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

Alexander Kostrov [1] **, Mikhail Mamonov** [2]

February 20, 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. 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. 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. 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 model.

1 Introduction

Bank capital has to satisfy the official regulatory requirements in all times. However, during the periods of financial turbulence, bank capital adequacy ratio can fall below not only the regulatory thresholds but even further below zero, depending on the size of shocks, thus rendering bank capital negative. In the latter case, troubled banks have huge incentives to falsify their financial statements by reporting artificially positive levels of capital to keep their licences, thus

[1]Chair of Mathematics and Statistics, University of St. Gallen, Bodanstrasse 6, 9000 St Gallen, Switzerland, +41 71 224 2475, alexander.kostrov@unisg.ch

[2]CERGE-EI, Politickych veznu 7, 111 21 Prague 1, Czech Republic, +420 776 071 490, mikhail.mamonov@cerge-ei.cz

We are grateful to Francesco Audrino, Vasily Korovkin, Grayson Krueger, Nicolas Mittag, and Juan-Pablo Ortega for their helpful comments and suggestions.

engaging in fraud and insider abuse (James, 1991; Kang et al., 2015; Cole and White, 2017).[3] These banks essentially operate with hidden negative capital (hereinafter, HNC) which is likely to bear certain — and, possibly, very large — losses to the 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 and we develop an empirical setting, which allows us to distinguish between those banks that are still operating with positive capital and those banks that are still operating but are likely to hide negative capital.

Previous research has identified several forces that lead to negative capital in banks (James, 1991; Kang et al., 2015): higher portion of non-performing loans, incomes 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 and, to the best of our knowledge, neither of them were examined in this context. In this paper, when investigating the process of HNC formation, we focus on product mismatch, leaving risk and liquidity mismatches for future work. 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 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 already analyzed forces 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 banks that they choose than private depositors do[4], 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, 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

[3]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 into future due to high monetary and/or non-monetary costs associated with the closure of particular troubled banks in current period (the regulatory forbearance effect, Kang et al. (2015)).

[4]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.

treatment effect) consistent with the informational asymmetry view.[5]

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, and if they eventually fail, society incurs a variety of losses. 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 because of either informational asymmetry and possible political pressure (Brown and Dinç, 2005; Kang et al., 2015), a government budget deficit when the banking system is weak (Brown and Dinç, 2011), or the risk of missing the opportunity to sell failing banks to healthier banks (Bennett and Unal, 2014; Granja et al., 2017). However, a recent study by Berger and Bouwman (2013) shows that bank capital matters for retaining market shares and for sustaining the stability of banking services provision: For smaller banks this effect holds across all phases of business cycles; for larger banks during periods of crisis. Therefore, either bank capital must become positive again (after a shock dies out) or the central bank eventually reveals HNC, which entails a bank failure.

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, particulary 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.[6] Non-monetary costs 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), which, in turn, are affected by a decreased availability of bank credit (Kang et al., 2015). 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 infer that a self-disappearing HNC is unlikely to be a very frequent event in practice.

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. This is because financial regulators have to bear certain reputational costs

[5] Larger banks have more options to diversify their assets and liabilities than smaller banks because of greater confidence and transparency, 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).

[6] 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.

of publishing these data because the existence of HNC *per se* may question the validity of the central bank's prudential policy (valid prudential regulation should exclude HNC) and, in addition, may undermine the confidence to operating banks (they can 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 finance-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 will help to generalize our finding.[7]

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 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.

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 regression analysis for the probability of a bank failure with HNC and conditional size of HNC, and use it with 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 deposit interest rates and the returns on loans for both the United States and Russia — even on the aggregate level — clearly support the idea of a reduced interest rate margin of a mismatching 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 a higher probability that HNC exists and with a larger size of HNC.

Third, for both countries, we analyze the heterogeneity of the mismatch effect on HNC.

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

Specifically, we split our samples into two asset classes, smaller and larger banks, and consider the transmission of mismatch in each clas. This allows us to study the possible heterogeneity of the mismatch effect and how this heterogeneity differs between U.S. and Russian banks.

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.

In essence, our estimation results suggest that Heckman selection approach is 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. Conversely, in Russia, mismatching appears both from the side of larger and smaller banks. Although we do not specify the reasons of the latter given the limitations of the data at hand, we attribute these reasons to the "pocket bank" problem and the gambling of larger banks with the financial regulator in Russia. Clarifying these reasons could be an avenue for future research.

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 were rather focused on bank failures *per se*, motivating their analysis by the fact that bank failures are dangerous for the economy because they destruct relationship lending and reduce the total supply of credit to the economy (Aschcraft, 2005) and increase the GDP losses associated with them (Boyd, Kwak and Smith, 2005). However, there has not been undertaken a systematic attempt to predict both bank failures and the size of negative capital.

