NATIONAL RESEARCH UNIVERSITY
HIGHER SCHOOL OF ECONOMICS

Dmitry I. Malakhov, Nikolay P. Pilnik,
Igor G. Pospelov

STABILITY OF DISTRIBUTION OF RELATIVE SIZES OF BANKS AS AN ARGUMENT FOR THE USE OF THE REPRESENTATIVE AGENT CONCEPT

BASIC RESEARCH PROGRAM

WORKING PAPERS

SERIES: ECONOMICS
WP BRP 116/EC/2015

Dmitry I. Malakhov[1], *Nikolay P. Pilnik*[2], *Igor G. Pospelov*[3]

STABILITY OF DISTRIBUTION OF RELATIVE SIZES OF BANKS AS AN ARGUMENT FOR THE USE OF THE REPRESENTATIVE AGENT CONCEPT[4]

We propose a new theoretical model of the large-scale banking system of an open economy. It is shown that distribution of relative sizes of individual banks is stable over time and does not depend on the volume of deposits. Our findings provide an additional argument in favor of use of the representative agent concept in banking sector modeling.

Empirical testing shows that using generalized versions of Pareto and Normal distribution, distributions of relative sizes can be approximated with high accuracy and, moreover, distributions are stable over time. Moreover, banks move wothin this distribution, thus distribution of the general population of banks is stable over time.

Key words: size distribution of banks, representative agent, general equilibrium.

JEL classification: E10, L11, G21

---

[1] National Research University Higher School of Economics. Laboratory for Macrostructure Modeling of Russian Economy; Department of Applied Economics, Assistant Professor; E-mail: dmalakhov@hse.ru

[2] National Research University Higher School of Economics. Laboratory for Macrostructure Modeling of Russian Economy, Senior Research Fellow; Department of Applied Economics, Associate Professor; E-mail:u4d@ya.ru

[3] Federal Research Center "Informatics and Control" of Russian Academy of Science, professor, corresponding member, E-mail: pospeli@yandex.ru

[4] The reported study was supported by Russian Science Foundation (RSF), research project No. 14-11-00432.
We thank Ivan Stankevich, Stanislav Radionov and Sergey Pekarskiy for helpful comments.

Introduction

Microfounded macroeconomic models appeared as a result of the Lucas critique (Lucas (1976)). Modern studies, which analyzes the banking sector in DSGE models, typically use the microfounded approach (this class of models has been developing since the last financial crisis, see, for example, Gertler, Kiyotaki (2010) , Gertler, Karadi (2011), Gertler, Kiyotaki (2013)). These models use the representative agent approach, thus they omit potentially important effects of heterogeneity of real agents (Chang, Kim, Schorfheide (2013)). Agent-based models were developed in attempt to overcome this problem (Farmer, Foley (2009), Borshchev, Filippov (2004)). But this approach has disadvantages such as complexity and the absence of a conventional approach to modeling.

In this paper we show how stability of distribution of relative sizes of banks can be used as an argument for the use of the representative agent concept in general equilibrium models. As proxy for relative bank sizes we used their fraction of assets, because for banks, being financial firms, assets are an accurate approximation of scale. Moreover, the volume of a bank's assets is rather sensitive to the requirements of their products and financial policy (Jagtiani, Khanthavit (1996), James, Houston (1996)). In our paper we paid attention to specific activity of banks such as deposit accumulation and interbank money to investigate the evolution of the structure of the banking industry.

While early money multiplier models, such as Johannes, Rasche (1979), Bernanke, Blinder (1988), Carpenter, Demiralp (2012) did not pay much attention to industry structure and its evolution, the last financial crisis in 2007-2009 revealed that the structure of the financial system is crucial for stability, for the forecasting of its development and for characteristics analysis. So attention from the famous 4 «L», leverage, liquidity, losses, and linkages has now shifted to the last «L», because risk measures for the first 3 «L» are somewhat investigated, but not the last one. Modern studies, which analyze the structure of the financial industry (Acemoglu, Ozdaglar, Tahbaz-Salehi (2015), Billio, Getmansky, Pelizzon (2012), Iori, De Masi, Precup, Gabbi, Caldarelli (2008)) show that the number of linkages and their characteristics are very important for good risk resistance, but these models in fact don't take into account to the reasons for the development of such structures in the banking system. Moreover, analysis of the dynamics or evolution of the industry is very important because it reveals the mechanisms involved in changes to the industry's structure and helps us to make more precise forecasts.

This paper is presented as follows: In the first section we discuss the theoretical model of the banking system. In the second section we provide the result of the empirical testing of our model using data from Russian banks. And finally, we make our conclusions.

1. Model of interbank transactions

We analyzed the distribution of money among banks and the dynamics of the assets of banks. In this section a large scale open economy with a large amount of perfectly competitive banks is discussed.

We showed how the mechanisms of money distribution affect the structure of the banking industry. We didn't impose any restrictions on the banks' heterogeneity, but noted that the distribution of the fraction of bank assets is stable over time. If empirical tests of these results show this to be correct then this can be used as an additional argument for the representative agent concept because if banks are really heterogeneous, then the empirical distribution of their fraction of assets can be potentially multimodal or unstable or not well approximated by standard distribution.

This section is mainly based on Malakhov, Pospelov (2014). The main tool that we used below is the backward Kolmogorov equation for Markov processes. In our case, in fact, it was an accurate total probability formula, which allows to describe all possible variants of history based on their probabilities. Solution of this equation determines the dynamics of the distribution of the elements of the system according to investigated indices.

This approach is used in economic problems related to the description of the dynamics of systems consisting of large numbers of anonymous agents. Similar problems are solved in population game theory the main results are described in Sandholm (2010). In this field there are two distinct approaches to aggregate behavior dynamics: the deterministic approach, based on differential equations, and the stochastic approach, based on Markov processes. The second approach is directly related to our work. Both of these approaches are very complex from a mathematical point of view, so they are not yet the standard for the solution of economic problems.

However, such attempts are undertaken regularly. For an example, also see in Radionov, Pospelov (2014) the discussion on the dynamic of industry with monopolistic competition as in Melitz (2003). In this article, we use it to model the dynamic of the banking system in order to analyze the evolution of distribution of the overall set of banks, rather than behavior of individual banks. Due to this approach we are able to avoid assumptions about the rational behavior of individual agents and the existence of an equilibrium, for example, as in Ericson, Pakes (1995).

1.1. General description of the model

New money appears in the economy through two channels:

1) Loans from outside of the banking system. Residents and non-residents can put their money into national banks. National banks can give credit to residents and non-residents or can use other financial instruments, such as debt emissions.
2) Credit emission. A bank can give credit to a client with a corresponding creating/changing his/her current account.

In this paper we do not consider in this paper the impact of the monetary policy pursued by the central bank. In this sense, we can assume the monetary policy regime does not switch. Also we assume that other regulatory aspects stay the same. Of course, we imposed rather rigid assumptions because, as it was shown in Jagtiani, Khanthavit (1996), regulation can greatly affect the banks' sizes.

In our model we do not differentiate between these two means of money accumulation. . Moreover, we could consider interbank credit as a special case of the transactions listed above. The value of accumulated money $m(t)$ depends only on the number of clients and is independent to conjuncture. We propose that all of the clients are identical to each other (if the number of clients is great (which is true for developed banking system) then this assumption is realistic, and if not, we can divide transaction of one individual into several smaller ones). So, for example, with economy growth the number of clients or value of their emissions could increase, but this does not affect our assumptions.

