MASTER-ESSAY

The Determinants of Bank Failure: the Evidence from Ukraine and Russia

Authors: Oksana Bagatiuk, Valeriia Dzhamalova

Supervisor: David Edgerton

Lund University, Department of Economics

Seminar date: June, 4 2009

Abstract

Title: The Determinants of Bank Failure: the Evidence from Ukraine and Russia

Seminar date: 2009-06-04

Authors: Oksana Bagatiuk, Valeriia Dzhamalova

Advisors: David Edgerton

Key words: financial ratios, bank failure prediction, cross-country model, linear probability model, binary choice models.

Purpose: To investigate which financial ratios are significant in the prediction of bank failure for the sample of Russian and Ukrainian banks and to see if there any similar ratios, which drive the bank to bankruptcy in both countries for the first decade of this century.

Methodology: As an instrument for our analysis we use the linear probability and the binary choice models. They are used to analyze the significance of different financial ratios drawn from the publicly available annual balance sheets for the period 2002-2008.

Theoretical perspective: The theoretical framework is the previous research in the field of bank failure prediction and credit-risk management methodology.

Conclusions: For the Ukrainian sample the capitalization and profitability measures influence the bank failure negatively, as well as a size of the bank. For the sample of Russian banks, the return on assets and the loans to total assets ratios are found to have a negative influence on the default, in contrast to the capitalization, which positively influence the bankruptcy according to the estimated coefficients. Moreover, we conclude that despite some similarities, the differences exist in determinants of the banks to failure in two countries; hence the regulatory policy of central banks on the local levels still remains the most important for the healthy functioning of the banking sector.

Table of Contents

1. Introduction

1.1 Background

The bankruptcy of any company causes different reactions in society depending on its size and publicity. However, the insolvency of financial institutions and banks specifically always is alarming. As a default of the bank can cause insolvency of healthy companies, displeasure of private customers and as a consequence distrust in the local financial system. That is why regulation and overseeing of the banking sector is a major concern of the government of any particular country. Moreover, on the international level, the existence of such organization as the Bank of International Settlements, which "serves as a bank for central banks,"[1] emphasizes the importance of banking control not only on the national level but among the countries as well. One of the major objectives of this organization is the promotion of monetary and financial stability, which is realized through designing the international standards on capital adequacy; the Core Principles for Effective Banking Supervision; and the Concordat on cross-border banking supervision.

Among the researchers the question of identifying the factors, which influence the bank failure on the local, as well as on the international levels is also brought up. For instance, Mongomery, Santoso et al. (2005) find evidence in favour of a cross-country model for Indonesia and Japan in comparison with domestic ones[2]. They use a logit analysis on financial ratios of commercial banks. Other economists study macroeconomic variables, which can be associated with systemic banking crises in different countries. Among them Demirguc-Kunt and Detragiache (1998) who use multivariate logit for identifying if such factors as inflation, GDP growth, interest rate etc. can influence the financial stability of the banks. Using a large sample of developed and developing countries in the period from 1980 to 1994, they find that where the macroeconomic environment is weak there is a tendency of banking crisis eruption. Many other studies are concentrated on the micro-level within one country. For example, Whalen (1991) examines the usefulness of proportional hazard model (PHM) as an early warning tool in prediction of bankruptcy of the bank. Taking a sample of more than 1500 American banks for the period of 1987-1990, he finds that the PHM constructed even with quite a small number of explanatory variables can serve as an effective early warning tool. Some authors, such as Gonzalez (1999), Peresetskiy (2004) recognize both micro- and macro-components as important identifiers of the

[1] http://www.bis.org/stability.htm

[2] In this essay we use the term domestic model when we analyze each country in turn and the term cross-country model when the pooled variables of two countries are analyzed together.

distress of the banking sector. The details of these studies as well as a review of other literature in this field are presented in Section 3.1.

To continue with clarifying the background for our topic we think it is necessary to give some facts which underline the important role of stability of East European banking sector in the stability of the whole European financial system. When the Soviet Union broke up, a new huge financial market was opened. At the end of the nineties of the last century several large western financial groups have entered it. For instance, *The Wall Street Journal* (February 19, 2009) reports that nearly 80 % of the profit of Austrian Raiffeisen Zentralbank Osterreich AG over the past decade came from that region. Austrian banks assets in the region came to about two-thirds of the country's gross domestic product. But at the same time a new profit was associated with a new risk. And in February 2009 Moody's Investors Service Inc. notification about possible downgrading of the banks with large investment in Eastern Europe has provoked a sharp drop of stock prices of these banks.[3] This news was on the front pages of many European newspapers and the reason is quite obvious. The banking crisis in Eastern Europe and in ex-Soviet countries in particular, may cause considerable credit losses not only inside the region, but also in the European Union. Hence, the timeliness and efficiency of the government regulation of the banking sector are very crucial for the current situation as well as for the future trust a region enjoys among investors.

Historically and geographically, Russian and Ukrainian economies are closely related. Moreover, the political decision made in one country can provoke the sharp changes in economic growth of another country. This argument is illustrated by the example of well-known gas-conflict. Hence, the State Statistic Committee of Ukraine[4] reports that Ukraine's industrial output in January 2009 dropped over a 30% compared with a year earlier. And this is partly a result of a cut gas supplies. As Reuters states, "Much of the industry [chemical and petrochemical] ground to a halt…when Russia cut gas supplies to Ukraine over a pricing dispute".[5] Furthermore, the Russian capital comes up to 21% of the total amount of foreign banking capital in Ukraine. And the most important that, in words of Ukrainian media,[6] many of these banks have the "Kremlin's tracks". Thus one of them is under control of Russian government, another is subsidiary of central Bank of Russian Federation and lastly, the Vnesheconombank is a state corporation, which has the prime-minister of Russia as a head of supervisory board.

3 Champion, M., Bart, K., Connaghan C. (2009-02-18), "Eastern Europe Shakes Banks", *The Wall Street Journal (Europe)*, vol. XXVII No.13.
4 http://ukrstat.gov.ua
5 http://in.reuters.com/article/asiaCompanyAndMarkets/idINLH37296720090217
6 http://www.epravda.com.ua/publications/4a0914d48ee17/

Consequently, the amount of Russian government's capital in Ukrainian banking sector, as well as interdependence of economic processes in two countries let us assume that there are some similar factors which can drive the banks to bankruptcy in Russia and Ukraine . And possibly, the cross-country analysis of these factors can improve the regulatory arrangements in the field of banks bankruptcy.

1.2 Specification of the problem

Taking into consideration the interdependence of financial sectors stability of the former Soviet republics and other countries outside the region, common political and economic factors, driving the development of Russian and Ukrainian economies; we find it interesting to see if there are any similar micro-level financial factors, which are driving the banks in Ukraine and Russia to bankruptcy. Moreover, as far as we know there is not much existing research in the field of banks failure prediction modeling for the Ukrainian case and our ambition is to fill up this gap.

1.3 Purpose

The purpose of this study is dual: 1) to investigate which financial ratios are significant in the prediction of bank failure for the sample of Russian and Ukrainian banks for the first decade of this century; 2) to see if there are any similar factors, which drive the bank to bankruptcy in both countries. As an instrument for our analysis we employ the linear probability and the binary choice models. They are used to analyze the significance of different financial ratios drawn from the publicly available annual balance sheets for the period 2002-2008[7]. The choice of the examined financial ratios is limited to the extent of availability of the data. Available data include the capital of the bank, total assets, deposits, loans and profit after tax.[8]

1.4 Limitations

The primary limitation of the study, which we cannot avoid, is the quality and transparency of both Russian and Ukrainian accounting data. Unfortunately, in ex-Soviet Union countries it is still possible that all financial statements and especially the data on profits are easily manipulated. And of course, if the initial data is drawn from sources which do not reflect the true economic reality, then the estimations and conclusions based on this data are questionable.

Also despite many similarities, we realize the difference that exists between the economic and political conditions in each of the two banking systems. Future research should try to account for the regional differences and macroeconomic environment.

[7] The sample period differs for Ukraine and Russia; the details are described in Section 4.1.
[8] The details on selection of variables are presented in Section 4.2.

1.5 OUTLINE OF THE PAPER

Section 1 introduces the reader to the background of the problem raised in the essay. The specification of the problem, the purpose of the study and the limitations are stated. The current situation and the development dynamics of Ukrainian and Russian banking sectors are described in Section 2. A short review of the previous research in the field of bank failure prediction models, a description of the theory and method is presented in Section 3. In Section 4 a discussion of the estimation results and performance of the models are described. The conclusions are provided in Section 5.

2. A short description of the Ukrainian and Russian banking sectors

To give the general idea about the objects of our analysis, in this section we present the current situation and the development dynamics of the two banking systems.

The history of Russian and Ukrainian banking systems reflects the processes and tendencies that have taken the place throughout the countries evolution. These two countries are closely related in many spheres and in the banking sector as well. Both are a two-tier system and consist of the central bank and commercial banks of various types and forms of ownership. Until 1991 Ukraine and Russia had a unified economic and financial system, but after the Soviet Union broke up a new spiral in the development of both banking sectors arose. According to the information from the annual reports of Russian and Ukrainian central banks,[9] the both banking sectors demonstrate a high rate of growth during recent times. Some of growth rates are presented in Table 2.1.

Table 2.1 Some Rates of Growth of Banking Sectors in Ukraine and Russia

Parameters[10]	Ukrainian banking sector	Russian banking sector
Annual bank capital grows %	76.0	57.8
Capital to GDP ratio, %	10.0	8.1
Annual assets grows, %	76.0	44.1
Profit grows, %	160.0	36.7

Thus, the growth is especially significant in Ukraine, where the banking system developed rapidly despite the political and economic problems at the end of 2004 and the slowdown of the economic processes thereafter. Since 2001 the banking assets increased almost by 50% in nominal terms. “In 2007, the Ukrainian banking system gained a record high profit in the history of the domestic banking system, which is 1.6 times higher than in 2006 and 12 times higher than the profit gained in the year 2001. As of 01.01.2008, return on assets (ROA) was 1.5 %, return on equity (ROE) - 12.67%.” [11] By 01.01.2008 the total number of operating banks was 175 in Ukraine and more than one thousand in Russia. Among them there are 47 and 202 banks with participation of the foreign capital for two countries respectively. Credit institutions in both countries were characterized by the growth in loans to nonfinancial enterprises and households.

9 The information is taken as of 01.01.08.
10 Grows rates are taken for 2007 compared to 2006.
11 http://www.bank.gov.ua

Unfortunately from the second part of 2008 the situation in the banking sectors of both countries has changed. Thus, for the first four months of 2009, 13 Ukrainian banks were taken under the temporary administration of National Bank of Ukraine because of their insolvency. The total expenses of banks for the first quarter of 2009 exceeds the incomes almost on 10 % (in comparison, for the same period in 2008 the income were greater than expenses on 13%). There are still not any publications about the defaults of Russian banks this year. However the Moody's Global Credit Research 2008 presents the following information: "Moody`s outlook for the Russian banking system is negative, reflecting our concerns with regard to the system's structural weaknesses which have become particularly apparent in the course of the crisis in the Russian stock market - and the possible impact of negative external factors. Thanks to their focus on the domestic market in terms of assets, the global credit and liquidity crisis has not resulted in any direct losses for the Russian banks. However, in terms of their liabilities, the impact of the crisis has been more significant, with reduced access to market funding a major driver of a slowdown in growth." [12]

[12] Moody's Global Credit Research, 2008

3. Theoretical Framework

3.1 Review of Previous Research

In this sub-section we give a short overview of previous research which has been done in the field of domestic and cross-country bank-failure prediction models.