In most cases, the focus of previous studies on bank failures was 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 the best of our knowledge, while the former encompasses many countries around the globe.[8]

2.2 The determinants of negative capital

In the influential paper by James (1991), the 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 including book value of equity, core deposits, and a few types of non-performing assets. By applying OLS, James (1991) obtains a surprising result that the pre-failure equity capital was positively associated with the ex-post size of negative capital. In attempts to clarify this finding, he attributes it to fraud and insider abuse that are more prevalent in better capitalized banks among those who failed. From 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 relationships estimated. The research following James (1991) has developed the analysis of negative capital in many directions. Among the latter one can find the examination of regulatory forbearance effects (Brown and Dinç, 2011; Kang et al., 2015; Cole and White, 2017, etc.), understanding of the real effects of failing banks (Bennett and Unal, 2014; Granja et al., 2017), analyzing the differences between high-cost and low-cost failures (Shaeck, 2008), and addressing the sample selection bias concerns (Balla et al., 2015). Though it would be redundant to consider all these studies in details, we have to acknowledge that they exploit basically the same determinants of negative capital that can be extracted from the U.S. banks' financial reporting.

[8]Indeed, one can find a great number of studies on the probability of bank failures dealing with either U.S. banks (Cole and White, 2012; DeYoung and Torna, 2013; Cleary and Hebb, 2016; Audrino et al., forthcoming), just 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 – 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).

From the studies on regulatory forbearance we can borrow a number of determinants of negative capital beyond those used by James (1991) and mentioned above. 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 commitment and letters of credit decrease the 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 of bank failure with HNC; however, unlike in James (1991), they provide strong evidence on 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 a 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 the 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 are rather acquired than liquidated.

Research on the differences between high-cost and low-cost failures brings another portion of relevant predictors for HNC. Shaeck (2008) emphasizes that regulators are mainly concerned with the expensive failures and that the factors driving the high- and the low-cost failures might be different. While not addressing the sample selection issue and using the sample of U.S. banks that failed in 1984-2003, 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 increases the cost of bank failures. Note that Shaeck (2008) is the first paper to incorporate liability structure into the empirical models of negative capital, the fact that we will also exploit in our estimations.

Finally, the paper by Balla et al. (2015) makes a first attempt to address the sample selection issue by analyzing both failed and surviving banks. The sample selection bias appears since some of the existing banks have *de facto* failed but go on 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 for the size of negative capital conditional on being selected. Importantly, Balla et al. (2015) show that the correlation between the selection and outcome regressions' errors is statistically significant confirming thus the existence of sample selection bias.

Therefore, we use the mentioned above papers to outline the most basic set of the determinants of HNC, taking into account possible data limitations that stem from the cross-county nature of research.

3 Model and data

In this section, we describe the main steps of our empirical strategy and then discuss the data used on U.S. and Russian banks.

3.1 Empirical strategy: Heckman selection models

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

We employ two versions of the Heckman selection model to explain the HNC formation. The benchmark version builds on the basic set of determinants used in the literature. The specification proposed in this paper extends the reference model by adding the mismatch variable, a cross-product of loans to firms and the deposits of households as Equation 3 defines. After that, we augment our suggested model with bank clustering by assets size (small banks and large banks) in order to tests the heterogeneous mismatch hypothesis and to study its transmission channels. Heckman model returns the predictions on whether a given bank in the sample is likely to fail with negative capital and the conditional size of its HNC.

The Heckman selection model requires the selection equation to have at least one variable that significantly affects selection but have no effect on the outcome. There is no significant correlation between the size of a bank and the relative size of HNC. This statement is intuitively reasonable because HNC is a consequence of wrong business decisions of a bank's managers (bad luck in the spirit of Berger and DeYoung (1997) or on the degree of illegal activities and falsifications in a bank. In both U.S. and Russian samples used in our analysis, there are large banks with small relative HNC and small banks with large relative HNC. This allows us to use the size of a bank as the variable identifying the model.The resultant specification of the Heckman selection model takes the following form:

$$
\begin{aligned}
D_{i,t} &= \alpha_1 + \sum_{i=1}^{n} \beta_1 \text{BASIC}_{i,t-k} + \gamma_1 \text{Mismatch}_{i,t-k} + \delta_1 \text{Size}_{i,t-k} + \epsilon_{1i,t}, && (1) \\
\frac{\text{HNC}_{i,t}}{\text{Liabilities}_{i,t}} &= \alpha_2 + \sum_{i=1}^{n} \beta_2 \text{BASIC}_{i,t-k} + \gamma_2 \text{Mismatch}_{i,t-k} + \epsilon_{2i,t}, && (2)
\end{aligned}
$$

where $\text{HNC}_{i,t}$ is the absolute size of hidden negative capital of bank i at time t uncovered by the regulator (zero for operating banks); Liabilities are the value of total liabilities officially reported in financial statements; D 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; Size is the natural logarithm of bank total assets, the identifying variable. α, β, and γ are 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 regression analysis and report key results for larger forecasting horizons in Appendices.

The set of n BASIC predictors that stems from the previous literature and allows us to estimate the same model for U.S. and Russian banks includes:

- Capital — the ratio of bank equity 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.

$Mismatch_{it}$ reflects mismatch in a bank's strategy: funding with relatively expensive deposits from households and granting loans to firms at a lower interest rate. We define mismatch for bank i at time t as the product of the household deposits to liabilities ratio and corporate loans to assets ratio:

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

where LnsF is a proxy for loans to firms (as a share in total assets) and DepH is a proxy for household deposits (as a share in total liabilities). We say that the product mismatch hypothesis is accepted if γ_2 in equation 2 is grater than zero and statistically significant, meaning that larger mismatch increases HNC in a bank.