Withdrawals from a bank occur only when the bank repays its debts or its client's debt relief. We assume that emission generates interest income (which can be potentially negative). Also we propose that all losses are covered and all profits are derived.

A bank can potentially transfer a certain amount of its liabilities to other banks. The bank which initiates the transaction transfers money from the client's account to the correspondent account of the receiving bank. So only the structure of the liabilities of the banks will change (the value will be the same). Also the transaction between clients does not affect the value of the assets of the banking system.

Assume that all credits are repaid with a frequency which is proportional to the duration of the credits β. All assets are identical to each other in the sense of duration (only the moments of creation of bank's assets are changing). Also we consider that all emissions are equal in size and the value of the bank's assets depend only on the number of the bank's clients (we propose that one client during one moment of time can induce only one single emission). Potentially, in reality we can separate all transactions into tranches to hold this assumption.

If we analyze a developed banking system with a large amount of highly competitive banks and identical clients over a rather long period of time, then assumptions about durations and deposit sizes are relevant because we can then average all of the transactions.

Proposing that there is a set of banks B. The fraction of an individual bank's assets in the overall amount is n_b , $1=\sum_{b\in B} n_b$. We assume that set B is stable over time.

Bank b induced the initial emission m, and the transaction is needed with probability $\tau(\cdot)$. We can assume that each emission induces a chain of emissions (as in banking multiplier models), if the amount of banks is great and $\tau(\cdot)$ is small then the corresponding series converge. Moreover, this assumption does not decrease any explanatory power of the model. So let's assume that with probability n_i transaction $b\to i$ is needed. The average assets change after transaction will be

$$\begin{cases} \mathrm{m}, i=b \\ \tau(\cdot)\cdot n_i\,\mathrm{m}, i\neq b \end{cases}$$

1.2. Stochastic process of assets change

Let's assume that at the moment t bank b has assets $\mathrm{a}_b(t)$

$$A(t)=\sum_{b\in B}\mathrm{a}_b(t),$$

During the period $[t,t+dt]$ one client of a randomly chosen bank b, independently of the others, initiates the emission with probability $\alpha(t)\cdot dt$ with size $\mathrm{m}(t)$, which is much smaller than $\mathrm{a}_b(t)$. Note, that $\alpha(t)$ is “real” demand and $\mathrm{m}(t)$ is a proxy for inflation and society welfare. Thus for simplicity we will further consider, that τ is constant.

With probability $\beta \cdot dt$ induced emissions are independently covered. Furthermore, we propose that β is big enough, so all loans are short-term (long-term loans can be divided into parts and analyzed as series of short-term loans) and we can assume that $\mathrm{m}(t)$ can change during the repayment time.

1.3. Dynamic of generalized moments

Now to discuss the averaged value of function $\chi(\mathrm{a})$ over the realization of stochastic process $\mathrm{a}(t)$:

$$\mathrm{X}(t) = E_{\mathrm{a}(t)}\left\{\chi\left(\mathrm{a}(t)\right)\right\},$$

where $E_{\mathrm{a}(t)}\left\{\chi\left(\mathrm{a}(t)\right)\right\}$– mathematical expectation of $\chi\left(\mathrm{a}(t)\right)$ over $\mathrm{a}(t)$. Calculate $X(t+dt)$, $dt > 0$ using the chain rule for mathematical expectation:

$$\mathrm{X}(t+dt) = E_{\mathrm{a}(t)}\left\{E_{[t,t+dt]}\left\{\chi\left(\mathrm{a}(t+dt)\right)\middle|\mathrm{a}(t)\right\}\right\}.$$

During the period $[t+dt]$:

1) with probability $1-\alpha(t)A(t)\,dt-\beta A(t)\,dt$ assets are not changed
2) with probability $\beta\,\mathrm{a}_b(t)\,dt$ assets of bank b decrease by $\mathrm{m}(t)$
3) with probability $\alpha(t)\,\mathrm{a}_b(t)\,dt$ initial emission $\mathrm{m}(t)$ occurs at bank b and with probability $\tau \cdot \dfrac{\mathrm{a}_c(t)}{A(t)-\mathrm{a}_b(t)}$ bank $c \neq b$ continues this transaction. Here we assume, that the probability of continuation of emission is proportional to the amount of assets.

Now we derive conditional mathematical expectation using the probabilities which were mentioned above:

$$E_{[t,t+dt]}\left\{\chi\left(\mathrm{a}(t+dt)\right)\middle|\mathrm{a}(t)\right\} = \left(1-\alpha(t)A(t)\,dt-\beta A(t)\,dt\right)\chi\left(\mathrm{a}(t)\right)+$$

$$+\beta\,dt\sum_{b\in B}\mathrm{a}_b(t)\,\chi\left(\left\{\mathrm{a}_i(t)-m(t)\delta_i^b\right\}_{b\in B}\right)+$$

$$+\alpha(t)\,dt\sum_{b\in B}\mathrm{a}_b(t)\left((1-\tau\cdot\chi\left(\left\{\mathrm{a}_i(t)+m(t)\delta_i^b\right\}_{i\in \mathrm{B}}\right)+ \qquad (5)$$

$$+\tau\sum_{c\in B\setminus\{b\}}\frac{\mathrm{a}_c(t)}{A(t)-\mathrm{a}_b(t)}\chi\left(\left\{\mathrm{a}_i(t)+m(t)\delta_i^b+m(t)\delta_i^c\right\}_{i\in B}\right)\cdot$$

Then we use (4):

$$\frac{d}{dt}\mathrm{X}(t)=E_{\mathrm{a}(t)}[-\left(\alpha(t)A(t)+\beta A(t)\right)\chi\left(\mathrm{a}(t)\right)+\beta\sum_{b\in B}\mathrm{a}_b(t)\chi\left(\left\{\mathrm{a}_i(t)-m(t)\delta_i^b\right\}_{b\in B}\right)+$$

$$+\alpha(t)\sum_{b\in B}\mathrm{a}_b(t)\cdot((1-\tau)\cdot\chi\left(\left\{\mathrm{a}_i(t)+m(t)\delta_i^b\right\}_{b\in B}\right)+$$

$$+\tau\cdot\sum_{c\in B\setminus\{b\}}\frac{\mathrm{a}_c(t)}{A(t)-\mathrm{a}_b(t)}\chi\left(\left\{\mathrm{a}_i(t)+m(t)\delta_i^b+m(t)\delta_i^c\right\}_{b\in\mathrm{a}}\right))]\cdot \tag{6}$$

We can rewrite this difference equation as (for a detailed solution see Appendix 1):

$$\frac{1}{\mathrm{m}(t)}\frac{\partial}{\partial t}\ln f(t,\mathrm{x})+\sum_{b\in B}\left(\left(\alpha(t)-\beta\right)\right)+\sum_{b\in B}\left(x_b\left(\left(\alpha(t)-\beta\right)+\tau\right)\cdot\frac{\partial}{\partial x_b}\ln f(t,\mathrm{x})\right)=$$

$$=-\left|B\right|\cdot\tau\cdot\sum_{b\in B}\frac{x_b}{\sum\limits_{c\in B\setminus\{b\}}x_c}, \tag{7}$$

where $f(t,x)$ - density function of x at time t, x – random variable, which describes the volume of assets of a particular bank, x and a are connected as a random variable and its realization.