The article which became a starting point of our essay is an article on a cross-country bank failure prediction model for Japan and Indonesia by Montgomery, Santoso et al. (2005). They use a logistic model (though they do not specify if they use a panel or a pooled logit) and a stepwise logistic model on 17 financial ratios for the entire population of banks in Indonesia in 1997-2003 and Japanese banks for period from 1978 to 2001. The authors construct two domestic and cross-country models and then compare their properties. Their domestic models show the importance of monitoring the portfolio of banks loans for prediction of bank failures. In both Japan and Indonesia they find such ratios as loans to deposits and loans to total assets statistically significant, in contrast with regulatory capital ratios, which are found not to be important predictors of bank failure for the period of the study. But the most important fact they present is that their cross-country model out-performs the domestic models in predictive power. This, in their words, can stimulate the regional cooperation on this issue.

A similar idea of identifying the factors, which influence banking crises in developed and developing countries is used by Demirguc-Kunt and Detragiache (1998, p.81). But in contrast to the previously presented article they analyze the macro-factors as the determinants of the banking crises for the range of countries from the list of International Monetary Fund for period 1980-1994. They use a pooled logit model in their study. And they find that “crises tend to erupt when the macroeconomic environment is weak, particularly when growth is low and inflation is high” as well as high real interest rate also indicates the problems in the banking sector. One more research on determinants of the financial crisis on the macro-level is made by Komulainen and Lukkarila (2003). These authors use a panel probit on aggregated data. Using 23 macroeconomic indicators as explanatory variables for several emerging market’s countries for the period 1980-2001, they conclude that such variables as unemployment, inflation and several indicators of indebtedness can help in explaining the currency crises. But this paper, as well as the two others mentioned above, does not concentrate on the econometric part of research and some details are not clear. In particular, researchers do not mention anything about possible problem of model misspecification.

However, among the literature on bank failure prediction models in the Russian case there are a number of articles, which implement a very thorough analysis as well as explain the ideas very clearly. For instance, Peresetsky et al. (2004) use binary choice models for the analysis of different ratios drawn from financial statements of the banks and also some macroeconomic variables. Firstly, the authors analyze the significance of 30 bank parameters using a pooled logit based on accounting data for 1569 Russian banks for a two-year period of crisis in Russia. Then they take quarterly balance sheet data for a five-year period, include macro-indicators and implement an analysis using a random effect probit on panel data. They also show that clustering (or grouping the banks by size or other characteristics) helps to improve the predictive power of pooled logit. Macroeconomic variables are shown to improve the model's performance. One more paper on failure prediction of Russian banks is written by Konstandina (2006). She uses a multivariate panel logit (but also does not indicate if it is random or fixed effect) and proportional-hazard models for two panels of Russian banks (annual and quarterly basis). Both models attempt to capture the effect of micro and macro factors on bank failure. In contrast with Peresetsky et al. (2004) this study does not confirm the significance of macroeconomic variables.

As to the research on bank failure prediction models for Ukrainian case, we have not found so many empirical studies. This is probably connected with the fact that not so much research in this field exists and even less is available online. We can mention only two Master theses of students at the Economic Education and Research Consortium in Kiev, who analyzed this question. Popryga (2001) use a limited dependent variable model for analyzing the financial ratios of 111 Ukrainian banks for the period from 1995 to 1996. Her main result is that some traditional variables (such as return on equity) are not significant for soundness of Ukrainian banks. Location of the bank and the duration of its business do not play a role in this issue. One more similar study is made by Nikolsko-Rzhevskyy (2003). He performs various methods of analysis of bank failures during 1998-2003. He finds that one of the major factors which have the influence on bankruptcy of banks is their size.

3.2 THEORY

Below we present and compare the theories and models, which are designed to analyze the default risk of companies. The motivation for why we choose the analysis of financial ratios as a basic tool for our study is also described in this section.

According to Crosbie and Bohn (2005), default risk is the uncertainty surrounding a firm's ability to service its debts and obligations. A plethora of approaches is developed to model this event and to control it. In Table 3.1 a short description of four popular methods of assessment of the default probability is presented. All of them are used for the credit-risk management inside banks, but they are also applicable to valuation of commercial banks. The table is based on theories presented in the book of Saunders and Allen (2002) “Credit Risk Measurement.” Table 3.1 illustrates that taking into consideration the necessary initial data, simplicity and transparency of the analysis; we find the credit scoring system is the most appropriate for our analysis. However, as we do not score any credits here, but rather evaluate the significance of different financial ratios in prediction of the bank failure, hereafter we will call this approach a financial ratio analysis. According to Saunders and Allen (2002, p.20), the idea behind this method is “to pre-identify certain key factors that determine the probability of default and combine or weight them into a quantitative score”. One of the first analyses of financial ratios for predicting the bankruptcy of companies is an Altman Z-score model. Altman (1968) use a multiple discriminant analysis (MDA) to analyze a set of financial ratios for a grouped data set of failed and surviving firms. In his paper the discriminant function has the following form:

$$Z = v_1x_1 + v_2 x_2 + \dots + v_nx_n \qquad (3.1)$$

Where v_1, v_2, ... v_n are discriminant coefficients and x_1, x_2, . . . x_n are independent variables, representing different financial ratios. According to Altamn (1968, p.592) this function “transforms individual variable values to a single discriminant score or Z value which is then used to classify the object” as default or non-default.

Table 3.1 The Comparative Analysis of Theories on Default Probability

Approach	Short Description	Advantages	Shortcomings	Could it be used in our study?
External Rating Systems	International credit rating agencies rate the creditworthiness and quality of the different institutions.	Assumed to be independent; recognized all over the world.	As the ratings are made on a payment base, their objectivity is questionable; rate's updates are often not appropriate.	No. Only few banks in Russia and Ukraine are rated.
Artificial Neural Networks	Simulates the human learning process. Based on analysis of input and output variables that can give the probability of default.	Does not depend on human factors, able to make educated "guess" when the information is incomplete.	Does not give transparent information about final output, and relative importance of input variables.	Possibly, but knowledge of sophisticated statistical techniques is required.
Credit scoring systems	Credit score are calculated based on certain input factors and their weights.	Simple, transparent, easy to implement even with a limited data set.	The linear form of underlying model does not describe the real relationship; cannot catch the negative dynamics of a quickly deteriorating company.	Yes. We have necessary input information and we are familiar with the techniques used for this analysis.
Contingent Claim Models	Calculate the probability of default (PD) for a particular firm within a certain period of time. The market value of assets and equity connected to each other through Merton's option pricing model.	In addition to PD the recovery rate can be calculated, application of stock prices allows catching the changes in financial conditions of the company.	Only those companies that are publicly traded can be analyzed.	No. We do not have enough initial information. As only few banks in Russia and Ukraine are quoted on the stock market.

From the set of 22 potential explanatory variables Altman constructs the following function of five financial ratios that are weighted by the estimated coefficients:

(3.2) $$Z=1.2x_1+1.4x_2+3.3x_3+0.6x_4+1.0x_5$$

where X_1 = working capital/total assets ratio;

X_2 = retained earnings/total assets ratio;

X_3 = earnings before interest and taxes/total assets ratio;

X_4 = market value of equity/book value of total liabilities ratio;

X_5 = sales/total assets ratio.

If the value of the z-score is below the critical value (originally in Altman's study it is equal to 1.81), then the firm is classified as bankrupt. A number of criticisms about Altman's Z-score are raised in the economic literature, among them Gharghori et al. (2006, p.208) emphasize two problems with accounting-based measures of default. The first is that financial statements report a firm's performance for the past period, which is inconsistent with a "forward-looking measure of default risk". And the second problem is that we can only guess intuitively which ratios should be used in predicting the default since no precise theory exists.

Nevertheless we decided to use this model as a basic for our study because despite the existence of more sophisticated techniques, the financial ratio analysis is still one of the simplest, most understandable and transparent. Moreover, it is widely applicable in the modern regulatory policy of central banks; hence the conclusions and suggestions made in our essay can be more appropriate for the specialists of these organizations than conclusions based on other unfamiliar to them techniques. For instance, Halling et al. (2006) describe that the regulatory function of many central banks is realized through on-site and off-site inspection of commercial banks. As on-site inspection is usually very costly and takes a considerable amount of time, the decision of performing it for a particular bank is based on the analysis of information, available to supervisors. The Russian and Ukrainian central banks use balance sheets and income statements for the evaluation of financial stability of commercial banks. In calculations of the ratios drawn from the financial statements, they use either their own methodology or methodology developed by the Bank of International Settlements, and the International Monetary Fund. According to the information posted on the web-pages of the central banks of analyzed countries, all indicators of financial stability of commercial banks can be placed in one of several groups: the indicators of the adequacy of a capital, the indicators of profitability, liquidity indicators, and indicators of the bank's assets structure. Aziz et al. (2006, p.23) have also given the empirical evidence for the argument that the financial ratios analysis still remains the most popular technique. They examine 46 articles, which report 89 empirical studies of corporate bankruptcy prediction. They find that "more than 60 per cent of the studies used financial ratios as the only explanatory variables, about 7 per cent used cash flow information while the remaining 33 per cent employed a mix of financial ratios and other variables (including macroeconomic, industry-specific, location, and other firm-specific variables)".

However, different methods can be employed for evaluation the financial ratios of interest. The descriptions of the methods, which can be used in this type of analysis, are presented in the next sub-section.

3.3 METHOD

3.3.1 SELECTION OF THE APPROPRIATE MODEL

The following three methodological forms are the most commonly used for the analysis of financial ratios: (1) binary choice models (probit and logit) (2) multiple discriminant analysis, (3) proportional hazard models. Returning to the statistics reported by Aziz et al (2006) we can say that more than 30 per cent of the research examined by them use multiple discriminant analysis, while another 21 per cent prefer the logit model. A total of 77 per cent of all studies on corporate bankruptcy prediction use statistical models. The ranking of relative performance of statistical models suggests that MDA and logit models can be more reliable than other statistical, artificially intelligent expert system (AIES) and the theoretical models.

Different authors criticize or defend the properties of the statistical models mentioned above. For instance, Whalen (1991) emphasizes that such models as discriminant analysis, logit and probit are designed only to model the probability that a bank with a given set of characteristics will be ranked as default or non-default. However, these models cannot predict the time, when the event will happen. In contrast to them the proportional hazard model allows one to estimate the time to failure, moreover it does not make any assumptions about the distributional properties of the data, which can be easily violated. Nevertheless, existing research gives different empirical evidence concerning the benefits of binary choice and time survival models. For example, Laviola et al. (1999) in their study of fragility of Italian banks conclude that the proportional hazard model outperforms logit and probit, though Lee and Urrutia (1996) find that logit and hazard models differ in number of significant coefficients and their use should be combined. The choice between binary choice models and MDA is not clearly identified among the researches. Thus, Canbas et al. (2005) use both types of these models in the prediction of banks failures for the Turkish case and find that in some instances, MDA can give a more accurate prediction than logit or probit. The study of Kim et al (2006), however, illustrates that logit has better performance than discriminant analysis.

The statistics presented above are based on articles, which predict the corporate bankruptcy for different type of companies. However, if we consider bank failure prediction models in particular, to our knowledge, binary choice models still remain the most popular for modeling the defaults of banks. Consequently in this essay we mostly focus our attention on binary response models (the linear probability model is also implemented, the motivation is discussed below).