Then, we divide our samples of U.S. and Russian banks in two parts with respect to 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. We augment the proposed model (equations (1)–(2)) with size dummies (denoted Small and Large) and cross-products between the mismatch and respective dummies both in the U.S. and Russian cases.

$$\begin{aligned} D_{i,t} &= \alpha_1 + \sum_{i=1}^{n} \beta_i \text{BASIC}_{i,t-k} + \gamma_1 \text{Mismatch}_{i,t-k} + \theta_{11} \text{Mismatch}_{i,t-k} \text{Large}_{i,t-k} \\ &\quad + \theta_{12} \text{Mismatch}_{i,t-k} \text{Small}_{i,t-k} + \delta_1 \text{Size}_{i,t-k} + \epsilon_{1i,t}, \qquad (4) \\ \frac{\text{HNC}_{i,t}}{\text{Liabilities}_{i,t}} &= \alpha_2 + \sum_{i=1}^{n} \beta_i \text{BASIC}_{i,t-k} + \gamma_2 \text{Mismatch}_{i,t-k} + \theta_{21} \text{Mismatch}_{i,t-k} \text{Large}_{i,t-k} \\ &\quad + \theta_{22} \text{Mismatch}_{i,t-k} \text{Small}_{i,t-k} + \epsilon_{2i,t}. \qquad (5) \end{aligned}$$

where $\text{Large}_{i,t-k}$ and $\text{Small}_{i,t-k}$ are size dummies.

We include a dummy variable for large banks in the outcome equation (small banks form a reference group) but not in the selection equation. This need to be clarified. First, one might argue that dummy variable Large should not affect the outcome equation. In Appendix A we show that the correlation between relative HNC and this dummy is negligible indeed so its inclusion into the outcome equation should not affect the estimation procedure qualitatively.

Second, Size and Large variables are highly correlated and we thus leave only size in the outcome equation as an identifying variable.

We estimate the selection and outcome equations (1)–(2) and (4)–(5) simultaneously by the maximum likelihood method (efficient two-step estimates appear in the robustness check, see Section 5). The sample selection bias is present in the model if error termsϵ_1 and ϵ_2 are not independent. We test it by applying the likelihood ratio test with the null hypothesis of no correlation between the selection and outcome regressions' errors.

3.2 Data

In this section, we describe the bank-level data on HNCs and and its potential determinants first for the U.S. banking sector and for the Russian banking sector.

The bank-specific financial characteristics of U.S. banks are coming from Call Reports on FDIC-insured banks for the period from 2004 to the second quarter of 2016. These reports are disseminated by FDIC in quarterly sets of files with detailed banking statistics.

We compile several sources of the statistics on Russian banks for 2007–2017. Monthly balance sheet (Form 101) and the quarterly profit & loss statement (Form 102) are made available by the Central Bank of Russia. Since 2007, information on HNCs in failed banks appears in the press releases of the Central Bank of Russia (in so-called *Vestnik Banka Rossii*). These data were manually collected. Moreover, in some cases the size of HNC is re-evaluated (usually increased) by the Central Bank of Russia so that we had to double-check all values in the financial news.

3.2.1 The data on U.S. banks

Our sample includes 50 quarters of data for 9936 unique banks. The information on the 525 failure cases with corresponding hidden negative capital is documented in the FDIC's Failed Bank list. Table 2 presents the descriptive statistics on the HNCs of 525 failed U.S. banks for the considered period.

Table 2: Descriptive statistics for failed U.S. banks: Hidden negative capital (HNC)

Absolute value of HNC	Mean	SD	Min	Max	Percentiles		
					P25	P50	P75
% of remaining total assets	0.243	0.137	0	0.754	0.141	0.229	0.335
% of remaining total liabilities	0.186	0.087	0	0.430	0.123	0.186	0.251

On average, U.S. banks have smaller HNC in comparison with Russian peers. This could be a sign of higher transparency and better supervision in the U.S. banking sector. Naturally, the value of 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 short before they fail. So, we here do not measure the size of HNC with respect to reported bank capital. In table 3 we compare descriptive statistics for failed and operating banks .

Table 3: Descriptive statistics for U.S. banks: Failed vs. operating banks, 2004–2016Q2

Indicator	Failed banks					Operating banks				
	Obs	Mean	SD	$Q_{1\%}$	$Q_{99\%}$	Obs	Mean	SD	$Q_{1\%}$	$Q_{99\%}$
Panel 1: Basic set										
Capital	517	0.01	0.03	–0.08	0.09	382067	0.12	0.08	0.05	0.47
NPL	517	0.16	0.10	0.00	0.45	379208	0.02	0.03	0.00	0.13
Liquidity	517	0.07	0.07	0.00	0.31	382549	0.16	0.13	0.00	0.59
TAgrowth	516	–0.12	0.16	–0.46	0.48	343296	0.12	7.13	–0.19	1.10
ROA, %	517	–6.56	6.46	–28.1	0.97	382067	0.78	3.58	–5.04	3.63
Size	517	12.4	1.39	9.68	16.35	382549	12.0	1.38	9.35	16.58
Panel 2: Additional set										
LnsF	517	0.08	0.07	0.00	0.35	382549	0.09	0.07	0.00	0.34
DepH	517	0.09	0.13	0.00	0.60	382067	0.03	0.15	0.00	0.34

Note: In our description all variables are taken with 1 quarter lag prior to failure (for dead banks) and as of 2016Q2 (for alive banks). The notations: LnsF – the ratio of all commercial and industry loans to total assets, DepH – the ratio of brokered deposits to total assets.