We could find a solution to this equation:

$$f(t,x)=\frac{\exp\left(h\left(\frac{x_1}{\sum\limits_{b\in B}x_b},\ldots,\frac{x_{|B|}}{\sum\limits_{b\in B}x_b}\right)\right)}{\exp\left(\ln\left(\sum\limits_{c\in B}x_c e^{-\int_0^t\left(\mathrm{m}(\xi)\left(\alpha(\xi)-\beta\right)\right)d\xi}\right)\sum\limits_{b\in B}\frac{x_b}{\sum\limits_{c\in B\setminus\{b\}}x_c}+\sum\limits_{b\in B}\int_0^t\mathrm{m}(\xi)\left(\alpha(\xi)-\beta\right)d\xi\right)}. \tag{8}$$

Notice, that density function $f(x)$ can be represented as a quotient of two typically different factors. The numerator $\exp\left(h\left(\frac{x_1}{\sum\limits_{b\in B}x_b},\ldots,\frac{x_{|B|}}{\sum\limits_{b\in B}x_b}\right)\right)$ depends only on the fractions of assets, and does not depend on time or the absolute value of assets. The denominator depends on the absolute values of the assets of the individual banks only in the expression $\sum_{b\in B}\frac{x_b}{\sum_{c\in B\setminus\{b\}} x_c}$. If this expression has a constant value over time, then it makes sense to only pay attention to the aggregate value of assets in the denominator. We consider the fraction of assets to be the relative size of the bank.

Since the function f is a function of density, of course, then the integration of all the values of x at any time t gives 1. After the change of variables in the integral, we can go to the $|B|-1$ indexes of fraction and the total absolute amount of assets. Indeed, if we know indexes of fraction of $|B|-1$ banks, and the total absolute amount of assets of banks, we can calculate the volume of the assets of each bank individually. Direct calculation of the Jacobian, which is required during replacing the variables in the integral, transforms our original expression, but allows us to integrate the numerator and denominator separately. It is important that the numerator will remain only fractions of banks' assets, and the denominator - only the total absolute amount of assets. The only variable that remains after the integration on fractions of assets and total amount of assets, is the variable of time t. But this variable is not in the numerators, thus the integral of the denominator will not depend on it Therefore, the function $\exp\left(h\left(\frac{x_1}{\sum_{b\in B} x_b},\dots,\frac{x_{|B|}}{\sum_{b\in B} x_b}\right)\right)$ can be regarded as up to a constant a function of density, depending on only a fraction of the banks' assets. Moreover, since it is not directly dependent on time, this function within our assumptions must be constant.

Thus, according to our model under an unchanging monetary regime, constant number of banks and the absence of structural shocks (such as risk requirements or Central Bank regulation policy) the distribution of relative sizes of banks is stable over time, so growth of the economy (which could increase the number of banks’ clients and/or the volume of their deposits) does not affect the structure of relative sizes of banks. So macro-agent usage could lead to the correct results without any loss of generality due to the stability and homogeneity of banking system.

2. Empirical testing

2.1. Model validation

To validate the model we provided empirical tests. We decided to use the financial statements of banks as our source of information as these are particularly informative[5]. Moreover, we decide to use information about Russian banks because the Russian banking system is rather developed and competitive, especially during the last 10 years, and the data about Russian banks is open and very detailed.

2.1.1. Data

We used information from the '101 turnover balance sheet' of individual credit organizations. The 101 turnover balance sheet is the trial balance with debit and credit subtotals per account, we can therefore get information about assets, deposits, credits and other financial variables from this report. We collected information only from credit organizations, both bank and non-bank organizations, which can provide banking services and are registered in Russia and report balance sheets publicly. The proportion of non-bank credit organizations is very small if we consider either the number of firms or the volume of assets. For simplicity we will name all credit organizations as banks.

Information about the 101 turnover balance sheet is collected from the official website of the Central Bank of the Russian Federation http://www.cbr.ru/. In our sample, on average for the period 2009-2015 we have approximately 99% of the overall number of banks and about 99% of the overall banking system assets for all time periods. So our sample is roughly equal to the amount of Russian banks (for details see Figure 3).

Sub-accounts are rather minor, so they are noisy and are not very representative indicators of the financial health of individual banks. We used aggregate variables because they are very informative, are not so noisy and the number of these variables is not very high. We decided to use the following variables:

1. Total amount of assets.
2. Fixed date deposits of banks and other credit organizations, including overdraft (below we will use abbreviations for financial variables, so this variable is Db)

[5] It is not correct for real sector firms (physical indicators (output, etc.) for firms sometimes are very useful, but rather subjective).

3. Fixed date deposits of non-residents (Df),
4. Fixed date deposits of individuals – residents (Dh),
5. Fixed date deposits of nonfinancial organizations (Da),
6. Fixed date credits to commercial non-bank organizations-residents, including overdraft (La),
7. Fixed date credits to individuals – residents (Lh),
8. Fixed date credits to foreign organizations (Lf).

All our variables are calculated by summing the corresponding sub-accounts of the 101 turnover balance sheet. We chose these variables because they are a significant fraction of the total amount of assets (liabilities). The final data are tables, where columns indicate the time period and rows indicate the bank's I.D. We have separated tables for each financial variable. The period of observation begins in January 2004 and ends in February 2015 (monthly data without any omissions). We have the actual data for each time period.

In our analysis we used the relative size of individual banks as the total amount of the particular variable. So, for example, the fraction of assets of Bank A is the amount of assets of Bank A at the end of month i, divided by the total amount of assets of all the banks in the sample at the end of month i. We used fraction s of assets instead of absolute assets because distribution of fractions is investigated in the first part of our work. Moreover, we needed to not deflate them as they give a relevant picture of the banking system structure. As mentioned before, we would name the fraction of assets as the relative sizes of banks, due to the fact, that the amount of assets is an accurate proxy for bank size.

The number of banks in Russia has changed over time, also the proportion of banks which give information to the Central Bank has changed too, so we had a different number of observations each month. Generally the number of banks didn't vary greatly. The number of banks with non-zero values is approximately 700 at the beginning of the time period and 1100 by the end. It is important to mention that we worked only with banks which provide information to the Central Bank.

We did not drop any banks from our sample, so we estimated distributions including those of the very big banks such as Sberbank and VTB. Moreover, we did not ignore the very small banks, which form a left tail of distribution. This section is mainly based on the paper Malakhov, Pilnik, Radionov (2015).

2.1.1. The possibility of aggregation of the banking system

Let's go back to the equation (8). We plot values of $\sum_{b\in B}\frac{x_b}{\sum_{c\in B\setminus\{b\}} x_c}$ for each month (vertical axes is value of corresponding parameter, horizontal axe is time (January 2004-Febrary 2015)).