3.3.2 BINARY CHOICE AND LINEAR PROBABILITY MODELS

First of all we need to clarify that in bank failure prediction models the outcome is default or no default, in other words the dependent variable y takes the value 0 (no default) or 1 (default). We want to explain this event with the help of several explanatory variables x. Using matrix notation, a general model can be written as:

$$P(y_i = 1 \mid \boldsymbol{x_i}) = F(\boldsymbol{x_i}'\boldsymbol{\beta}) \tag{3.3}$$

Where $0 \leq F \leq 1$, means that F is a cumulative distribution function (c.d.f.). As we only have information on zeros and ones on the left-hand side of equation there has to be a large number of observations to estimate the function F .The simple linear regression is not the best choice here and the simplest motivation for this is that in linear regression models the estimated value of y is not restricted to be in the interval between 0 and 1. However, Wooldridge (2002) asserts that despite the weaknesses of linear probability models (LPM) for binary response they often can give good estimates of the partial effect on response probability. Thus if we set $F(\boldsymbol{x_i}'\boldsymbol{\beta}) = \boldsymbol{x_i}'\boldsymbol{\beta}$, LPM for binary response y is specified as:

$$P(y = 1 | x_i) = \boldsymbol{x_i}'\boldsymbol{\beta} + \varepsilon_i \tag{3.4}$$

with $E(\varepsilon|\boldsymbol{x}) = 0$ and $Var(\varepsilon|\boldsymbol{x}) = \boldsymbol{x_i}'\boldsymbol{\beta}(1 - \boldsymbol{x_i}'\boldsymbol{\beta})$.

According to Verbeek (2008, p.200), to overcome the problems of LPM, the binary choice models (or binary response, latent variable models) are used. They are "designed to model the "choice" between two discrete alternatives" (e.g. default/no default, married/unmarried etc). Most commonly in these models the cumulative normal function and the logistic function are the specified for F in expression (3.3). The difference between these two functions is marginal (both have an expectation of zero, but the variance of the logistic distribution is equal to $\pi^2/3$) and in most of the cases both give almost the same results for marginal effect and for the sign of parameters. If we choose the standard normal distribution as a function for F:

$$F(w) = \int_{-\infty}^{w} \frac{1}{\sqrt{2\pi}} exp\left\{-\frac{1}{2}t^2\right\} dt \tag{3.5}$$

this leads to the probit model, while the standard logistic distribution, given by

$$F(w) = \frac{e^w}{1+e^w}, \tag{3.6}$$

leads to the logit model.

Apart from the signs, the coefficients in binary choice models cannot be interpreted directly. Knowing the sign of the estimated coefficient we can judge if the effect of particular variable has positive or negative effect, but to find its magnitude, the marginal effect of changes in the explanatory variables should be calculated. The marginal effect Ψ_j of the i-th regressor is:

(3.7) $$\Psi_j(x_i) = \frac{\partial E(y|x_i)}{\partial x_{ij}} = \beta_j f(x_i'\beta)$$

Where f (s) = ∂F(s)/ ∂s is the p.d.f. Empirically, the marginal effect for the average observation is usually calculated. For our further discussion it is also important to underline that in binary choice models the ordinary residuals e_{oi} do not have the same meaning as in the linear regression. Instead the generalized residuals e_{gi} play an important role. The usefulness of generalized residuals derives from the fact that by multiplying them by each of the regressors, the score vector is obtained. This score vector then can be used in Lagrange Multiplier tests (the details of the tests used in this study are described in Section 3.3.5). The generalized residuals are derived from the first order condition of the maximum-likelihood estimates[13] and they may be regarded as orthogonal to regressors. In the Table 3.2 three types of residuals are summarized and $\hat{p}_i = 1 - F(-x_i'\hat{\beta})$ there, is the fitted probability, e_{si} is standardized residuals (fitted probability divided by theoretical standard deviation), which also will be used in the next chapters. The expressions in the table are taken from Eviews manual.

Table 3.2 Three Types of Residuals Mentioned in our Study

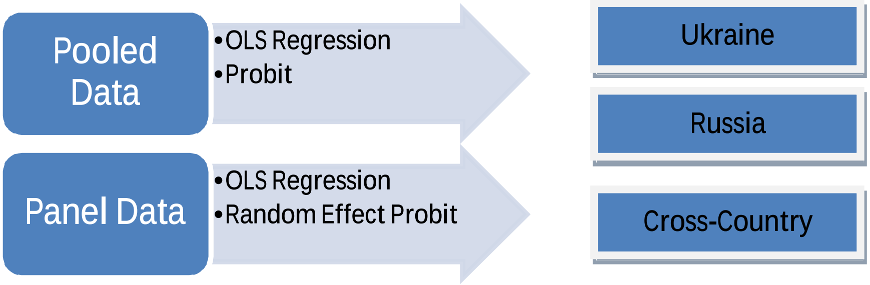
Ordinary	$e_{oi} = y_i - \hat{p}_i$
Standardized	$e_{si} = \frac{y_i - \hat{p}_i}{\sqrt{\hat{p}_i(1-\hat{p}_i)}}$
Generalized	$e_{gi} = \frac{(y_i - \hat{p}_i) f(-x_i'\hat{\beta})}{\hat{p}_i(1-\hat{p}_i)}$

3.3.3 Panel and pooled data representation

To make the further discussion clear, in the Figure 3.1 we illustrate the simplified scheme of our research. The terminology is described below.

[13] The maximum likelihood method is most commonly used for the estimation of the parameters in binary choice models. Our estimation will be implemented with the help of two programs: Eviews and Stata, both of them allow using this method.

Figure 3.1 The Simplified Scheme of Empirical Analysis



It is very important to define what we mean by pooled and panel data in this essay. According to Wooldridge (2002 p.128-129) the idea behind the pooling cross section over time is "that during each year a new random sample is taken from the relevant population. Since distributions of variables tend to change over time, the identical distribution assumption is not usually valid, but the independence assumption is. " He also states that in using pooled data it is useful to include the time dummies and even interact them with some explanatory variables to "allow the partial effect to change over time". In the panel data the same groups of cross-sectional units are followed over time, what allows to control for individual cross-sectional heterogeneity.

In our case we have the observations on 943 Russian banks for the period of six years and on 186 Ukrainian banks for period of seven years. In the Table 3.3 we present the classification of linear probability model used for in our estimation. The indices here have the following interpretation: j = Ukraine, Russia; $i = 1,\ldots, Nj$ (number of banks); $t = 1,\ldots, Tj$ (number of years). For the logit and probit the classification is the same.

Table 3.3 Classification of linear probability model

	Domestic	Cross-Country
Pooled	$y_{jit} = \alpha_j + x_{jit}^{'}\beta_j$	$y_{jit} = \alpha + x_{jit}^{'}\beta$
Panel	$y_{jit} = \alpha_{ji} + x_{jit}^{'}\beta_j$	$y_{jit} = \alpha_{ji} + x_{jit}^{'}\beta$

The motivation behind such classification is that, first of all, the outcomes of the binary choice models are very sensitive to the variation in the dependant variable y. And as the bankruptcy of the bank is not a very frequent event, the number of zeroes (no default) significantly dominates the number of ones (default), we need to try different techniques to get the output. For instance, initially we only had about 2% of defaults in our sample and then the pooled probit for Ukraine did not give any results (unless the constant term was excluded from the regression) before we added to the initial sample the information for the year 2008 and the number of defaulted banks increased almost twice for Ukrainian case[14]. One more way to overcome the problem of small variation in the dependent variable is the use the random effects (RE) probit for the panel data. The further discussion on panel data is presented below

3.3.4 FIXED AND RANDOM EFFECTS FOR PANEL DATA

As it was shortly mentioned above, the panel data are two-dimensional (even though, it can have a more complicated structure), and contains the time series observations for a number of cross-sectional units. Hsiao (2005) distinguishes between different advantages of the panel data. Among others he emphasizes that panel data usually gives more accurate and effective predictions because of a greater number of degrees of freedom and less multicollinearity comparing with cross-sectional data. He also questions the accuracy of estimations if the data is pooled rather than taking into account the data for each individual. He reckons that if data are conditional on the same variables, panel data allow taking into account individual heterogeneity by observing the behaviors of other individuals. Hence, this gives more precise information on individual behavior in comparison with another one.

According to Verbeek (2008) in the estimation of linear or nonlinear models using panel data, a fixed or random effects specification has to be chosen. To explain the difference among them, we present here the binary choice model in terms of an underlying latent variable y_{it}* (this variable represents the unobserved characteristics, which influence the outcome of qualitative variable):

(3.8) $$y_{it}^* = x_{it}'\beta + \alpha_i + u_{it}$$

where y_{it} =1 if $y_{it}^*>0$ and $y_{it} = 0$ otherwise. The error term u_{it} has a symmetric distribution function F(.), var (u_{it}) = σ_u^2, Cov(u_{it}, u_{js}) = 0, $\forall\ i \neq j$ or $t \neq s$. The term α_i is the so called the individual heterogeneity. We can assume it to be a fixed parameter, which should be estimated or we consider it as a random error term, accordingly we have then the fixed effects or random effects specification of the model.

Greene (2004) emphasizes practical and methodological shortcomings of the non-linear fixed effects model. The first is the difficulty of computing the maximum likelihood estimation for probably thousands of dummies and the second is the "incidental parameter" problem, which brings into question the statistical properties of ML estimator. He argues that this estimator is inconsistent when T, the number of time period is fixed. Verbeek (2008, p.396) underlines that in

[14] Due to the financial crisis for the first four months of year 2009 13 Ukrainian banks were placed under the temporary administration of National Bank because of their insolvency and was indicated as default in year 2008 in our data set.

the fixed effects binary choice model, "only individuals who changed the status at least once are relevant for estimating β". Others are simply excluded from the estimation as they do not provide any information about β. Since the banks in our case have very few defaults, they do not change their status very often, and thus using random effects specification seems more reasonable. At the same time, we have to be aware that the random effects probit assumes the independence of individual effects (the α_i) of the regressors (the x_i). But the correlation between these two terms often exists in practice and this can lead to inconsistent estimators. Besides the choice between the random or fixed effects, we have to distinguish if we use logit or probit in our analysis. The common approach for the panel data is to use the multivariate normal distribution, which leads to a random effects probit model. The motivation behind this is that we assume the normal distribution for both α_i and $\boldsymbol{u}_{it}$. It is, of course possible to assume the logistic distribution for one or both of the terms, but then the sum of this distributions could lead to estimation of probabilities, which does not correspond to probit or logit model.

Consequently, in our paper we use the random effects probit model for the analysis of the panel data. We also implement the standard probit estimation for pooled data and, despite its shortcomings the linear probability model is used. According to Maddala (1987), if we ignore the correlation among the error terms, the linear regression can give consistent (but not efficient) estimates.

3.3.5. Specification tests in the models

For specification tests in the models, we found it important to implement the heteroskedasticity and the normality tests. As it was mentioned in Section 3.3.3, we can suspect the presence of heteroskedasticity when we are pooling the cross-sectional data over time. Moreover the presence of nonnormality and heteroskedasticity can make the estimation of binary choice models biased. Below we describe consequently the effect of heteroskedasticity and nonnormality on linear and nonlinear models as well as testing strategies.