At first sight, failed and operating banks are of similar size. Differences in most meaningful financials are consistent with our expectation. Alive banks report much higher capital adequacy, rely on expensive sources of funding (brokered deposits), and hold assets of poor quality.

3.2.2 The data on Russian banks

The sample size of operating banks (i.e. those still keeping the license and reporting the Forms) is 925 at the beginning of the sample period and is 531 at the end. These banks cover approximately 95% of the total banking system assets. The sample size of failed banks is 378.

Table 4 contains the descriptive statistics on the HNCs of failed Russian banks averaged across the sample period.

Table 4: Descriptive statistics for failed Russian banks: Hidden negative capital (HNC)

Absolute value of HNC as:	Mean	SD	Min	Max	Percentiles		
					P25	P50	P75
% of remaining total assets	3.564	8.725	0.008	105.21	0.330	0.961	3.105
% of remaining total liabilities	0.310	0.135	0.008	0.498	0.199	0.329	0.431

On average, the HNC of the failed Russian banks is dramatically large being equal to more than a half of their liabilities (or, equivalently, 3 times their capital reported prior to failure). First, we compare the samples of banks that failed with non-zero HNC and operating banks (in the end of 2017). We conduct this preliminary statistical exercise with two sets of variables: the basic one includes the variables that were used in the previous research on negative capital and the additional one covers the two variables needed to test the product mismatch hypothesis (see Table 1).

Table 5: Descriptive statistics for Russian banks: Failed vs. operating banks

Indicator	Failed banks					Operating banks				
	Obs	Mean	SD	$Q_{1\%}$	$Q_{99\%}$	Obs	Mean	SD	$Q_{1\%}$	$Q_{99\%}$
Panel 1: Basic set										
Capital	377	0.18	0.13	0.02	0.64	141873	0.22	0.17	0.04	0.86
NPL	374	0.05	0.08	0.00	0.28	137148	0.04	0.09	0.00	0.45
Liquidity	377	0.02	0.06	0.00	0.28	141873	0.02	0.05	0.00	0.22
TAgrowth	374	0.25	0.78	–0.56	3.89	126819	0.37	3.75	–0.56	3.37
ROA, %	376	–1.54	16.2	–61.1	37.8	138569	1.81	9.20	–17.0	22.6
Size	377	1.35	1.63	–1.95	5.82	141873	0.95	2.04	–3.91	6.43
Panel 2: Additional set										
LnsF	377	0.44	0.24	0.00	0.90	141873	0.34	0.21	0.00	0.81
DepH	377	0.39	0.25	0.00	0.81	141873	0.27	0.21	0.00	0.73

Note: In our description all variables are taken with 3 month lag prior to failure (for dead banks) and as of 2017M12 (for alive banks). LnsF – total loans to firms over total assets, DepH – total deposits of households over total assets.

Several evidences emerge on the descriptive statistics. First, the failed banks are smaller on average in terms of assets than the alive 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 4 times higher compared to that of the alive banks. Third, the reported quality of total loans to firms and households is surprisingly higher and less volatile for the dead banks compared to the alive banks. This gives the first notion of balance sheet falsifications and misreporting. Fourth, in terms of liquidity, dead banks are unsurprisingly worse than the alive banks. Fifth, returns and capital reported are also lower for the dead banks compared to the alive banks. In fact, 3 months prior to failure, banks already had visible negative ROA.

Switching to the variables that we use to test the product mismatch hypothesis also delivers 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 alive banks. As we already stated in the Introduction, this situation must significantly reduce the banks' margins to the relationships between the observed interest rates on deposits and loans to households and firms (see Table 1). Given the latter, the normal situation would be to attract deposits and grant loans to the same type of agents, not different. We leave the discussion on the reasons why the dead banks behave in this mannerand just emphasize that this product mismatch is among the main driving force of their negative ROA discussed above. Thus, product mismatch can kill the bank.

3.3 Data cleaning

We apply the same cleaning procedure to U.S. and Russian datasamples. In order to fight with outliers among operating banks we remove:

- observations outside the 5^{th} and 95^{th} percentiles for ROA.
- observations outside the 1^{st} and 95^{th} percentiles for the equity ratio (Capital).
- 5% of observation with the highest growth rate of total assets (TA growth) .

Data for failed banks is left without changes.

4 Estimation results

In this section, we present the estimation results for the benchmark and for the mismatch augmented versions of the Heckman selection model (equations 1–2). Section 4.1 describes the results for U.S. banks and Section 4.2 for Russian banks. Further, Section 4.3 presents the estimation results for the heterogeneous mismatch specification of the Heckman selection model (4–5) comparatively for U.S. and Russian banks. Finally, Section 4.4 describes the results of the out-of-sample exercises for both banking systems. For the sake of simplicity, we disclose the detailed estimation results for the last periods only.