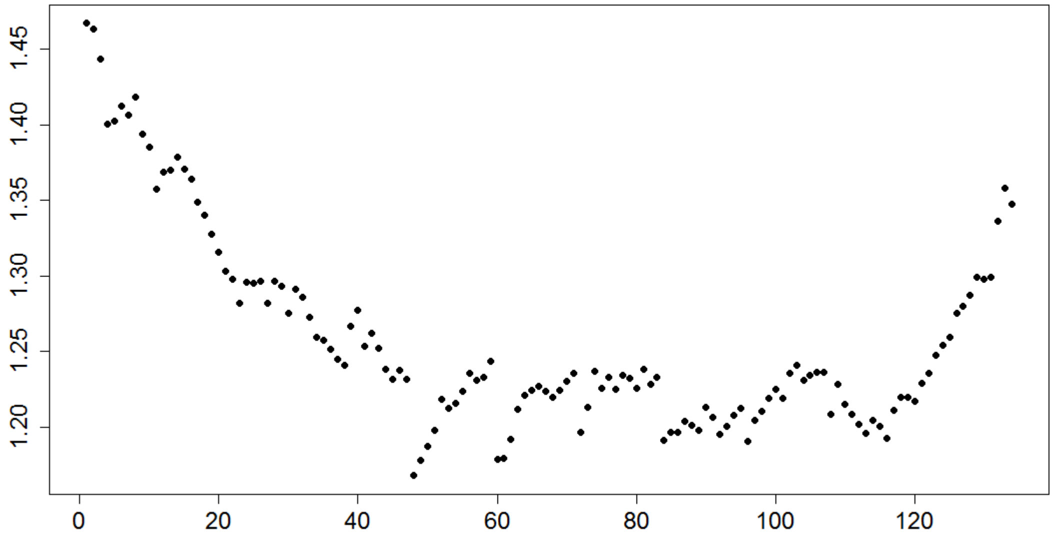


Figure 1. Values of the main factor of (22).

As we can see, for a fairly long period of time from March 2008 (approximately 50th point) to January 2014 (120th point) the value of $\sum_{b\in B}\frac{x_b}{\sum_{c\in B\setminus\{b\}} x_c}$ was rather stable. Consequently, during this period, it can be expected that the distribution of relative sized banks did not change as much as in the remaining periods. Shocks in $\sum_{b\in B}\frac{x_b}{\sum_{c\in B\setminus\{b\}} x_c}$ could be possibly connected with sample changes, monetary policy regime switches and other banking system regulation aspects. It is quite interesting that the "stability" period includes part of the global financial crisis, but it might be connected to the fact that in USA this crisis began at the end of 2007 and by the time the shocks affected the Russian economy Russian banks and the Central Bank already knew how to react and prepared accordingly by providing assistance.

Also we investigate the dynamic of $\ln\left(\sum_{c\in B} x_c\right)$. The chart clearly shows three periods: before the 50th point, from the 50th to 120th and after the 120th. Thus figure 1 represents some realistic trends, not “statistical artifacts”.

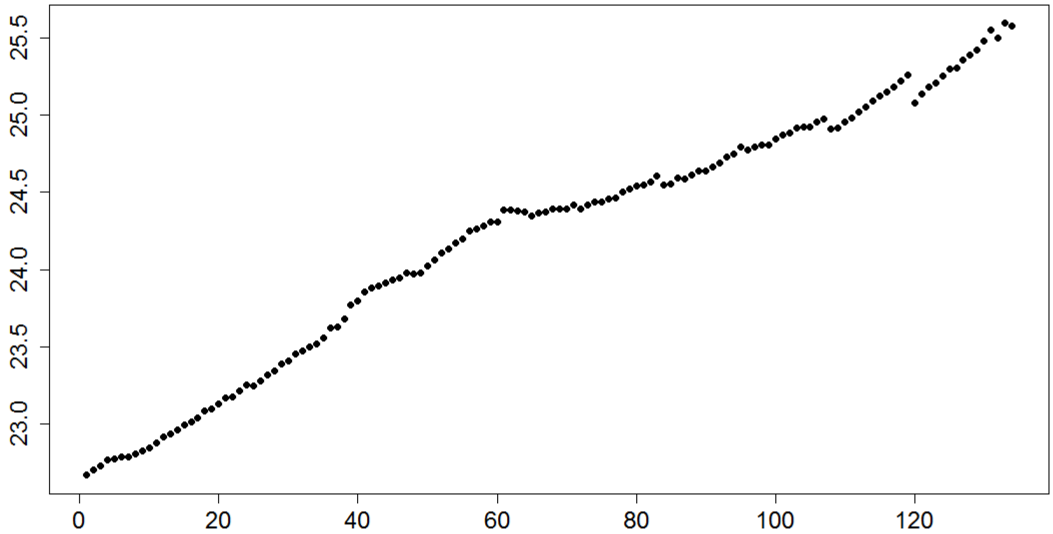


Figure 2. Logarithm of the total assets of the Russian banking system.

2.2. Approaches to modeling distributions of firms’ sizes

The theoretical model showed that distribution of the relative size of banks is stable over time, so these results are highly connected with industry evolution and development. Today there are many works connected with the evolution of particular industries. The classical work is Gibrat (1931), in which the following hypothesis is formulated: a firm’s size and its growth rate are independent. If this hypothesis is true then firm growth rate is independent of its size, so large and small firms have approximately equal growth rates. Researchers tested this hypothesis for many different industries. Generally it is difficult to say if this Law is correct for a real economy or not: firms in some industries grow independently of their size, but firms in other industries don’t (Javonovich (1982)).

Lotti, Santarelli, Vivarelli (2003) provided an empirical test of Gibrat’s Law for young firms. The authors postulate that there are three approaches to Gibrat Law testing. The first approach is the most general: Gibrat’s Law is correct for all firms, independent of their bankruptcy status during the period of observation. The second approach: Gibrat’s Law is correct only for firms which are functioning during the period of observation. The third approach: Gibrat’s Law is correct only for firms whose size is bigger than the minimum efficient scale. The authors provide an extensive research survey (actual for 2003) in which they show that for some industries this Law is correct, and for others - not.

If Gibrat’s Law is correct then firm size can be approximated by lognormal distribution (for example, Gibrat(1931), Axtell (2001)). But fat tails of distribution of firm sizes may occur because positive feedback can exist. So Pareto distribution may also be helpful.

Prescott, Janicki (2006) investigate the data of American banks during 1960-2005 . The authors postulate that lognormal and Pareto distributions are good approximations for bank size, but the right tail of empirical distribution is much fatter than the lognormal one, so they use lognormal distribution as the main distribution for the central part and left tail of data, but the right tail is approximated by Pareto distribution. Also Prescott, Janicki (2006) show that Gibrat’s Law is correct for American banks.

In the paper Cont, Moussa (2010) a quantitative methodology for analyzing the potential for contagion and systemic risk in a network of interlinked financial institutions is presented. This methodology is applied to a data set of mutual exposures and capital levels of financial institutions in Brazil in 2007 and 2008, and the role of balance sheet size and network structure in each institution's contribution to systemic risk is analyzed. Results emphasize the contribution of heterogeneity in network structure and the concentration of counterparty exposures to a given institution in explaining its systemic importance. In this paper, which analyzes the interbank sector, it is shown that the right tail of distribution can be approximated by Pareto distribution. For testing this hypothesis Cont, Moussa (2010) show that linear regression models are good approximations of the logarithmic data.

In Andreev, Pilnik, Pospelov (2009) the authors analyze rang distribution of Russian banks and come to the conclusion that it can be approximated by Pareto distribution with high quality. Moreover, this distribution is stable over time.

A specific parametric form of distribution function of the relative sizes of banks can also be interesting for researchers (see McDonald (1984), Fishlow (1972) for the selection of specific distribution function of household incomes). Moreover, a lot of attention is paid to the connection between income distribution and economic development of the country (see, Galor, Zeira (1987), Greenwood, Jovanovic (1989)).

The selection of a specific functional form of distribution helps us to understand properties of random variable, forecast more precisely its dynamics and calculate indexes of inequality more accurately (see Atkinson, Bourguignon (2000), McDonald , Xu (1995), Kleiber, Kotz(2003)).