In the linear regression model one of the assumptions, which has to be made to produce the efficient estimates of the coefficients, is the assumption of homoskedasticity or constant variance of conditional distribution of the error terms ε given the matrix of explanatory variables x:

$$Var(\varepsilon|\boldsymbol{x}) = Var(\varepsilon) = \sigma^2\boldsymbol{I} \qquad (3.9)$$

where $\boldsymbol{I}$ is identity matrix (diagonal matrix with diagonal elements equal to one).

But if the variance of the error term varies over observations and does not equal to expression 3.9, then this problem is referred to heteroskedasticity. As a consequence the OLS estimates are not efficient, but still unbiased and consistent. [15] To overcome this problem in the linear regression model, modern econometric theory proposes either using an alternative estimator, reconsidering the model specification or adjusting the standard errors of the OLS estimator for heteroskedasticity.

As to normality, for exact statistical inference the assumption about the distribution of the error terms has to be made. Most commonly, the error terms are assumed to be jointly normally distributed. If this restriction is violated, interpretation of the results could not be correct.

In the case of binary choice models the heteroskedasticity and nonnormality have other consequences. According to Wooldridge (2002, p.479), if we remember that in latent variable model we focus on $P(y = 1|\mathbf{x})$, then heteroskedasticity in Var $(\varepsilon|\mathbf{x})$ "entirely change the functional form for $P(y = 1|\mathbf{x}) = E(y|\mathbf{x})$ and "it makes little sense to care about consistent estimation for $\boldsymbol{\beta}$ when $P(y = 1|\mathbf{x}) \neq F(\mathbf{x}_i'\boldsymbol{\beta})$. Nonnormality in the error term in probit model means that F(w) ≠ Φ(w), where Φ(w) is the standard normal density and $P(y = 1|\mathbf{x}) \neq \Phi(\mathbf{x}_i'\boldsymbol{\beta})$. Hence, both of these problems mean that the distribution assumed for binary choice model is not correct and it leads to inconsistent estimator.

Taking into consideration the different consequences for linear and binary choice models and probit in particular, the testing strategies are also quite different. We summarized the tests, which are used in this essay in Table 3.4. However, one more important thing we need to mention is that the discussion above covered only for the case with pooled data. In a random effects panel data model the things are becoming more complicated. According to Verbeek (2008, p.373), "most of the tests that can be used for heteroskedasticity or autocorrelation in the random effects model are computationally burdensome". Moreover, as it is presented in Appendix 5, the tests for both linear and nonlinear models in the pooled set of data (except the probit for Ukraine) reject the null hypothesis of homoskedasticity. Normality is rejected in all cases except the probit for the set of Russian banks and weakly for the Ukrainian pooled probit. That is why we assume that in the random effects probit model we have heteroskedasticity (except the case for Ukraine) and nonnormality (except the case for Russia) as well.

[15] For the proof see, for example, Section 4.1 in Verbeek (2008).

Table 3.4 Tests for Normality and Heteroskedasticity used in the Study [16]

The Test	The Null Hypothesis	The Testing Procedure	Test Statistic
		Linear Model	
Normality Test	Skewness: $E(\varepsilon_i^3)=0$, Kurtosis: $E(\varepsilon_i^4-\sigma^4)=0$ under the normal distribution	Check: If $E(\varepsilon_i^3) \neq 0$, the distribution of ε_i is not symmetric around zero If $E(\varepsilon_i^4-3\sigma^4) > 0$, the distribution displays excess kurtosis	
Breusch-Pagan Heteroskedasticity Test [17]	H_0: $\sigma_i^2=\sigma^2$ H_1: $\sigma_i^2=h(\boldsymbol{z}_i'\alpha)$	1) Assume simple linear regression: $y_i = \alpha + \beta x'_i + \varepsilon_i$ 2) Regress y on explanatory variables x and get the residuals ε; 3) Regress ε^2 on a constant and $\mathbf{z}$; 4) Test if coefficients of $\mathbf{z}$ are equal to 0.	1) Usual F-test 2) $n{*}R^2$ from regression (3) is asymptotically χ^2 under H_0 with df equal to the number of variables in $\boldsymbol{z}$
		Binary choice model	
Normality Test (Lagrange Multiplier (LM) framework)[18]	H_0: $\gamma_1=\gamma_2=0$	1) Assume the restricted distribution function, which allows for skewness ($\gamma_1 \neq 0$) and excess kurtosis ($\gamma_2 \neq 0$), then probit is described as following: $P(y=1\|\boldsymbol{x}) = \Phi(x_i'\boldsymbol{\beta} + \gamma_1(x_i'\boldsymbol{\beta})^2 + \gamma_2((x_i'\boldsymbol{\beta})^3)$ *2)*Run the auxiliary regression of vector of ones upon[19] $e_{gi}\, x_i'$, $e_{gi}\,(x_i'\boldsymbol{\beta})^2$, $e_{gi}\,(x_i'\boldsymbol{\beta})^3$	1) Usual F-test 2) $n{*}R^2$ from regression (2) is asymptotically χ^2 under H_0 with two degrees of freedom (df)
Heteroskedasticity (Lagrange Multiplier (LM) framework)	H_0: $\sigma_i^2=\sigma^2$ H_1: $\sigma_i^2=h(\boldsymbol{z}_i'\alpha))$	1)Run the probit on explanatory variables, retrieve the fitted probabilities $\hat{p}$ and fitted index function $x_i'\hat{\beta}$ 2) Run the following auxiliary regression with standardized residuals[20] ε_{si} on the left-hand-side $e_{si} = \frac{f(-x_i'\hat{\beta})}{\sqrt{\hat{p}_i(1-\hat{p}_i)}} x_i' b_1 + \frac{f(-x_i'\hat{\beta})(-x_i'\hat{\beta})}{\sqrt{\hat{p}_i(1-\hat{p}_i)}} z_i' b_1 + u_i$	The explained sum of squares from regression in (2) is asymptotically χ^2 under H_0 with df equal to the number of variables in z

16 In this table the methodology described in Verbeek (2008), Wooldridge (2002) and Eviews manual is used.
17 In this table n denotes the number of observations; z is a vector of exogenous variables.
18 LM tests are used when unrestricted model is difficult to estimate. Thus we estimate only restricted model and see if the restrictions, which were made, violate the hypothesis about unrestricted model.
19 e_{gi} stated for generalized residuals here, which described in Section 3.3.2
20 The formula for standardized residuals is presented in Section 3.3.2

4. Empirical Study

4.1 Data Collection

In our study we use the financial ratios, which are drawn from the publicly available financial statements of Russian and Ukrainian banks. The data are taken on the annual basis from 01.01.2003 to 01.01.2008 for Russia and from 01.01.2003 to 01.01.2009 for Ukraine. This period is chosen because the best data set is available for it. Also it is important to take the data for 2008 for Ukraine because it allows us to increase the variation in the dependent variable. As it was mentioned before, for the first four months of the year 2009, 13 Ukrainian banks were placed under the temporary administration of National Bank due to their insolvency and they are indicated as default in year 2008 in our data set. The annual balances for Russian banks for the year 2008 are still not available for a moment.

Our sample includes the accounting data on 943 Russian credit institutions, which are available on the web-page of the central bank of Russia and 186 Ukrainian banks, which were in business in the period 2002-2008. The total number of observations comes up to seven thousand for two countries. The banks are classified as failed if their licenses are withdrawn or the banks are under the temporary administration of National Bank of Ukraine or Agency for Restructuring the Credit Organizations in Russia.

As the sources of our data we used the web-pages of Ukrainian and Russian central banks. And we need to mention here, that these web-pages are not designed to get the necessary information easily. That is why the data collecting took the major of time, which was given for writing this paper. The information, which posted on Ukrainian web-page, is presented in aggregated tables for all the banks. On the one hand, it is easy-to-use, but on the other hand it took us very long time to discover if a specific bank was failed, merged or just renamed if it disappears from the table in the forthcoming period. Such information is not available on the web-page of National Bank of Ukraine, so we had to search for it in different sources on the web. Moreover, the information in aggregated tables does not allow us to use all desirable variables, which are available in the balance sheets of Russian Banks, for instance. The web-page of Russian central bank is much better organized than Ukrainian and gives the separate balance sheet for each of the bank. But due to the fact that we were drawing 10 parameters from each of the balance sheet for more than 900 banks for 6 years, it was very time-consuming and finally the special computer programme was written by specialist to optimize the process of data collecting for Russian banks.

4.2 SELECTION OF VARIABLES

The dependant variable *y* is constructed in the following way: it is equal to zero if the bank is non-failed and it is equal to 1 if the bank is failed. To each value 1 of the bank, failed at time t corresponds its financial ratios at time ***t-1***.

Traditionally, in the financial ratios analysis the ratios of profitability, liquidity, and solvency are used. Unfortunately, in our study we have to skip the liquidity ratios because of the lack of appropriate indicators for Ukrainian banks and necessity to keep the equality in model for two countries estimated together. In our choice of explanatory variables we follow the previous research at this field. Montgomery, Santoso et al.(2005) take 17 ratios for their analysis, Peresetsky et al. (2004) use about 30 parameters in different models. But as we are limited by our initial data and also we are aware of the problem of multicollinearity (the sum of all possible parameters can come to one). "In general the term multicollinearity is used to describe the problem when an approximate linear relationship among the explanatory variables leads to unreliable regression estimates." (Verbeek, p.43). Tucker (1996) emphasizes that this problem is typical when the financial data are modeled because all the variables are drawn from the same balance sheet or income statement. To minimize the presence of this problem we include in our study only those coefficients, which we find interesting for the analysis. Initially we took ten financial ratios, but then excluded two of them because of high correlation with other variables. Furthermore, in addition to the financial ratios we planned to include the macro-indicators in our study. But we were confronted by a problem of perfect correlation between the variables, because the macro-indicators are the same for each bank for a particular year.

In table 4.1 the list of variables included in the model as well as their expected signs are presented.

Table 4.1 Variables Included in Estimation[21]

No	Variable	Description	Expected sign
1	CTD	Capital/Deposits	-
2	CTA	Capital/TotalAssets	-
3	LTA	Loans/Total Assets	+
4	LTD	Loans/Deposits	+
5	ROA	Return/Total assets	-
6	ROC	Return /Capital	-
7	DTA	Deposits/Total assets	-
8	Size	Natural logarithm of total assets	-

We expect capital to deposits and capital to total assets ratios to have the negative influence on probability of default because better capitalized banks have more reserves to cover their loan losses and meet obligations in case of bankruptcy. We assume the ratios loans to total assets and loans to deposits have the positive sign. The first ratio measures the credit risk and the second shows how much the bank depends on borrowed funds. We expect the ratios which measure the profitability (return on assets and return on capital) to have the negative sign because the high profitability usually indicates good performance and management of the bank. As Konstandina (2006, p.11) emphasizes “some aggressive and risky banks also could be profitable a few period before the failure”. But as it was mentioned before, the policy of tax minimization (and profit-minimization correspondingly) is typical for ex-Soviet Union companies, that is why the expression above is not valid in our case. The deposits to assets ratio indicates the level of investors trust in the bank. The proxy for the bank size is taken as the natural logarithm of the total assets. It is not clear in which direction it influences the probability of default, but similar to other studies we assume a negative sign.