4.1 Hidden negative capital: The case of U.S. banking system

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 negative capital of U.S. banking failures. The second pair of columns add the mismatch variable to the set of HNC determinants. In the both models, the selection equation enters the first column and the outcome equation enters the second column. All models are estimated on the bank-level data for 2016Q2 (still operating banks) and the historical failure events accumulated from 2004 to 2016Q1 (failed banks).

The estimation results suggest that the benchmark specification supports the previous findings on the cost of U.S. banking failures. That is, first, the reported equity capital tends to decrease both the probability of selection and the size of HNC conditional on being selected, in line with the previous studies on negative capital (Kang et al., 2015; Cole and White, 2017, etc.) 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 prefious findings. Third, liquidity seems have expected negative effects on both the probability of selection and the conditional size of HNC, but these effects are not significant. 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

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

	Benchmark Heckman		Heckman with Mismatch	
	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: In our basic regression analysis all variables are taken with 1 quarter lag prior to failure. ρ – the correlation between regression errors in the selection and the outcome equations

***, **, * – a coefficient is significant at the 1%, 5%, 10% level, respectively. Standard errors appear in the brackets under the estimated coefficients.

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 until they get positive dividends). Fifth, the growth of total assets increases both the 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, bank's size (log of assets) has negative effect on the probability of being selected, thus, supporting the too-big-to-fail view on U.S. banks (O'Hara and Shaw, 1990). In our setting this finding imply 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 falsify their accounts).

The further estimation results show that when we add the mismatch variable into the Heckman selection model we do find support to our mismatch hypothesis while we still get the same effects from the benchmark control variables as above. That is, the mismatch matters: the higher the mismatch (i.e. more deposits attracted from households and more loans granted to non-financial firms), the higher both the probability of selection and the conditional size of HNC. Turning to the magnitudes of implied economic effects, consider the increase in mismatch variable by its 1 sample standard deviation. Approximately, this corresponds to an increase in both components of the mismatch variable by 4–5 percentage points (i.e., the ratio of household deposits to liabilities and the ratio of corporate credit to assets). The model suggests that the resulting increase in the probability of selection is huge and equals 25 percentage points and the increase in the size of HNC is 6.1 percentage points.

Notably, the use of the mismatch variable caused a doubled the correlation between the selection and outcome regressions' errors (in absolute terms), given almost the same precision of the estimate. This implies that the mismatch variable crucially helps in the identification of sample selection bias. Therefore, we conclude that mismatching between different types of clients may provide a valid signal of deteriorating banking stability and should thus be controlled for by financial regulators in advance.

From the technical standpoint, the use of the both benchmark and extended versions of the Heckman selection model for U.S. banks is justified by the presence of statistically significant correlation between the regression errors in selection and outcome equations.

4.2 Hidden negative capital: The case of Russian banking system

Now we can 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 the Heckman selection model appear in Table 7. 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.

The estimation results obtained for the Russian banks share many similar features with those for U.S. banks; though, some notable differences also emerge. Most importantly, Heckman selection models are identified and the mismatch hypothesis cannot be rejected. That is, the correlation between the selection and outcome regressions' errors is significant, indicating the presence of sample selection bias in the data. Further, greater mismatch leads to higher probability of selection and higher conditional size of banks' HNC, exactly as in the case of U.S. banks.

Among other similar features are the effects of capital, ROA, the growth of assets, and size. However, capital and ROA significantly affect only the selection process, what makes a difference in comparison with the U.S. case. Similarly, liquidity does not effect the conditional size of HNC in our data for Russian banks, as was true for U.S. banks. This has an interesting implication from the standpoint of capital–liquidity nexus: that is, holding more or less liquidity

Table 7: Heckman selection models for Russian banks

	Benchmark Heckman		Heckman with Mismatch	
	Selection	Outcome	Selection	Outcome
Mismatch			2.315***	0.103*
			(0.463)	(0.061)
Capital	–2.802***	–0.064	–1.857***	–0.014
	(0.459)	(0.060)	(0.489)	(0.065)
NPL	–1.945***	–0.261***	–1.478***	–0.231**
	(0.491)	(0.093)	(0.480)	(0.091)
Liquidity	–2.183**	–0.069	–1.734*	–0.041
	(0.945)	(0.122)	(0.944)	(0.121)
ROA	–4.542**	–0.242	–4.840***	–0.258
	(1.800)	(0.193)	(1.774)	(0.195)
TA growth	0.861***	0.022**	0.828***	0.020*
	(0.187)	(0.011)	(0.182)	(0.011)
Size	–0.321***		–0.308***	
	(0.036)		(0.019)	
Constant	1.100***	0.301***	0.549***	0.277***
	(0.148)	(0.019)	(0.182)	(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: In our basic regression analysis all variables are taken with 3 months lag prior to failure. ρ – the correlation between regression errors in the selection and the outcome equations

***, **, * – a coefficient is significant at the 1%, 5%, 10% level, respectively. Standard errors appear in the brackets under the estimated coefficients.

is not important for those still operating banks that are already with HNC. Importantly, as in the case of U.S. banks above, we do find the support to the too-big-to-fail view for Russian banks using the size variable. It is also possible that, when increasing their size, Russian banks are more successful in hiding their problems from the Central Bank of Russia, as are the U.S. banks.