Moreover, correct selection of specific distribution can help fill the gaps in data. Today there is a class of works (Miao (2005), Ericson, Pakes (1995)), which discusses the evolution of industry and the connection of firm characteristics with their sizes. But the authors are unaware of any such works which discuss the banking system.

2.3. Preliminary analysis of data

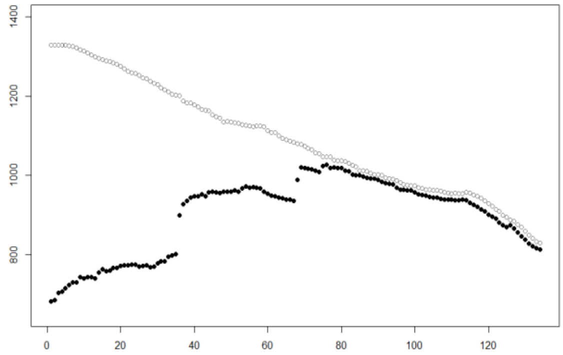
In this part we focus only on the relative sizes of Russian banks[6]. We calculate descriptive statistics for all of the time periods, but for clearer visualization we print only time means:

Table 1. Descriptive statistics

	Value
Stand. deviation	0.01345341
Min	6.414726e-08
Max	0.3835583

So we can see that the difference between the mean smallest bank and the mean largest bank[7] is rather big. So we have some large banks, such as Sberbank, VTB, and very small banks. It is important to mention that the mean relative size is $\frac{1}{n}$, n – number of banks. Moreover, we plot the dynamic of number banks on our sample, standard error of relative sizes, skewness and kurtosis (vertical axes are values of corresponding parameters, horizontal axes are time periods (January 2004 - Febrary 2015)).

[6] Due to space restrictions.
[7] Values are stable over time.



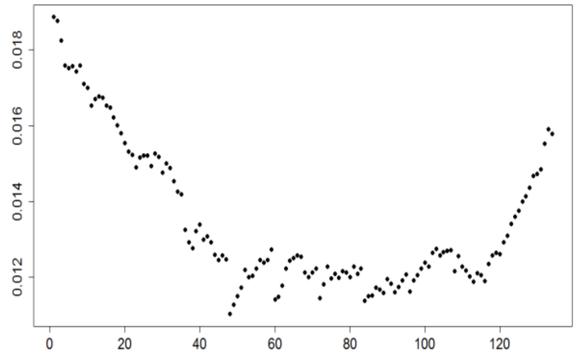
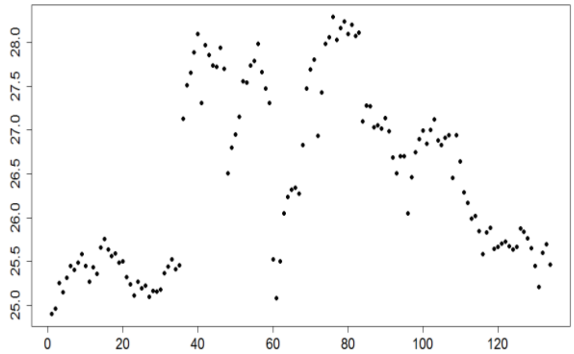


Fig. 3. Number of banks in our sample (black points) and the amount of Russian banks (grey points).

Fig. 4. Dynamic of standard error of relative sizes of banks.



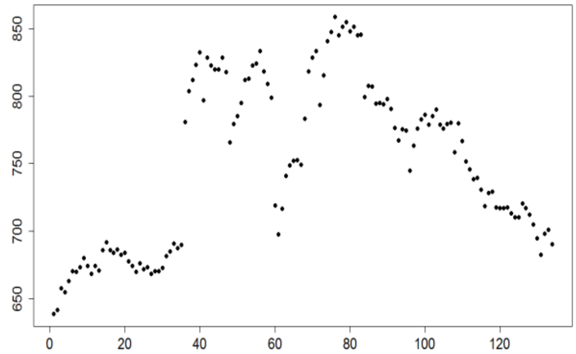


Fig.5. Dynamic of skewness of relative sizes.

Fig.6. Dynamic of kurtosis of relative sizes.

It is clear that since October 2009 (70th point) our sample is approximately equal to the amount of Russian banks. Inequality (measured as standard error) has a nontrivial dynamic, it has been rising since the 70th point. Skewness and kurtosis are far from "normal" ones so we expect that distribution functions will be nontrivial.

2.4. Distribution approximation

In our research we used two families of distribution: Pareto-related distributions (Pareto, distribution, Generalized Pareto distribution (Gen. Pareto), Wakeby distribution, Pareto IV type, Generalized Beta of the second kind (Beta prime distribution)) and Normal-related distribution (Normal, Generalized normal distribution (Gen. normal distribution), Skew normal distribution, asymmetric exponential power distribution (asymmetric generalized error distribution or simply AEP), Generalized lambda distribution (Gen. lambda distribution)). For a detailed description of the distributions and motivation for their selection see Appendix 3.

We used R software for our analysis.[8] Maximum likelihood and L-moments approaches are used because these methods give estimators with "good" properties and they don't depend on chosen distances (Asquith (2015), Hosking (2015), Yee, Wild (1996)).

The quality of approximation is rather stable in terms of time and for all financial variables. Also it is important to mention that we modeled the overall set of banks by the entire distribution, thus we have included in our analysis extremely small or extremely big banks. As an example, we present a graph of the cumulative distribution function (only for one month (May 2012)). We used the relative sizes of banks, as mentioned above. For clearer visualization the top 3-4 banks with very big assets for the family of Pareto distributions were not shown (this will not affect the analysis results because the pattern will be the same).

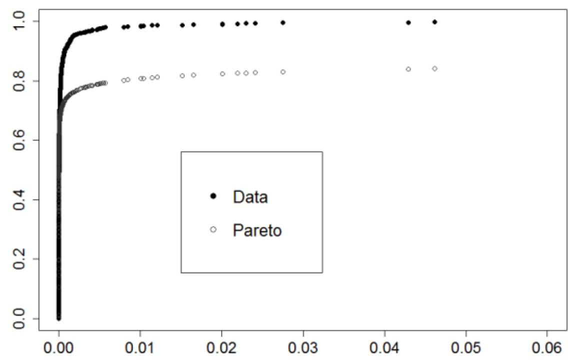


Figure 7. Relative sizes, Pareto distribution.

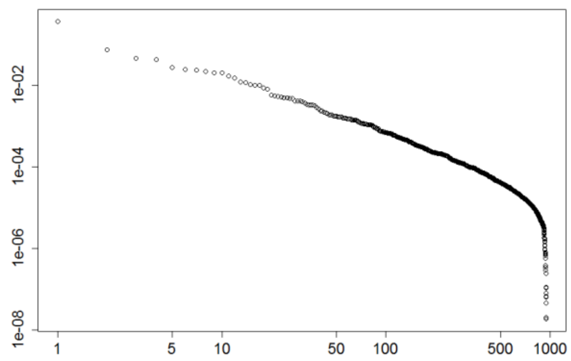


Figure 8. Rang distribution of relative sizes in both logarithmic axes.

The Pareto approximation is not very good however this can easily be explained by low number of parameters and the inflexible functional form. Moreover, the rang distribution graph shows us that this data cannot be correctly approximated by Pareto distribution because it is not a strict line in the double log axes: we can see the curve on the right side of graph (the cluster of small banks). A similar situation occurs with other time periods and financial variables.