The relative means of the chosen coefficients, grouped by failed and reliable banks, are presented in Figures 4.1 and 4.2. A visual analysis of the figures does not allow indicating clearly the direction of the influence, which can be expected in each of the countries, moreover contradictory conclusions can be made if we compare the means for each of the groups between the countries. For instance, the mean value for the loans to deposits ratio for failed banks in Russia are higher than for surviving banks, but the situation in Ukraine is inverse. The same picture is indicated for return on capital variable. Moreover, the fact that in Ukraine the return on capital is higher for failed banks than for the surviving is quite surprising, because it measures

[21] The full list of variables used in econometric models, including dummies, is presented in Appendix 1

the profitability of the bank, and it should be lower if the bank is in poor condition. Though, this could be connected with the fact, that sometimes when the bank has big losses, its own capital can be reflected in the balance sheet as a negative value. Consequently, the ratio of two negative figures, gives the positive ratio and this could explain the illustrated difference in the means. Hence, in contrast to theoretical suggestions, we can expect this ratio to have the positive sign for the probability of default in Ukrainian banks, though return on assets could have a negative sign. According to Figure 4.2 we can guess that capital to deposits and loans to deposits ratios have a positive influence on probability of default in the Russian case.

Figure 4.1 Relative Means for Failed and Reliable Banks (Ukraine)

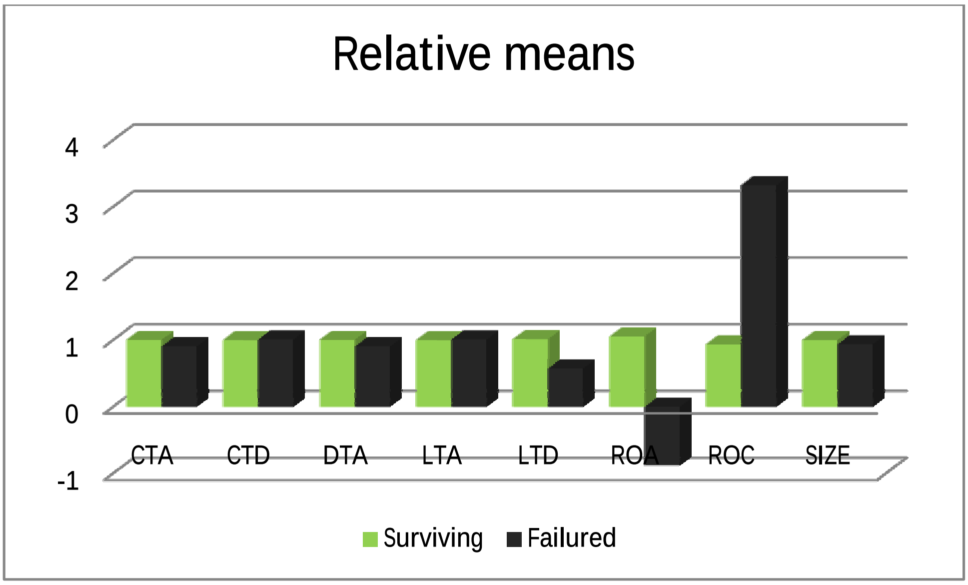
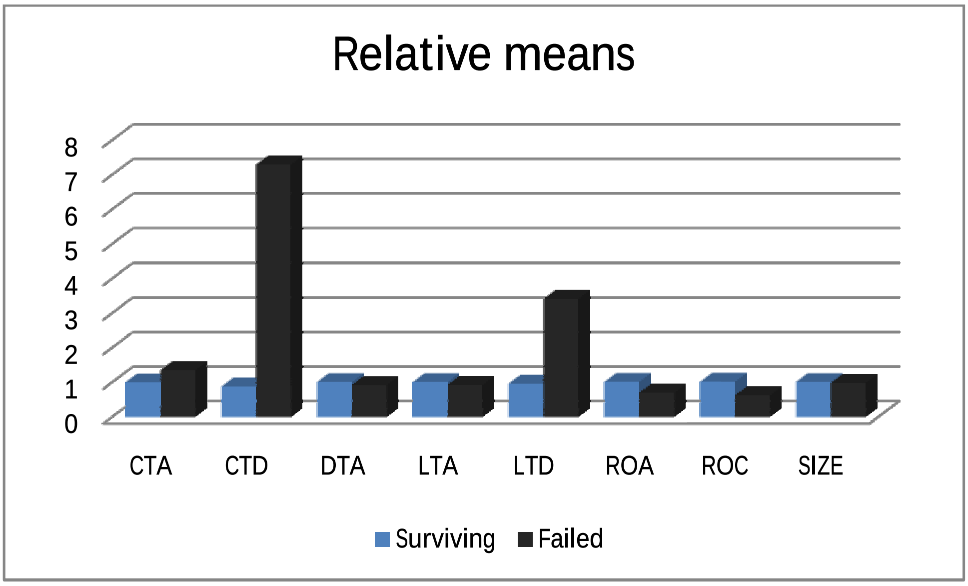


Figure 4.2 Relative Means for Failed and Reliable Banks (Russia)



4.3 *DISCUSSION OF ESTIMATION RESULTS*

In this sub-section we discuss the estimation results and highlight the coefficients which are found to be statistically significant for the banks failures for a given sample of Russian, Ukrainian banks and the cross-country model.

In the sample of Ukrainian banks (Appendix 2) for the most of the coefficients we get the meaningful signs in the pool, panel probit and with some exclusions in the linear regression as well. For instance, the capitalization (CTA) and profitability measures (ROA) influence a bank failure negatively, as well as the size of the bank, i. e. the more profitable and bigger the bank is - the smaller the probability of its insolvency. The sign of the measure of a bank's dependency on borrowing funds (LTA) also corresponds to the theoretical expectation, though it is insignificant in all implemented models. Furthermore if we return to our discussion on relative means of variables presented in Figure 4.1, where we expect the return on capital ratio (ROC) to have a positive sign (as its mean value is greater for the group of failed banks in comparison with surviving institutions), we can see that actually in linear regression and panel probit it has a positive sign and it is not statistically significant in any of the three presented models. Thus, the quality of initial data is crucial for the output of the estimation. Moreover, we have to be careful with inferences of the results as they could be biased because of heterokedasticity and non-normality in pooled models, and in addition for the panel probit the estimation results can be biased because of an unbalanced panel.[22] In Appendix 5 we present the results of specification tests, which show that we do not have the evidence of homoskedasticity except for the pooled probit for the Ukrainian case. This is probably connected with the fact that the number of cross-sectional units here is smallest in comparison with other samples. Moreover, we cannot reject the null hypothesis of normality at 1% significance level also. Consequently, based on pooled probit, which seems to be correctly specified, we can conclude that for a given period in Ukraine such explanatory variables as the capital to total assets, return on assets ratios as well as the size are found to be statistically significant. And the sign of their coefficients is in line with theoretical.[23] In the linear regression the error terms are heteroskedastic and not normally distributed for the models in all three samples (Appendixes 5, 6). Though, they are adjusted for heteroskedasticity with Newey-West standard errors.

[22] Several institutions in our sample were merged or appeared in business not in the beginning of the concerned period. For further discussion on unbalanced panel see Amemiya et al. (2001)

[23] We should mention that before we have included in Ukrainian sample the data for the year 2008, which made the variation in the dependent variable greater, all the coefficients in the pooled models were insignificant and only two significant coefficients were in the panel probit model (SIZE and ROA). By this fact we want to emphasize again how sensitive the outcome of binary choice models to the variation in the dependent variable.

For the sample of Russian banks (Appendix 3), we single out three coefficients, which are significant in the LPM, pooled and panel probit. These are the coefficients for capital to total assets ratio (CTA), return on assets (ROA) and loans to total assets (LTA). Though, only one of them, ROA, has the expected sign. Opposite to theoretical, the coefficient of capital to total assets ratio (CTA) has a positive sign, but it is in line with the results of Konstantidina (2006) for the sample of Russian banks for the period from 1999 to 2003. She supposes that this could mean that less capitalized banks may recognize their weaknesses and do not undertake risky operations, in contrast with banks with big capital, which are more confident and thus can actively work with subprime mortgages, for instance. The negative sign of loans to total assets (LTA) ratio can be explained probably by the fact that the banks with big loan portfolio are more profitable (as the loan is the one of the important income items for banks), which indicates the good performance of its business. However, in the case with the Russian data set the specification tests (Appendix 5) do not allow us to say that our nonlinear model is correctly specified. And even though the normality is not rejected for the pooled probit, error terms are still heteroskedastic and the estimates of the model are not likely to be close to the true parameters.

The common result for the two countries according to the local models: is the influence of the profitability (ROA) on the bank failure prediction (it is significant and have the same direction of influence for the both countries) and the capital to total asset ratio (CTA) (it significant in both countries, but have the different direction of influence).

In the cross-country model as the dependent variables we have the pooled variables of Ukrainian and Russian banks. To introduce the country effect we use the country dummy (CD)[24], which has the value of 1 for Russia and 0 for Ukraine. To see if there are any similarities in forces driving the banks to default in two countries, we include in the model variables specified according to each country. In other words, we include explanatory variables multiplied by the country dummy (CD_CTD, CD_CTA etc). Unfortunately we can not specify a country-time dummy because the estimation failed due to the perfect collinearity between the variables, so they were removed from the equations[25]. According to the estimates presented in Appendix 4, those coefficients of

[24] Though, in the case of pooled model it is probably more precise to say that this variable indicates if the model is correct or not.

[25] The solution which could help to overcome this problem (excluding from the estimation only the time dummy multiplied by county dummy for the year 2008, for which we do not have any data for Russia) was explored in the end of writing of this paper. The changes couldn't be made due to the time limitations and has to be taken into account for the future research.

financial ratios, which are created by pooling the parameters of two countries in one series, are all insignificant (except the ROC, which is significant on 10% level in OLS regression). The country dummy is not significant in three statistical models, but we cannot argue that the country effect does not exist in cross-country model. Because some other coefficients of the variables, which are designed to capture the differences between the countries, in particular CD_CTA, CD_ROA, are significant in pooled, panel probit and linear regression. Therefore we perform the Wald test with the null hypothesis that all the country specific coefficients (CD, CD_CTA, CD_ROA, etc.) are equal to zero. The test results, presented in Appendix 7 (Table A7.1), allows us to reject the null hypothesis only on the 5 % significance level (for pooled and panel regression), but not on the 1%. This could mean that there are some similar factors, which drive the banks to bankruptcy; however, the significant difference between two countries exists. Moreover, the effect of the influence of the same financial ratios on the bank failure could be opposite for each of the country (as it was described for the two local models above). But nevertheless we want to aware the reader again, that this conclusion also has to be taken very carefully because the tests in Appendix 5 and 6 indicate that the model is misspecified for cross-country data.

However, there is only one effect about which impact we can be sure. It is the time effect. Almost all the time dummies are significant in domestic and cross-country models. In addition, we perform the Wald test with the null hypothesis that all the time dummy coefficients are equal to zero. The results of the test are presented in Appendix 7 (Table A7.2). According to the test results we can reject the null on 5 % significance level for all the statistical models and on 1% significance level for all models, except the OLS. So, we can conclude that in almost all the cases yearly events have significant impact on the banks insolvency. Significance of the time dummies indicates that there are strong differences in failures in different periods. For example, in 2004-2005 in Ukraine there were very specific economic and political conditions which could have impact on banks` failure probability. According to the estimation results D2004 is significant in all types of them only for Ukraine.

And even though, our results do not show that many of the coefficients are statistically significant, we need to compare the models in their predictive power and the results are presented in the next subsection.

