–0.643*** (0.229)	–0.544** (0.229)
Log Likelihood	532.8	530.4	529.9	525.5

Table IV: Heckman selection models for Russian banks:
Different lag structures of regressors

	Forecasting horizon			
	Basic: 3M	6M	3Q	4Q
Mismatch: selection equation	2.315*** (0.463)	2.338*** (0.464)	2.045*** (0.469)	1.69*** (0.443)
Mismatch: output equation	0.103* (0.061)	0.096* (0.060)	0.096* (0.060)	0.087 (0.061)
N obs.	799	796	808	813
N censored	440	446	469	484
N observed	359	350	339	329
ρ	0.28*** (0.238)	0.311** (0.146)	0.326** (0.140)	0.367*** (0.142)
Log Likelihood	227.4	232.2	219.1	253.2

Note: Dependent variable is the probability of HNC formation (in the Selection equation) or the size of HNC conditional on being selected (in the Outcome equation). All explanatory variables are taken with 1 (basic), 2, 3 or 4 quarter lags for U.S. banks and 3 (basic), 6, 9 or 12 month lags for Russian banks. Control variables are not reported.

***, **, * indicate that a coefficient is significant at the 1%, 5%, 10% level, respectively. Standard errors are clustered at the bank level and appear in the brackets under the estimated coefficients.

Appendix E. Robustness to the too-many-to-fail effect

Table V: Heckman selection models: Too-many-to-fail effect

	U.S. banks: 2016Q2		Russian banks: 2017M12	
	I (basic)	II	III (basic)	IV
Panel 1: Selection equation				
Mismatch	78.78***		2.315***	
	(22.32)		(0.463)	
Mismatch×Small		74.82***		3.422***
		(19.88)		(0.420)
Mismatch×Large		–329.01***		1.212
		(81.51)		(2.374)
Panel 2: Outcome equation				
Mismatch	0.599***		0.103*	
	(0.208)		(0.061)	
Mismatch×Small		0.482***		0.119*
		(0.171)		(0.062)
Mismatch×Large		4.194		0.172
		(3.733)		(0.180)
System Capital		–1.814**		–2.437***
		(0.843)		(0.649)
N obs.	2371	2371	799	799
N censored	1867	1867	440	440
N observed	504	504	359	359
ρ	–0.707***	–0.608*	0.28*	0.348**
Log Likelihood	532.8	536.6	–227.4	–245.1

Note: Dependent variable is the probability of HNC formation (in the Selection equation) or the size of HNC conditional on being selected (in the Outcome equation). All explanatory variables are taken with 1 quarter lag for U.S. banks and 3 months lag for Russian banks. Selection and outcome equations are estimated simultaneously using maximum likelihood (ML). System Capital$_{i,t-k}$ is not included in the selection equation because of ML convergence problems. ρ is the correlation between regression errors in selection and outcome equations.

***, **, * indicate that a coefficient is significant at the 1%, 5%, 10% level, respectively. Standard errors are clustered at the bank level and appear in the brackets under the estimated coefficients.

BOFIT Discussion Papers

A series devoted to academic studies by BOFIT economists and guest researchers. The focus is on works relevant for economic policy and economic developments in transition / emerging economies.

2018	No 1	Zheng (Michael) Song and Wei Xiong: Risks in China's financial system
	No 2	Jennifer N. Carpenter, Fangzhou Lu and Robert F. Whitelaw: The real value of China's stock market
	No 3	Bing Xu: Permissible collateral and access to finance: Evidence from a quasi-natural experiment
	No 4	Stefan Angrick and Naoyuki Yoshino: From window guidance to interbank rates. Tracing the transition of monetary policy in Japan and China
	No 5	Veronika Belousova, Alexander Karminsky and Ilya Kozyr: Bank ownership and profit efficiency of Russian banks
	No 6	Chengsi Zhang and Chao Dang: Is Chinese monetary policy forward-looking?
	No 7	Israel Marques II: Firms and social policy preferences under weak institutions: Evidence from Russia
	No 8	Ivan Lyubimov, Margarita Gvozdeva and Maria Lysyuk: Towards increased complexity in Russian regions: networks, diversification and growth
	No 9	Jeannine Bailliu, Xinfen Han, Mark Kruger, Yu-Hsien Liu and Sri Thanabalasingam: Can media and text analytics provide insights into labour market conditions in China?
	No 10	Sanna Kurronen: Oil price collapse and firm leverage in resource-dependent countries
	No 11	Marlene Amstad, Huan Ye and Guonan Ma: Developing an underlying inflation gauge for China
	No 12	Michael Funke, Rongrong Sun and Linxu Zhu: The credit risk of Chinese households – A micro-level assessment
	No 13	Yiyi Bai, Tri Vi Dang, Qing He and Liping Lu: Does lending relationship help or alleviate the transmission of liquidity shocks? Evidence from a liquidity crunch in China
	No 14	Hao Wang, Jan Fidrmuc and Yunhua Tian: Growing against the background of colonization? Chinese labor market and FDI in a historical perspective
	No 15	Paul-Olivier Klein and Laurent Weill: Bank profitability and economic growth
	No 16	Zuzana Fungáčová, Paul-Olivier Klein and Laurent Weill: Persistent and transient inefficiency: Explaining the low efficiency of Chinese big banks
	No 17	Oleksandr Faryna and Heli Simola: The transmission of international shocks to CIS economies: A global VAR approach
	No 18	Michael Funke, Andrew Tsang and Linxu Zhu: Not all cities are alike: House price heterogeneity and the design of macro-prudential policies in China
	No 19	Timo Korkeamäki, Nader Virk, Haizhi Wang and Peng Wang: Learning Chinese? The changing investment behavior of foreign institutions in the Chinese stock market
	No 20	Bingyang Lv, Yongzheng Liu, Yan Li and Siying Ding: Fiscal incentives, competition, and investment in China
	No 21	Tho Pham, Oleksandr Talavera and Andriy Tsapin: Shock contagion, asset quality and lending behavior'
	No 22	Johannes C. Buggle and Steven Nafziger: The slow road from serfdom: Labor coercion and long-run development in the former Russian Empire
	No 23	Eeva Kerola: In search of fluctuations: Another look at China's incredibly stable GDP growth
2019	No 1	Çağatay Bircan and Orkun Saka: Lending cycles and real outcomes: Costs of political misalignment
	No 2	Lucy Chernykh, Denis Davydov and Jukka Sihvonen: Financial stability and public confidence in banks
	No 3	Yin-Wong Cheung and Shi He: Truths and myths about RMB misalignment: A meta-analysis
	No 4	Yuping Deng, Yanrui Wu, Helian Xu: Political connections and firm pollution behaviour: An empirical study
	No 5	Sophia Chen, Lev Ratnovski and Pi-Han Tsai: Credit and fiscal multipliers in China
	No 6	Alexander Kostrov and Mikhail Mamonov: The formation of hidden negative capital in banking: A product mismatch hypothesis

BOFIT Discussion Papers
http://www.bofit.fi/en • email: bofit@bof.fi
ISSN 1456-4564 (print) // ISSN 1456-5889 (online)