We estimated the normal distribution for the log data (for motivation see Appendix 2). As we can see, the quality of approximation is moderate (but higher than for Pareto distribution),

[8] Computer code and dataset could send for request.

because the left tail is approximated (quantiles less than -12) not very well and there is noticeable bias in the middle section (quantiles is around -8), which is connected with the calibration of the right tail. So we move to the analysis of the rest of the distributions.

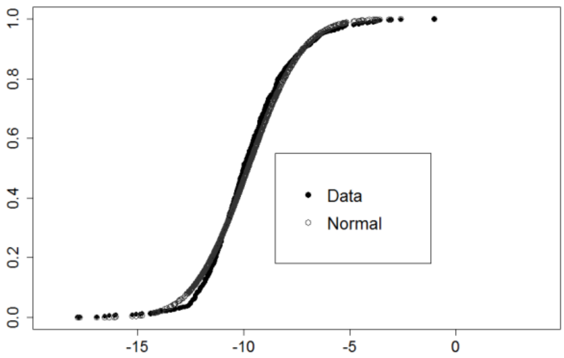


Figure 9. Logarithm of relative sizes, normal distribution.

Below, we discuss the findings of the graphs for 4 variables: fraction of assets, fraction of credits to firms (La), fraction of deposits of households (Dh), and fraction of interbank deposits (Db) (see Appendix 3). We have presented the graphs for May 2012 only because distribution is rather stable over time and patterns will be the same. We decided not to show results for the generalized beta of the second kind because its results are equal to those of the Pareto IVs therefore the graphs become more complicated. Wakeby distribution is turned into Pareto IV for our data, so we excluded this distribution too. Also we assumed that the location parameter for Pareto IV type is 0. This assumption is reconciled with data and does not lead to any limitations (Brazauskas (2003)).

The quality of approximation is rather high for all of the variables. We eliminate skew normal distribution from our further analysis because the estimation results are unstable and their quality is rather low. Pareto and Generalized Pareto distribution cannot approximate the left tail of empirical distribution due to the specific values of estimates of the parameters. The typical normal distribution is not a good approximation as the empirical distribution is in fact rather asymmetric and fat tailed. However, we used Pareto distribution and lognormal distribution as benchmarks. Generalized normal, AEP and Generalized lambda distributions are very accurate approximations. Pareto IV type also gives a very high quality of approximation.

According to the graphical analysis results it is difficult to conclude which distribution is better so we calculated two distances between empirical and theoretical distributions for each month, so called extreme and average distances, respectively:

1) $\max\left|F^{emp}(x) - F^{theor}(x)\right|$

2) $\frac{1}{n_k}\sum_{i=1}^{n_k}\left|F^{emp}(x_i) - F^{theor}(x_i)\right|$, where *n* – number of observations in *k*-th month.

We show maximum, minimum, average and standard error of distances. The following two tables are calculated for relative sizes of banks.

Table 2. Extreme distance for relative size distribution

	Pareto	Gen. Pareto	Pareto IV
Maximum	0.48313	0.18795	0.02455
Minimum	0.34992	0.10712	0.01191
Average	0.41238	0.14349	0.01709
Stand. Err.	0.03508	0.02008	0.00254

	Normal	Gen. Normal	AEP	Gen. Lambda
Maximum	0.06809	0.04451	0.02064	0.04773
Minimum	0.03485	0.02327	0.00972	0.01185
Average	0.05275	0.03178	0.01805	0.01924
Stand. Err.	0.00803	0.00470	0.00314	0.00468

Table 3. Average distance for relative size distribution

	Pareto	Gen. Pareto	Pareto IV
Maximum	0.21907	0.07107	0.00720
Minimum	0.17427	0.03248	0.00323
Average	0.19391	0.05107	0.00506
Stand. err	0.01084	0.01077	0.00093

	Normal	Gen. Normal	AEP	Gen. Lambda
Maximum	0.03006	0.01635	0.00734	0.02029
Minimum	0.01363	0.00929	0.00268	0.00334
Average	0.02378	0.01267	0.00498	0.00575
Stand. err	0.00523	0.00149	0.0008	0.00198

Pareto IV is the best distribution among the Pareto family of distributions and asymmetric exponential power distribution is the best among normal-related distributions. We can easily notice that difference of approximation between asymmetric exponential power and Pareto IV distributions is rather insignificant. For clearer visualization, we present graphs for these distributions for November 2004 and November 2014. November is one of the months with little seasonality effect. Also, it is interesting to investigate the differences in distributions which occurred during the last 10 years.

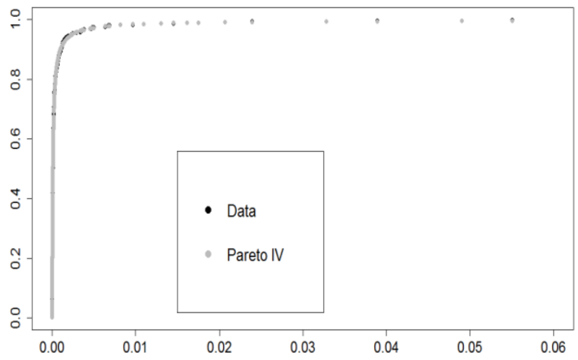


Figure 10. Relative sizes, Pareto IV distribution (November 2004)

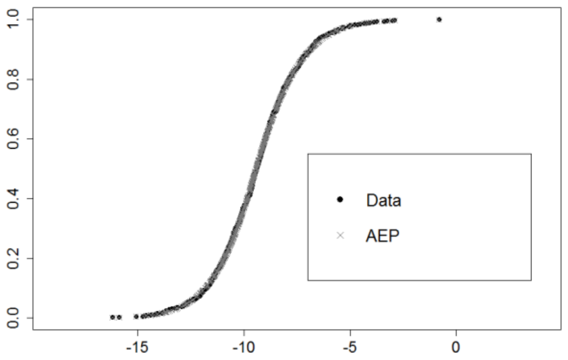
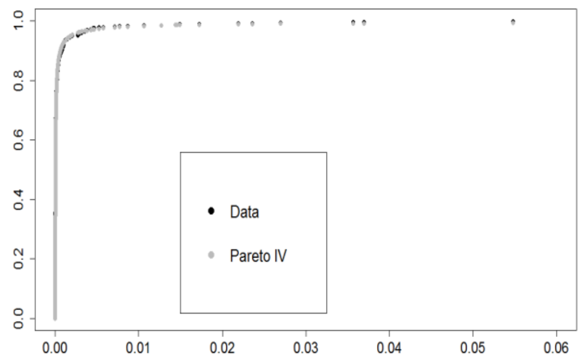


Figure 11. Logarithm of relative sizes, asymmetric exponential power distribution (November 2004)



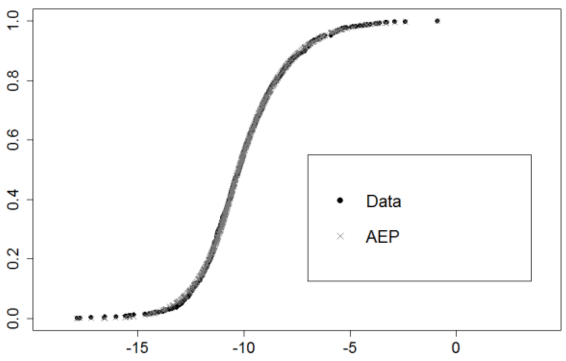


Figure 12. Relative sizes, Pareto IV distribution (November 2014)

Figure 13. Logarithm of relative sizes, asymmetric exponential power distribution (November 2014)

We can see that distribution of relative sizes of banks has not changed significantly during the last 10 years. Both distributions have become more inclined, so in general the banking system has become more homogenous. The recent crisis did not greatly affect (in terms of relative sizes) the banking system because there are no significant changes in the form of distribution. The quality of approximation is very high. Thus we can conclude that the functional form of distributions of relative sizes is stable over time.