4.4 MODEL COMPARISON

To compare the predictive power of domestic models with the cross-country in this subsection we present the results of the in-sample and out-of- sample forecasts.

4.4.3 IN-SAMPLE FORECAST

As all three statistical models (LPM, pooled and panel OLS) show more or less similar results in terms of significance of the coefficients, we have chosen only pooled probit for testing the predictive power (in-sample as well as out-of-sample) [26]. In Table 4.2 we present the evaluation of percentage of correct predictions of zeroes and ones for two domestic and cross-country models.

Table 4.2 Expectation-Prediction Evaluation for Binary Specification

	Estimated equation								
	y=0	y=1	total	y=0	y=1	total	y=0	y=1	total
	Russia			Ukraine			Cross-country		
P(y=1)≤ C	3532	46	3578	930	9	939	4460	47	4507
P(y=1)>C	2367	79	2446	171	21	192	2385	92	2477
Total	5899	125	6024	1101	30	1131	6845	139	6984
Correct	3532	79	3611	930	21	951	4460	92	4552
% Correct	59.87	63.20	59.94	84.47	70.00	84.08	65.16	66.19	65.18
% Incorrect	40.13	36.80	40.06	15.53	30.00	15.92	34.84	33.81	34.82

But to proceed with this type of evaluation, we need to choose the cut-off point. If the value of the estimated probability is less or equal to this point, the bank is classified as default, and non-default in the opposite case. As it is suggested by Eviews manual, it is useful to take the value of this threshold equal to the percentage of ones in *y*. In our case we take the success cut-off point for Ukraine equal to 0.04. For Russian and the cross-country models the threshold value is equal to 0.02. Consequently, based on the specified cut-off point, we can see that the pooled probit for the Russian sample correctly predicts 59.87% of the surviving and 63.2 % of failed banks. For the Ukrainian sample this model correctly predicts 84.47 % of reliable and 70 % failed banks. In the cross-country model the percentage is 65.16 % and 66.19% correspondingly. Thus, overall in-sample forecast indicates that the pooled probit correctly predicts 59.94% of the observations

[26] We also perform the forecast for the panel probit models but as far as we get worst results for the 2 from 3 models we decided not to include these calculations in discussion, but they can be given by request. The main results of in-sample forecast are the following: overall panel probit model accuracy for Ukraine is 97.52%, for Russia 57.93% and for the cross-country model is 61.99%

for Russia, 84.08% for Ukraine and 65.18% for the cross-country model. Hence, we do not find the evidence here that our cross-country model outperforms the models for Ukraine and Russia.

4.4.4 Out-of-sample forecast

One more way, which is commonly used in evaluating the performance of statistical models, is an out-of-sample forecast. Sometimes the predictive power of the estimated model is evaluated taking the data for future periods of time and if those are not available, the data for the past periods are used. In our case the data on the bank's parameters nor for future, neither for the past periods for the both of the countries is available. That is why we apply procedure similar to the predictive Chow-test. Following the previous researchers, for instance, Peresetsky et al. (2004), we randomly select observations from our main sample of reliable and failed banks (we kept the percentage of failures equal to initial sample). Then these observations are excluded from the main sample. As a next step, the pooled probit is estimated for the bigger sample and evaluated for the smaller one. Hence, for Russia we get two samples: the first contains 5823 observations among which 115 failures and the second contains 200 observations with 10 failures. For Ukraine we get the estimation sample of 931 observations (22 failures) and the prediction sample consists of 200 observations with 9 bankruptcies and for the cross-country model the estimation and prediction samples consist of 6785 (137) and 400 (19) observations, respectively. The prediction results are shown in the Table 4.3. For the two domestic and the cross-country models overall accuracy in predicting the default of the banks is almost equal (close to 95%), but only for the correct prediction of survival banks. The number of correct predictions of the defaults for both domestic and cross-country models is zero. But these results are not surprising, taking into account the fact that only few coefficients are found to be statistically significant and the model is not correctly specified for Russian and cross-country sample.

Table 4.3 Out-of-sample forecast implemented with pooled probit

	y=0	y=1	total	y=0	y=1	total	y=0	y=1	total
	Russia			Ukraine			Cross-country		
P(y=1)≤C	190	10	200	191	9	200	381	19	400
P(y=1)>C	0	0	0	0	0	0	0	0	0
Total	190	10	200	191	9	200	381	19	400
Correct	190	0	190	191	0	191	381	0	381
% Correct	100.00	0	95.00	100.00	0	95.50	100.00	0	95.25
% Incorrect	0	100.00	5.00	0	100	4.50	0	100.00	4.75

In addition, similar to Nikolsko-Rzhevskyy (2003) we would like to present the ranking list of the banks according to the estimated probability of bankruptcy (using pooled probit). We create two groups of 12 banks[27] with a small and big probability to fail for Ukrainian, Russian and the cross-country models. We name them the "white list" and the "black list", respectively (Appendix 8). But according to the information published on the web-pages of central banks of Russia and Ukraine, not any of the banks from our "black list" has financial difficulties at the moment. This gives us one more evidence that the variables, which we include in the model and the form of the model does not predict the bank failure for a given sample and a given period of time correctly.

[27] We choose only "alive" banks to see what probability of their failure gives our model.

5. Conclusions

In this Master essay we use the linear probability and the binary choice models for the analysis of financial ratios, which are drawn from the publicly available financial statements of Russian and Ukrainian banks for the period 2002-2008. The two purposes are specified for the study: 1) to investigate which parameters are significant in the prediction of bank failure for the sample of Russian and Ukrainian banks for the first decade of this century; 2) to see if there are any similar factors, which drive the bank to bankruptcy in both countries.

First of all we need to emphasize that the output of econometric probability models is very sensitive to the quality of initial data and the variation in the dependent variable. For instance, the number of defaults in Ukrainian sample is the double of defaults in Russian sample and moreover according to the specification tests the model is correctly specified. Consequently, we have got the results that Ukrainian pooled probit shows the greatest number of correctly predicted defaults in comparison with Russian and cross-country models for in-sample forecast. Even though it does not indicate the good predictive power out-of-sample, the estimates of significant coefficients have the meaningful sign. For instance, such variables are found to be significant for the bank failure: the capitalization and profitability measures and the size of the bank. All of them influence the default of the bank negatively, thus, the more profitable and bigger the bank is - the smaller the probability of its insolvency. Nikolsko-Rzhevskyy (2003) in his analysis of Ukrainian banks for period 1998-2003 also finds that bank`s size does matter for the failures of the banks for a given period. However, in contrast to him we do not find the significant influence of the deposits to total assets ratio on the bank`s ability to survive. The possible explanation for this is that in comparison with his period of study, today people's confidence in the banks in Ukraine has increased and it is less important (at least based on the evidence from our data) than profitability and capitalization measures.

For the sample of the Russian banks the size of the bank is not found to be significant for the probability of default. Only the coefficient of return on assets ratio (ROA) is significant and has the expected sign. The capital on assets ratio (CTA) is also significant, but it has the opposite direction of influence in comparison with Ukrainian one and the theoretical suggestions. Nevertheless, it is in line with the previous study of Russian banks failures by Konstandina (2006). She explains this by the possible risk aversion of the less capitalized banks.

As to the second objective of our study, we only have one significant variable with the same effect on bank failure, which is common for the both countries. It is the return on assets (ROA). And this is not surprisingly as it is often found to be an important indicator of the bank performance. For instance, Sheldon (1996) at his study of the bank's default probability in international comparison based on accounting data, used return on asset as a variable that has a great impact on banks probability to fail. The capital to total assets ratio (CTA) is also significant in both countries, but if in the Ukrainian case it influences the failure of the bank negatively, then for the sample of Russian banks it has the opposite sign. The possible explanation for this fact can be given as following: to our knowledge (according for example the Intelace research[28]) Ukrainian banks are undercapitalized in comparison with other Central and East European banks, but some of the Russian banks are overcapitalized. Probably the less capitalized banks (which dominate in Ukraine) may recognize their weaknesses and do not undertake risky operations, in contrast with banks with big capital (dominating in Russia), which are more confident and thus can actively work with subprime mortgages, for instance. Hence, the capitalization of the banks in Russia influences the probability of their defaults positively in contrast to Ukrainian sample. One more difference between the countries is that the size of the bank negatively influences the default of the bank for Ukrainian sample and does not have any influence for the sample of Russian banks. We can guess that the regional difference inside Russian Federation is much more significant than the difference in economic conditions in different regions of Ukraine, hence some other regional factors can have greater influence on the default of the bank in Russia than its size. Also, such variable as loan to total assets (LTA) is not found to be significant in Ukraine, but it is significant in Russia. Intuitively, this variable is expected to have the influence on the default of the bank because if the bank has the great number of issued loans and they are not paid back, this makes the bank insolvent. However, the insignificance of this variable for Ukrainian case possibly can be caused by the smaller sample of Ukrainian banks in comparison with Russian.

Finally, we conclude that despite some similarities, there are the differences in forces driving the banks to default in two countries. Such conclusion partly contradicts the results of Montgomery, Santoso et al. (2005), who find that their cross-country model out-performs the domestic models in predictive power for the case of Indonesia and Japan. This, in their words, can stimulate the regional cooperation on this issue. We cannot argue that our findings reject the necessity of cooperation between the former Soviet republics in the issue of prevention of systemic banking crisis. Inversely, the cooperation is important and it is even already exists. For instance, in

[28] See http://www.intelace.com.

January 2009 Vnesheconombank of Russia has bought 75% of shares of one of the insolvent Ukrainian banks. This fact can be taken as an example of cooperation in prevention of financial crisis in Ukraine, because the problem bank is reactivated now and starts the repayment of its debts. But nevertheless, too close cooperation always has the danger to become the expansion. That is why, we reckon, the regulatory policy of central banks on the local levels still remains the most important in healthy functioning of the banking sector.

Further research in this area could focus on estimating the binary choice models under the relaxed assumptions (for example semi-parametrically), calculating the marginal effect of the influence of each variable and using some additional variables, for example the liquidity ratios, macroeconomic variables and country dummies multiplied by year dummies.

References

Altman, E.I. (1968), "Financial Ratios, Discriminant Analysis and the Prediction of Corporate Bankruptcy," *Journal of Finance*, 23(4), 589-609.

Amemiya, T., Hsiao C., Morimune, K., Powell, V. (2001). *Nonlinear statistical modeling: proceedings of the thirteenth International Symposium in Economic Theory and Econometrics: essays in honor of Takeshi Amemiya,* Cambridge University Press.

Aziz, A., M., and Dar Humayon, A. (2006), "Predicting corporate bankruptcy: where we stand?" *Corporate Governance,* 6:1, pages 18-33.

Canbas, S., Cabuk, A., & Kilic, S. B. (2005), "Prediction of commercial bank failure via multivariate statistical analysis of financial structure: The Turkish case", *European Journal of Operational Research*, 166(2), 528-546.

Champion, M., Bart, K., Connaghan C. (2009-02-18), "Eastern Europe Shakes Banks", *The Wall Street Journal (Europe),* vol. XXVII No.13.

Crawford, D., Champion, M. (2009-02-19), "Austrian Banks on Firing Line", *The Wall Street Journal (Europe),* vol. XXVII No.14.

Crosbie, Peter, Bohn, Jeff (2005), "Modelling Default Risk", *Moody's KMV Company*, http://www.moodyskmv.com/.