Conversely, the NPL variable is strongly significant and appears with a reverse sign both in the selection and the outcome equations, dissimilar to results for U.S. banks. This could reflect the pattern that operating banks with large disclosed non-performing loans may have proportionally smaller conditional sizes of HNC.

For the computation of the magnitudes of implied economic effects, consider again, as in the case of U.S. banks, an increase in mismatch variable by 1 sample standard deviation (or, equivalently, by 14–15 percentage points the both components of the mismatch variable). Then, the model for Russian banks suggests that the resulting increase in the probability of selection is rather notable and equals 5 percentage points and the increase in the size of HNC is also visible and equals 9.6 percentage points. Therefore, the probability effect is 5 times smaller compared to that obtained above for U.S. banks and the outcome effect is 1.5 times larger than that in U.S. case.

4.3 Heterogeneous mismatch effect: smaller vs larger bank groups

Having established the existence of the mismatch effect, we now turn to the analysis of possible heterogeneity of this effect for banks in different size clusters. The estimation results appear in Table 8. 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; Panel 2 of the table — coefficients on respective variables from the outcome equation. The first two columns of the table describes 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.

The estimation results suggest that the mismatch effect is indeed heterogeneous and that it differs to some extent between U.S. and Russian banks. First, the results on the selection equation for U.S. banks indicate that the average mismatch effect is due to small banks only. Specifically, while the mismatch effect is positive for the small banks, it is insignificantly negative for large banks. We can 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 corporate loan market (see Chodorow-Reich (2014) for the details). Additional confirmation to this conclusion comes from the outcome equation that shows the absence of any significant influence of mismatch on the conditional size of HNC for large 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 much lesser extent. In this sense, the obtained differences in estimation results are rather expected.

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

Table 8: 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*** (22.32)		2.315*** (0.463)	
Mismatch×Small		73.93*** (23.00)		2.278*** (0.464)
Mismatch×Large		–278.81 (196.23)		5.191** (2.098)
Size	–0.731*** (0.151)	–0.656*** (0.165)	–0.308*** (0.037)	–0.328*** (0.040)
Panel 2: Outcome equation				
Mismatch	0.599*** (0.208)		0.103* (0.061)	
Mismatch×Small		0.593*** (0.208)		0.101* (0.061)
Mismatch×Large		4.286 (2.800)		0.189 (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: ρ – the correlation between regression errors in the selection and the outcome equations.

***, **, * – a coefficient is significant at the 1%, 5%, 10% level, respectively. Robust standard errors appear in the brackets under the estimated coefficients.

Finally, we are interested in to what extent the in-sample predictions of the size of HNC are different when we include or exclude the mismatch variable and how it differs between the U.S. and Russian banks. Given the estimated models described above, we predicted the modeled values of HNC without and with the heterogeneous mismatch effects and then plotted the densities for the both banking systems. Results appear in Appendix B and they suggest that the role of mismatches in the in-sample prediction is visible for the 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 changes only slightly for U.S. banks. This is another reflection of the bank-based vs.

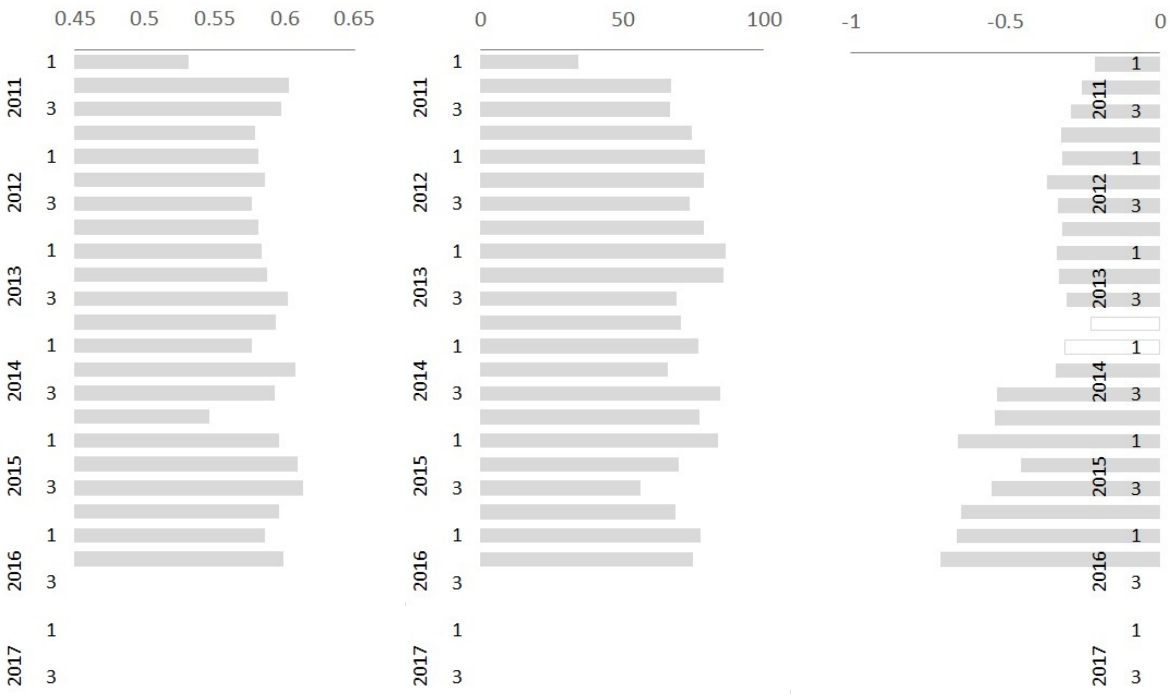
finance-based banking system features.