We decided to use the Kolmogorov-Smirnov test to compare the two nearest empirical distributions. The selection of empirical distribution as a subject of analysis is due to the fact that empirical distributions are consistent estimates of true ones and it is difficult to conclude which theoretical distribution (Pareto IV or AEP) is better. The mean p-value for comparing the two closest distributions is 0.94, which means that the hypothesis that "both empirical distributions come from the same continuous theoretical distribution" cannot be rejected at a 5% confidence level. If the difference between distributions is more than 8 months, then for some pairs the null hypothesis can be rejected at the 5% confidence level. Thus, the distribution of relative sizes of banks is rather stable over time.

For modeling other variables (different types of deposits and credits, as mentioned earlier) we used all distribution, but AEP and Pareto IV distributions were the best approximation for these variable too (for results see Appendix 4, for space economy we only provided the results for these distributions). It is difficult to choose which distribution is better. Different financial variables are approximated better by different distributions (generally AEP distribution is better for most cases). The distances between empirical and theoretical distributions did not change during our time period (for all our financial variables), so our results are stable. We decided to select both, AEP and Pareto IV as the most appropriate distributions.

We also analyzed graphs of density functions for relative sizes of banks, built with kernel functions and discovered that these functions were rather typical, unimodal and without any "statistical artifacts".[9] Thus this fact can be used as an argument for the concept of representative agent.

We assumed that it is not the distribution of relative sizes of particular banks which is stable, but rather the distribution of banks within the particular banking system. So we explored how individual banks could change their position on the listing and found that there were really many changes in the banks' positions (for detailed results see Appendix 5). We could not find any clear patterns in the banks' movements, but middle sized banks overall had on average much more significant position changes (about 250 banks changed their rang by 400 points and more during the observation period). Also it is quite interesting that banks on average had a negative monthly rang trend, so banks tend to become relatively bigger, thus inequality could increase. This is highly connected to the findings of Malakhov (2015). Yet despite these microchanges, overall distribution is stable. Perhaps, if one bank loses clients/assets other banks accumulate them due to high levels of competition. This fact can also be used as the argument of the representative agent concept because if the system is stable and rather homogeneous, macrolevel variables can be forecasted using only previous values of macrolevel variables if the frequency of data is high enough to reflect changes in economic system.

2.6. Dynamic of estimates of parameters

We made sure that Pareto IV type and AEP are good approximations for distribution of relative sizes of banks. Also, according to formal and graphical analysis, the functional form of distribution is stable over time, so now it is important to define the stability of estimates of parameters of Pareto IV type distribution and AEP distribution. Figures 15-17 show the dynamic of values of parameters of Pareto IV distribution (the first point is January 2004 and the last point is February 2015).

[9] Corresponding graphs can be send for a request.

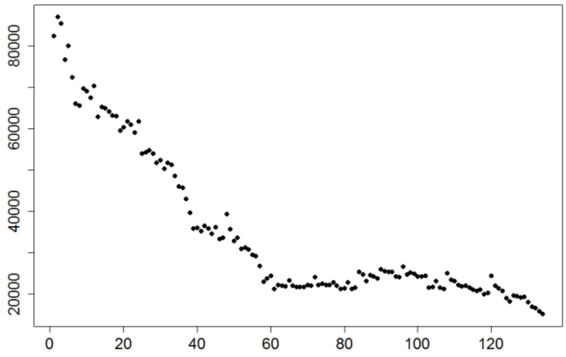


Figure 14. Dynamic of estimates of parameter of scale of Pareto IV distribution

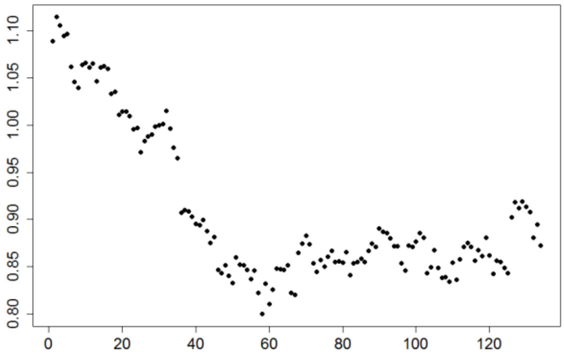


Figure 15. Dynamic of estimates of parameter of inequality of Pareto IV distribution

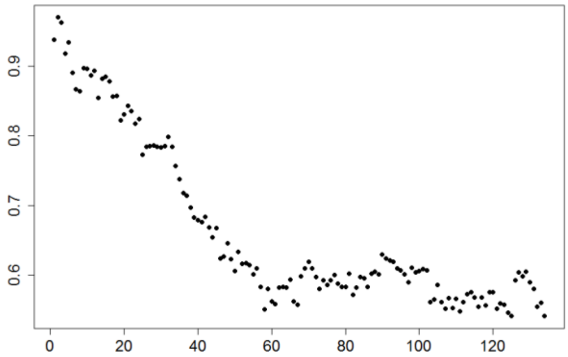


Figure 16. Dynamic of estimates of parameter of shape of Pareto IV distribution

After the 60th point, the values of estimates of parameters become stable (the 60th point is December 2008). We can see only a slight downward movement in estimates of shape and scale parameter and a slight upward trend in estimates of inequality parameter at the end of the period. Also it is important to compare the dynamics of estimates of parameters with the dynamic of $\sum_{b\in B}\frac{x_b}{\sum_{c\in B\setminus\{b\}} x_c}$ (Figure 1).

At the beginning of the paper we mentioned that if $\sum_{b\in B}\frac{x_b}{\sum_{c\in B\setminus\{b\}} x_c}$ is stable over time then distributions are stable over time too. Thus it can be easily seen that periods when $\sum_{b\in B}\frac{x_b}{\sum_{c\in B\setminus\{b\}} x_c}$ is rather stable (for example, the 50-120th point), estimates of the parameters of Pareto IV are rather stable too. Whereas during periods of instability (before 50th point or after 120th point), the estimates of the parameters are again unstable.

Analysis of the dynamic of estimates of parameters of AEP distribution (Figures 18-21) gives similar results. The location and first shape parameters are rather stable from the 60th point to 90th point and after the 90th point have a slight downward trend, possibly connected with the recent crisis. But estimates of scale and second shape parameters are not stable from October 2009 and have a very nontrivial dynamic in the last year. However, this fact can be potentially explained by the procedure of estimating the parameters of AEP distribution. Estimates of parameters of Pareto IV are estimated by a method of L-moments, but estimates of AEP distribution are estimated by a maximum likelihood estimates. Due to the very nontrivial functional form of likelihood function maximization procedure for AEP is not very robust, thus some proportion of deviation of estimate value can be explained by estimation procedure.