Demirguc-Kunt, A. and Detragiache E. (1998), "The determinants of banking crises in developed and developing countries", *IMF Staff Papers*, 45(1), 81–109.

Gharghori, P., Chan, H., Faff, R. (2006), "Investigating the Performance of Alternative Default-Risk Models: Option-Based Versus Accounting-Based Approaches", *Australian Journal of Management, 31(2), 207-234.*

Gonzalez-Hermosillo, B. (1999) "Determinants of Ex-Ante Banking System Distress: A Macro-Micro Empirical Exploration of Some Recent Episodes", *Working Paper of International Monetary Fund* WP/99/33.

Greene, William (2004), "The Behavior of the Maximum Likelihood Estimator of Limited Dependent Variable Models in the Presence of Fixed Effects", *Econometrics Journal*, volume 7, pp. 98–119.

Halling, M., Hayden, E. (2006), "Bank Failure Prediction: A Two-Step Survival Time Approach", http://www.efmaefm.org/efma2006/papers/674238_full.pdf.

Hsiao, Cheng (2005), "Why Panel Data?", *IEPR Working Paper* 05.33, http://www.usc.edu/iepr.

Kim, H., and Gu, Z. (2006), "Predicting Restaurant Bankruptcy: A Logit Model in Comparison with a Discriminant Model", *Journal of Hospitality and Tourism Research*, 30, 474.

Komulainen, T., and Lukkarila, J. (2003), "What drives financial crises in emerging markets?" *Emerging Markets Review*, 4(3), 248-272.

Konstandina, N. (2004), "Probability of Bank Failure: The Russian Case", *EERC working paper series, No 06/01*.

Laviola, S., Marullo-Reedtz, P. and Trapanese, M. (1999), "Forecasting Bank Fragility: The Evidence from Italy", *Research in Financial Services: Private and Public Policy* (Vol. 11), JAI Press.

Lee, S., H., Urrutia, J., L. (1996), "Analysis and Prediction of Insolvency in the Property-Liability Insurance Industry: A Comparison of Logit and Hazard Models", *The Journal of Risk and Insurance* 63:1, pages 121-130.

Maddala, G.S. (1987), "Limited Dependent Variable Models Using Panel Data", *The Journal of Human Resources*, 22, 307-338.

Montgomery, H., Tran, Bich Hanh, Santoso, W., Dwityapoetra S. Besar (2005) "Coordinated Failure? A Cross-Country Bank Failure Prediction Model", *ADB Institute Discussion Paper*, No.32, http://www.adbi.org.

Nikolsko-Rzhevskyy, O. (2003), "Bank Bankruptcy in Ukraine: What are the Determinants and can Bank Failure be Forecasted?" National University of "Kyiv-Mohyla Academy", *EERC, MA Thesis*, http://www.kse.org.ua/ .

Peresetsky, A., Karminsky, A. and Golovan, S. (2004), "Probability of Default Models of Russian Banks", *BOFIT Discussion Paper* 21/2004, www.bofit.fi.

Popryga, O. (2001), "Which Factors Cause Failure of Ukrainian Banks?" National University of "Kyiv-Mohyla Academy", *EERC, MA Thesis*, http://www.kse.org.ua/ .

Quantitative Micro Software, *EViews 6 User's Guide*, 2007, Irvine CA.

Saunders, A., Allen L. (2002), *Credit Risk Measurement. New Approaches to Value at Risk and Other Paradigms*, Second Edition, John Wiley & Sons, Inc., New York, USA, pp. 20-22.

Sheldon G., (1996), "Capital Asset Ratios and Bank Default Probabilities: An International Comparison Based on Accounting Data" *Swiss Journal of Economics and Statistics*, Vol. 132 (4/2), 743-754.

Tucker, J. (1996), *Neural Networks Versus Logistic Regression In Financial Modelling: A Methodological Comparison*, http://citeseerx.ist.psu.edu.

Verbeek, Marno (2008), *A Guide to Modern Econometrics*, Third Edition, John Wiley & Sons Ltd, Chichester, England.

Whalen, Gary (1991), "A Proportional Hazards Model of Bank Failure: an Examination of its Usefulness as an Early Warning Tool", *Economic Review*, Federal Reserve Bank of Cleveland, issue Q I, pages 21-31.

Wooldridge (2002), *Econometric Analysis of Cross Section and Panel Data*, Cambridge, Mass.: MIT Press.

Internet recourses:

http://www.bank.gov.ua

http://www.cbr.ru

http://www.alacrastore.com/research/moodys

http://ukrstat.gov.ua

http://in.reuters.com/article/asiaCompanyAndMarkets/idINLH37296720090217

http://www.epravda.com.ua/publications/4a0914d48ee17/

http://www.intelace.com.

Appendixes

APPENDIX 1

Variables description

Name of Variable	Description
CTD	Capital/Deposits
CTA	Capital/Assets
LTA	Loans/Total assets
LTD	Loans/Deposits
ROA	Profit /Total assets
ROC	Profit /Capital
DTA	Deposits/Total assets
SIZE	Natural logarithm of total assets
CD	Country Dummy
D2002	Dummy for year 2002
D2003	Dummy for year 2003
D2004	Dummy for year 2004
D2005	Dummy for year 2005
D2006	Dummy for year 2006
D2007	Dummy for year 2007
CD_CTA	Country Dummy multipled by Capital/Assets
CD_CTD	Country Dummy multipled by Capital/Deposits
CD_DTA	Country Dummy multipled by Deposits/Total assets
CD_LTA	Country Dummy multipled by Loans/Total assets
CD_LTD	Country Dummy multipled by Loans/Deposits
CD_ROA	Country Dummy multipled by Profit /Total assets
CD_ROC	Country Dummy multipled by Profit /Capital
CD_SIZE	Country Dummy multipled by Natural logarithm of total assets

APPENDIX 2

Estimation results[29] for the sample of Ukrainian banks

Series	Linear regression for pooled data[30]	Probit for pooled data	Probit random effect for panel data
CTA	-0.220585 (-1.900410)	-1.634459** (-2.194597)	-1.776257* (-1.78)
CTD	9.00E-05 (0.569888)	-0.002542 (-0.222859)	-0.0031509 (-0.23)
D2002	-0.144747*** (-2.999643)	-1.371281** (-2.412303)	-1.761087 (-1.64)
D2003	-0.103904** (-2.516241)	-0.923105* (-1.730951)	-1.209508 (-1.41)
D2004	-0.123764*** (-2.875946)	-1.838931** (-2.423553)	-2.322872* (-1.75)
D2005	-0.108231*** (-2.749991)	-1.235134** (-2.161011)	-1.552961* (-1.68)
D2006	-0.108788** (-2.556144)	-1.092300** (-2.017402)	-1.352361* (-1.67)
D2007	-0.046596 (-0.983384)	-0.311648 (-0.608720)	-0.3873689 (-0.62)
DTA	-0.080225 (-1.256677)	-0.799436 (-1.126320)	-0.9082015 (-1.00)
LTA	0.036156 (0.865585)	0.392348 (0.624137)	0.4569461 (0.58)
LTD	-0.000103 (-0.365933)	-0.036360 (-0.502195)	-0.0499911 (-0.52)
ROA	-0.245230*** (-2.715529)	-4.962900*** (-3.149303)	-5.676973** (2.57)
ROC	0.028192 (0.794574)	-0.009238 (-0.072405)	0.0071799 (0.05)
SIZE	-0.021760** (-2.132050)	-0.165049*** (-2.780481)	-0.1892631** (-2.08)
c	0.497211** (2.276339)	1.801760** (2.089966)	2.071793 (1.64)

[29] Corresponding t-statistics is in parentheses.
[30] The standard errors in the linear regression are adjusted for heteroskedasticity in all the estimation results below (Newey-West standard errors).
*** significant at 1% level.
** significant at 5% level.
* significant at 10% level.

APPENDIX 3

Estimation results[31] for the sample of Russian Banks

Series	Linear regression for pooled data	Probit for pooled data	Probit random effect for panel data
CTA	0.068088***	1.016878***	1.016833***
	(3.164607)	(3.175926)	(3.18)
CTD	7.93E-05	0.000753**	0.0007533**
	(1.398278)	(1.972890)	(1.97)
D2002	-0.018386**	-0.343360**	-0.3433236**
	(-2.525580)	(-2.506138)	(-2.51)
D2003	-0.014454**	-0.258744**	-0.2587149**
	(-1.990725)	(-1.975804)	(-1.98)
D2004	-0.002886	-0.067142	-0.0671256
	(-0.379594)	(-0.565457)	(-0.57)
D2005	-0.016090**	-0.324034**	-0.3240139**
	(-2.318448)	(-2.414727)	(-2.41)
D2006	-0.016462**	-0.333483**	-0.3334667**
	(-2.403341)	(-2.472794)	(-2.47)
DTA	0.008929	0.185528	0.1855243
	(0.891310)	(0.686921)	(0.69)
LTA	-0.027056**	-0.474011**	-0.4739938**
	(-2.323087)	(-2.570157)	(-2.57)
LTD	-1.92E-05	-0.000417	-0.0004168
	(-1.518648)	(-1.223902)	(-1.22)
ROA	-0.216184**	-4.374001**	-4.373835**
	(-1.989822)	(-2.161075)	(-2.16)
ROC	-0.001233	-0.265895	-0.2659144
	(-0.075861)	(-0.676793)	(-0.68)
SIZE	0.001383	0.023316	0.0233167
	(1.194290)	(0.910843)	(0.91)
c	0.011413	-2.170115***	-2.170067***
	(0.493405)	(-4.328789)	(-4.33)

[31] Corresponding t-statistics is in parentheses.
*** significant at 1% level.
** significant at 5% level.
* significant at 10% level.

APPENDIX 4

Estimation results[32] for cross-country model

Series	Linear regression for pooled data	Probit for pooled data	Probit random effect for panel data
CD	-0.003619	-0.839376	-0.8393585
	(-0.074643)	(-0.486361)	(-0.79)
CD_CTA	0.101970***	4.180962*	4.180823*
	(3.133996)	(1.919912)	(1.92)
CD_CTD	7.97E-05	0.005996	0.0059955
	(0.900311)	(0.051064)	(0.05)
CD_DTA	0.007942	2.290791	2.290703
	(0.305761)	(1.235877)	(1.24)
CD_LTA	-0.047942*	-2.679423	-2.679323
	(-1.903068)	(-1.423332)	(-1.42)
CD_LTD	-4.51E-05	0.918875	0.9188393
	(-0.375792)	(1.095804)	(1.10)
CD_ROA	-0.195309*	-5.336524**	-5.33638**
	(-1.797141)	(-2.207397)	(-2.21)
CD_ROC	0.009386	1.590929	1.590853
	(0.547363)	(0.855585)	(0.86)
CD_SIZE	0.001826	0.035631	0.0356307
	(0.480076)	(0.311779)	(0.31)
CTA	-0.033831	-3.167936	-3.167833
	(-1.369397)	(-1.470551)	(-1.47)
CTD	-3.87E-07	-0.005243	-0.0052428
	(-0.005713)	(-0.044653)	(-0.04)
D2002	-0.018076***	-0.355661***	-0.3556328***
	(-2.585208)	(-2.604057)	(-2.60)
D2003	-0.014335**	-0.269794**	-0.2697711**
	(-2.076838)	(-2.067212)	(-2.07)
D2004	-0.004055	-0.078039	-0.0780258
	(-0.564061)	(-0.659769)	(-0.66)
D2005	-0.015978**	-0.334030**	-0.3340138**
	(-2.410460)	(-2.496937)	(-2.50)
D2006	-0.016339**	-0.344543**	-0.3445295**
	(-2.503750)	(-2.562742)	(-2.56)
D2007	0.070359***	1.618528***	1.618489***
	(2.965270)	(2.745140)	(2.75)
DTA	0.001038	-2.111022	-2.110936
	(0.042340)	(-1.150352)	(-1.15)
LTA	0.020988	2.204247	2.204161
	(0.945148)	(1.176694)	(1.18)
LTD	2.59E-05	-0.919291	-0.9192555
	(0.216948)	(-1.096300)	(-1.10)
ROA	-0.021185	0.964779	0.9647707
	(-1.389676)	(0.731276)	(0.73)
ROC	-0.010750*	-1.856461	-1.856401
	(-1.908494)	(-1.021359)	(-1.02)
SIZE	-0.000444	-0.012968	-0.0129676
	(-0.121353)	(-0.116331)	(-0.12)
c	0.015065	-1.308337	-1.308315
	(0.328782)	(-0.785621)	(-0.79)

32 Corresponding t-statistics is in parentheses.
*** significant at 1% level.
** significant at 5% level.
* significant at 10% level.