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 2011–2017 in the United States and Russia. The out-of-sample forecasting horizon is fixed to be 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 hidden negative capital in banks h period ahead (at time $t+h$). This estimation procedure allows us to mitigate class-imbalance problem in 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 two countries. In every step, the parameters of the model are re-estimated.

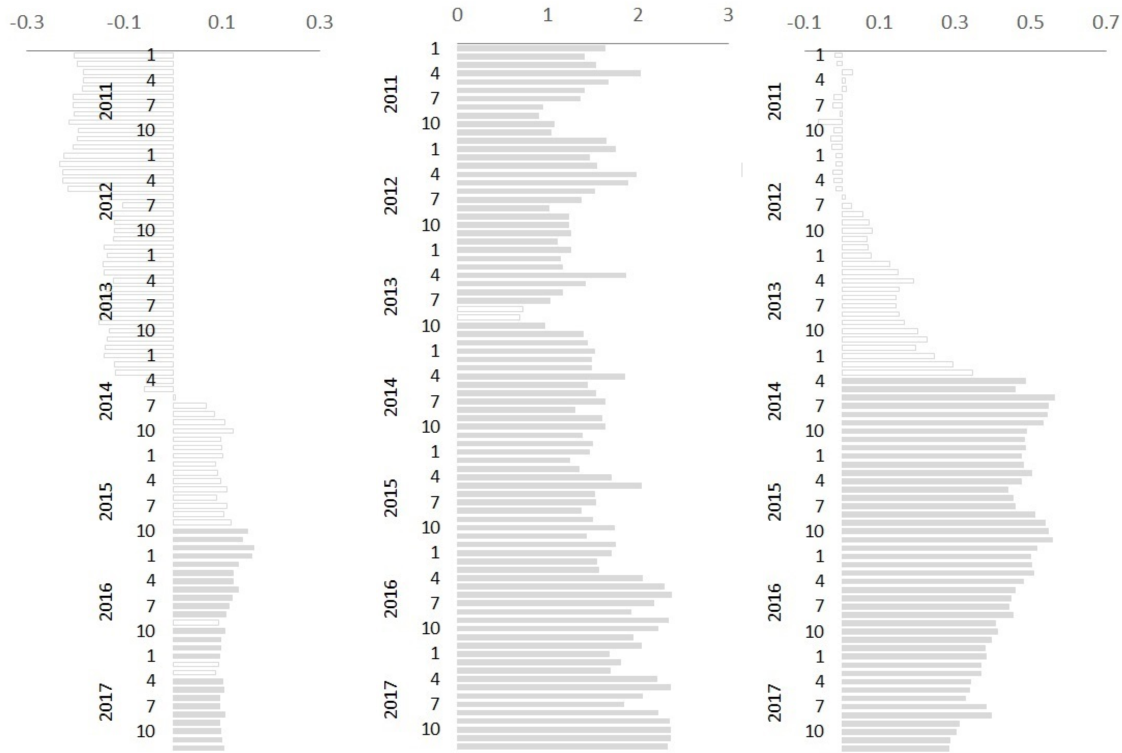
The main coefficients of our interest are Mismatch in the Heckman selection 1 and outcome 2 equations (verify the mismatch hypothesis) and ρ, a correlation between errors of selection and outcome equation (justifies the use of the Heckman selection model). We report the key results on every step of forecasting (Figures 1 and 2). As we can see, for the United States Mismatch is permanently significant and sample selection remains in data throughout the steps of forecasting (correlation ρ is statistically significant). It is notable that the only period when sample selection is not discovered is end of 2013 beginning of 2014. This could be attributed to appointment of Jerome Powell as the Chair of the Federal Reserve and the respective transition period. For the case of Russia, sample selection arises in data soon after the appointment of Elvira Nabiullina as a head of the Central Bank of Russia in mid-2013. She toughened supervision and started to clean up the Russian banking sector. Mismatch coefficient is positive and significant in the selection and outcome equations, which confirms our mismatch hypothesis.

Table 9 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 prediction the ratio of hidden negative capital to liabilities (as we define the size of HNC in equation 2). Although similar in the spirit, MSFE is more sensitive to large errors. We compare the performance of the suggested Heckman model with simple OLS regression that ignores the selection problem in 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 one. Second, Heckman model generate more precise predictions of negative capital in both countries, that confirms the importance of sample selection in data. Finally, t-test for the difference in mean errors votes for the statistically significant predominance of the Heckman sample selection model.