APPENDIX 5

Specification tests

	Test statistic	Chi-squared critical values 1% significance	5% significance	10% significance	p-valu
		Probit for pooled data			
Heteroskeasticity tests (LM test)					
Cross-Country model (22 df)	107.8747	40.289	33.924	30.813	0.000
Ukraine (14df)	29.16973	29.141	23.685	21.064	0.009
Russia (13 df)	46.61191	27.688	22.362	19.812	0.000
Normality test (LM test, 2 df**)**					
Cross-Country model	10.93408				0.004
Ukraine	7.273554	9.210	5.991	4.605	0.026
Russia	3.276221				0.194
		Linear regression Model			
Breusch-Pagan Heteroskedasticity test					
Cross-Country model (22 df)	119.8838	40.289	33.924	30.813	0.000
Ukraine (14df)	58.16560	29.141	23.685	21.064	0.000
Russia (13 df)	64.54841	27.688	22.362	19.812	0.000

APPENDIX 6 Histograms of linear regression models residuals

Figure A6.1 Sample of Ukrainian banks

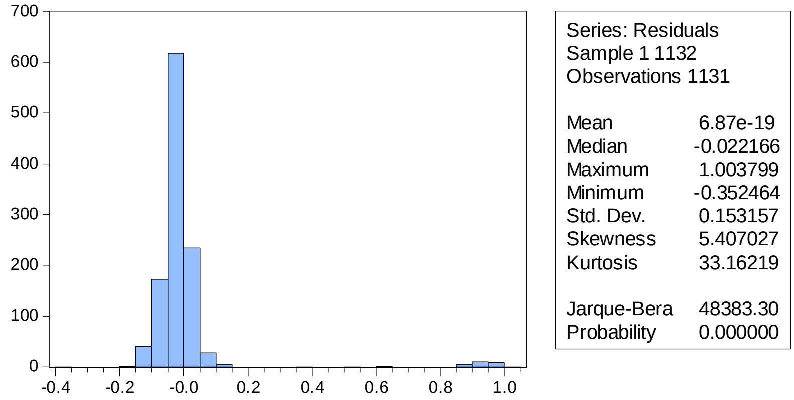


Figure A6.2 Sample of Russian banks

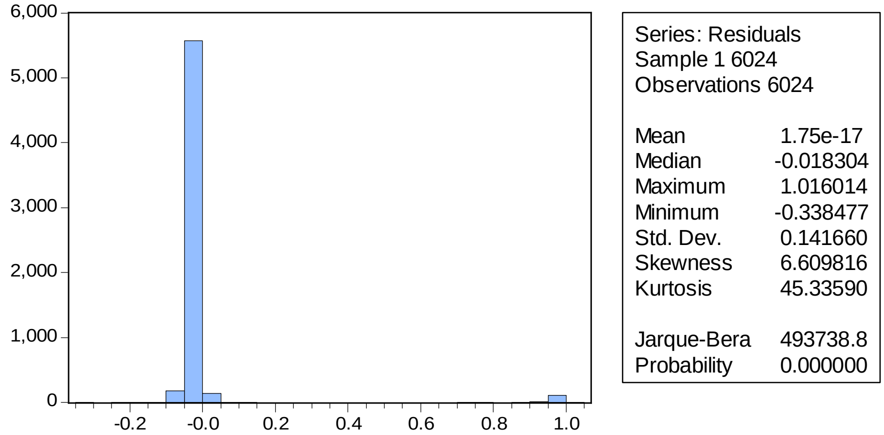
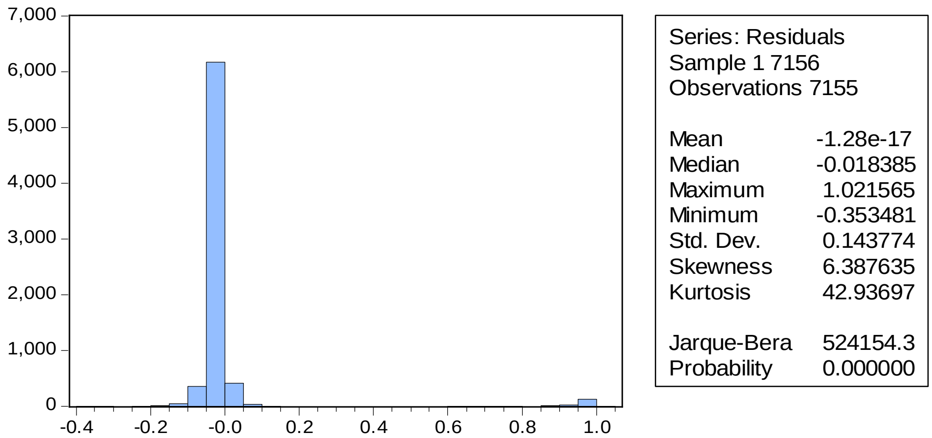


Figure A6.2 Cross-country sample



Appendix 7

Table A7.1 Results of Wald test with H_0: specific country coefficients (CD, CD_CTA, CD_CTD, CD_DTA, CD_LTA, CD_LTD, CD_ROA, CD_ROC, CD_SIZE) are equal to zero

	Chi-squared critical values (9 df)				
	Test statistic	1% significance	5% significance	10% significance	p-value
	Probit for pooled data				
Cross-Country model	17.78	21.666	16.919	14.684	0.037
	Probit random effect for panel data				
Cross-Country model	17.78	21.666	16.919	14.684	0.037
	Linear regression Model				
Cross-Country model	70.68	21.666	16.919	14.684	0.000

Table A7.2 Results of Wald test with H_0: all time dummies are equal to zero

	Chi-squared critical values				
	Test statistic	1% significance	5% significance	10% significance	p-value
	Probit for pooled data				
Cross-Country model (6 df)	24.04	16.812	12.592	10.645	0.000
Ukraine (6 df)	37.74	16.812	12.592	10.645	0.000
Russia (5 df)	13.45	15.086	11.070	9.236	0.019
	Probit random effect for panel data				
Cross-Country model (6 df)	24.04	16.812	12.592	10.645	0.000
Ukraine (6 df)	37.74	16.812	12.592	10.645	0.000
Russia (5 df)	13.45	15.086	11.070	9.236	0.019
	Linear regression Model				
Cross-Country model (6 df)	14.61	16.812	12.592	10.645	0.012
Ukraine (6 df)	47.01	16.812	12.592	10.645	0.000
Russia (5 df)	43.56	15.086	11.070	9.236	0.000

APPENDIX 8

Table A8.1 Lists of banks in Ukraine with the high and low probability to fail next year, predicted by the domestic probit model

"white list"		"black list"	
Name	Probability to fail	Name	Probability to fail
OTP bank	0.022402	Olympic Ukraine	0.132893
Industrialbank	0.032743	Accent-bank	0.132408
Khreschatuk	0.042011	Veles	0.124235
Kreditprombank	0.043139	Invest-Credit bank	0.116038
VAB bank	0.052127	Contract	0.115473
Active bank	0.052546	Agrarian Commercial bank	0.099257
Concord	0.063247	Coopinvestbank	0.097349
Partner bank	0.073174	TMM-bank	0.095471
Universal bank	0.078794	Trust-Capital	0.095311
Kredobank	0.083047	Home Credit bank	0.095057
Industrial-Export bank	0.084025	Pivdencombank	0.093824
Ikar-bank	0.084534	Grant	0.092039
Average probability	0.059316	Average probability	0.107447

Table A8.2 Lists of banks in Russia with the high and low probability to fail next year, predicted by the domestic probit model

"white list"		"black list"	
Name	Probability to fail	Name	Probability to fail
Morskoy bank	0.020991	Primorskiy territorialniy	0.027653
Uralprivatbank	0.021603	Smmit bank	0.027718
Dalcombank	0.022129	Aleksandrovskiy	0.027788
Unicreditbank	0.022392	Tempbank	0.028296
Energomashbank	0.022776	Baltiyskiy	0.030183
Russkobank	0.024049	Kemsocinbank	0.030353
Investbank	0.024064	Energobank	0.030408
Kolco Urala	0.025076	Centrocredit	0.031737
Avtovazbank	0.025209	Etalonbank	0.033144
Kuban	0.025299	Credit-Moskow	0.033867
Merkuriy	0.026032	Raschetno-Creditniy	0.035157
Selmashbank	0.026082	BKS-Investicionniy	0.043497
Average probability	0.023809	Average probability	0.03165

Table A8.3 Lists of banks in Russia and Ukraine with the high and low probability to fail next year, predicted by the cross-country probit model

"white list"		"black list"	
Name	Probability to fail	Name	Probability to fail
OTP bank	0.044702	Veles	0.124696
Accent-bank	0.057823	Ikar-bank	0.106187
Industrialbank	0.059391	Concord	0.102681
Kreditprombank	0.060109	Agrarian Commercial bank	0.09626
VAB bank	0.064014	Contract	0.091168
Industrial-Export bank	0.064727	Partner bank	0.090927
Khreschatuk	0.065795	Coopinvestbank	0.089489
Universal bank	0.066947	TMM-bank	0.086711
Home Credit bank	0.071564	Olympic Ukraine	0.08635
Trust-Capital	0.074658	Grant	0.08458
Pivdencombank	0.075506	Invest-Credit bank	0.083749
Kredobank	0.080165	Active bank	0.080166
Average probability	0.06545	Average probability	0.09358
Unicreditbank	0.015113	Credit-Moskow	0.032429
Centrocredit	0.025042	Primorskiy territorialniy	0.029560
Russkobank	0.021793	Aleksandrovskiy	0.026662
Avtovazbank	0.020301	Tempbank	0.027368
Energomashbank	0.023206	Smmit bank	0.030245
Kolco Urala	0.023448	Kemsocinbank	0.033295
Energobank	0.022388	Etalonbank	0.028900
Morskoy bank	0.021140	BKS-Investicionniy	0.040344
Dalcombank	0.018232	Raschetno-Creditniy	0.026029
Kuban	0.025657	Baltiyskiy	0.026718
Selmashbank	0.025934	Merkuriy	0.026076
Average probability	0.022064	Average probability	0.030537
Average probability for 2 countries	0.043757	Average probability for 2 countries	0.066562