(a) Mismatch in selection equation (b) Mismatch in outcome selection (c) Correlation ρ

Figure 1: 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 2: 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 9: 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; second — 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 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 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 could just capture the situations when 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 capital adequacy ratio in the banking sector at time t without 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 the 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 the HNC formation in 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 as a dependent variable in the outcome equation we consider relative, not absolute, size of HNC. 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: NHC 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 in small banks only 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, oppositely 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 finance-based, larger banks are likely to expand their activities beyond the traditional banking, while smaller banks have less 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.

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-Reicn, 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.

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.

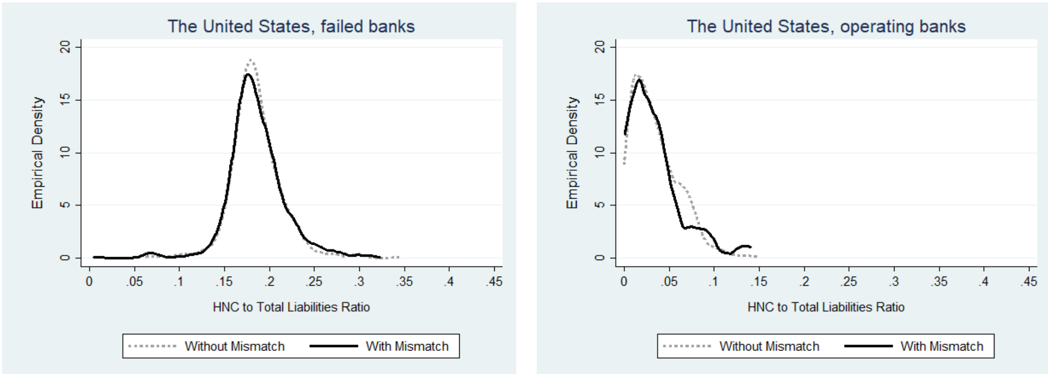
Wheelock, D. and Wilson, P. (2000) "Why do banks disappear? The determinants of U.S. bank failures and acquisitions", *The Review of Economics and Statistics*, Vol. 82, pp. 127–138.

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

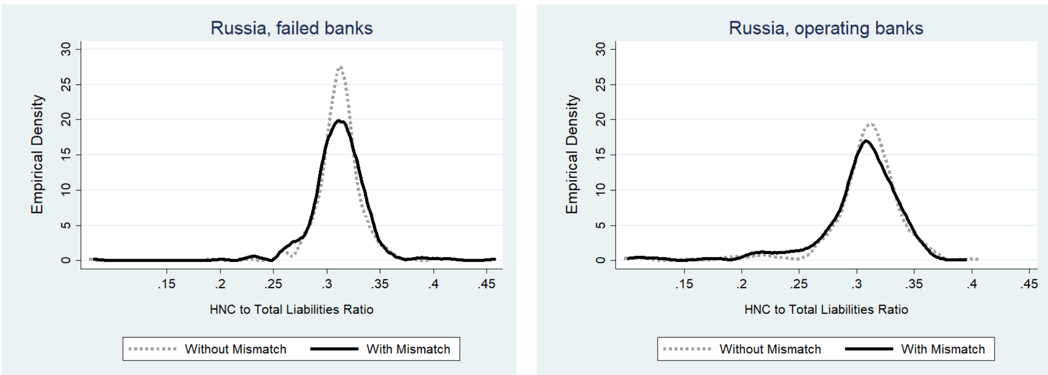
Appendix B. HNC in-sample predictions: The role of heterogeneous mismatches



(a) Failed banks

(b) Operating banks

Figure 3: Density of HNC predictions with and without mismatch: U.S. banks



(a) Failed banks

(b) Operating banks

Figure 4: 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: All underlying regressions include the mismatch variable. 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.

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

Note: ***, **, * – a coefficient is significant at the 1%, 5%, 10% level, respectively. Standard errors appear in the brackets under the estimated coefficients.

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

	Forecasting horizon			
	Basic: 1Q	6M (2Q)	9M (3Q)	12M (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: ***, **, * – a coefficient is significant at the 1%, 5%, 10% level, respectively. Standard errors 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	II	III	IV
Panel 1: Selection equation				
Mismatch	78.78*** (22.32)		2.315*** (0.463)	
Mismatch×Small		73.42*** (20.59)		2.293*** (0.446)
Mismatch×Large		–283.24*** (84.89)		5.147*** (1.913)
Panel 2: Outcome equation				
Mismatch	0.599*** (0.208)		0.103* (0.061)	
Mismatch×Small		0.467*** (0.170)		0.063 (0.060)
Mismatch×Large		5.243 (4.516)		0.613*** (0.134)
System Capital		–1.908** (0.837)		–2.448*** (0.646)
N obs.	2371	2371	799	799
N censored	1867	1867	440	440
N observed	504	504	359	359
ρ	–0.707***	–0.552**	0.28*	0.334**
Log Likelihood	532.8	536.9	–227.4	–216.5

Note: ρ – the correlation between regression errors in the selection and the outcome equations

***, **, * – a coefficient is significant at the 1%, 5%, 10% level, respectively. Robust standard errors appear in the brackets under the estimated coefficients. System Capital$_{i,t}$ is computed as an asset-weighted ratio of capital to assets in the banking system (excluding bank i) at time t.