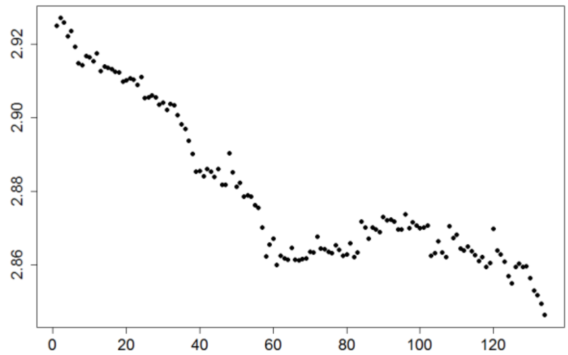


Figure 17. Dynamic of estimates of parameter of location of AEP distribution

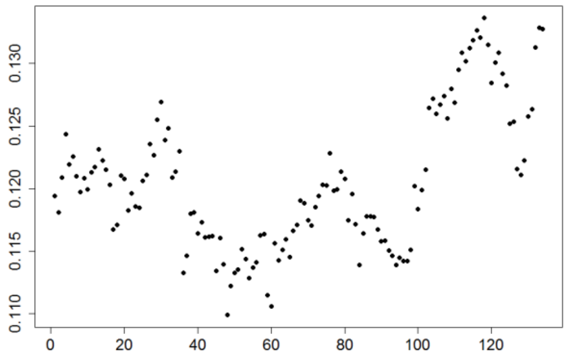


Figure 18. Dynamic of estimates of parameter of scale of AEP distribution

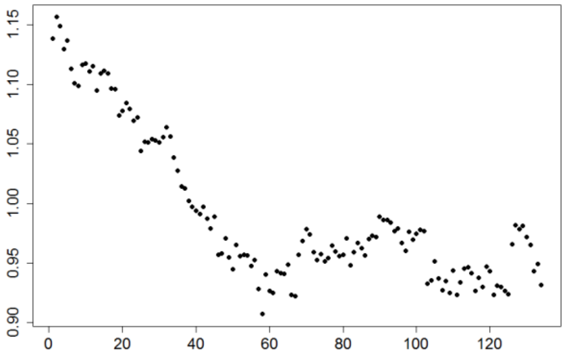


Figure 19. Dynamic of estimates of first parameter of shape of AEP distribution

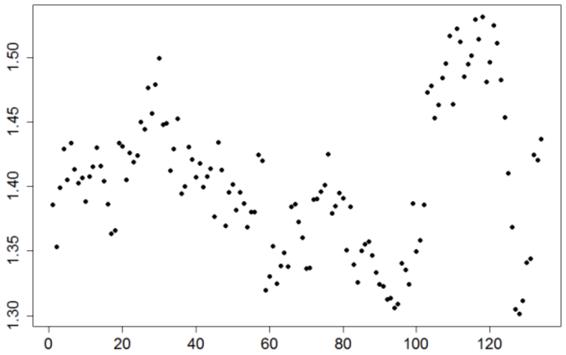


Figure 20. Dynamic of estimates of second parameter of shape of AEP distribution

We can see that estimates of parameters changed during the period of observation but in fact the distributions for 2004 and 2014 are identical, so estimated changes are minor in absolute value (except scale parameter for Pareto IV). Also, it is important to notice again that our sample was changing, so a certain part of the dynamic of estimates can be explained not only by institutional changes of the Russian banking industry, but also by sample changing (see Figure 3), but of course these processes are connected.

Therefore these slow changes in estimates of parameters of distributions in fact show that this banking system is affected by gradualism property. The banking system is changing rather inertly without any serious breaks.

Conclusion

Today there are two main approaches to modeling sectors in macromodels: representative agent and agent-based. We used the tools described in Melitz(2003) and Hopenhayn (1992a) to provide an argument for the use of the concept of representative agents in banking sector modeling.

We have demonstrated that for a large-scale open banking system under assumption, that monetary policy and regulatory aspects do not change the regime and the distribution of relative sizes of individual banks is stable over time. Thus economic growth, which can influence the number of banks clients and/or volume of deposits, does not affect the distribution of relative sizes of banks measured as a fraction of overall assets.

Data from the Russian banking system is used to validate the theoretical model. We discussed that among assets are other key variables such as deposits of households, credit of firms, interbank deposits, etc. Additional variables are selected to make model validation more precise. We show that Pareto IV and asymmetric exponential power distributions are very good approximations for all variables. Moreover, the quality of approximation is stable over time, thus we could say that the functional form of distributions is stable too.

We conducted Kolmogorov-Smirnov tests for empirical distribution functions of relative sizes of banks and found that if the distance between distributions is greater than 8 months, we can reject the null hypothesis for some pairs at a 5% significance level. Moreover, we discovered that individual banks could change their position in the distribution of relative sizes, but overall distribution is stable over time. We did not find any explicit patterns in the movements of Russian banks. Thus distribution of the general population of banks is stable over time.

Estimates of parameter values of Pareto IV distributions are mainly stable over time and values of estimates of location and first shape parameter for AEP distribution are mainly stable too, but estimates of scale and second shape parameters have a rather nontrivial dynamic. This result can be explained by the use of different estimation procedures. Moreover, on average absolute changes in the estimates of parameter values are not very significant. The question about the existence of connection of estimates' dynamic with sample size changes and industry's shocks, especially as monetary policy and banking regulation regimes switch, is relevant.

We can therefore say that our model passes the empirical test and the results are an argument for the use of the concept of representative agent in banking system modeling. Thus, an agent-based approach is not necessary for macromodels of the banking sector. This result could be also useful for precisians because it means that analysts need not pay much attention to differences between banks if they are interested in a macrolevel forecast, especially short run.

For future research projects it will be important to investigate the influence of changes of monetary policy and regulation requirements on the distribution of relative sizes of banks and analyze in detail the dynamics of estimates of parameters. It will be also useful to develop a forecasting technique for predicting the evolution of the banking system. Moreover, it is better to use the data of a banking system of a developed country, such as the USA, to test the theoretical model, as differences between the banking sectors in developed and developing countries can be potentially significant.

References

Acemoglu D., Ozdaglar A., Tahbaz-Salehi A. (2015) "Systemic Risk and Stability in Financial Networks", *American Economic Review*, 105(2): 564-608.

Andreev M. U., Pilnik N. P., Pospelov I. G. (2009) “Modeling of Russian modern industry”, *HSE Economic Journal*, vol. 13, issue №2, pp. 143-171 (in Russian).

Asquith W.H. (2015) “lmomco---L-moments, censored L-moments, trimmed L-moments, L-comoments, and many distributions”, R package.

Atkinson, A. B., & Bourguignon, F. *Handbook of Income Distribution.* 2000. Vol. 1. Elsevier

Axtell, R. L. (2001) “Zipf distribution of US firm sizes”. *Science*, 293(5536), 1818-1820.

Borshchev A., & Filippov A. (2004) “From system dynamics and discrete event to practical agent based modeling: reasons, techniques, tools”. *Proceedings of the 22nd International Conference of the System Dynamics Society*. 2004. Vol. 22.

Beena V. T., Kumaran M. (2010) “Measuring inequality and social welfare from any arbitrary distribution”. *Brazilian Journal of Probability and Statistics*, 24(1), pp. 78-90.

Benito E. (2008) “Size, growth and bank dynamics”. *Banco de España Working Paper No. 0801*.

Bernanke, B. S., Blinder, A. S. (1988) “Credit, money, and aggregate demand”, *NBER WP № w2534*

Billio, M., Getmansky, M., Lo, A. W., Pelizzon, L. (2012). “Econometric measures of connectedness and systemic risk in the finance and insurance sectors”, *Journal of Financial Economics*, *104*(3), 535-559.

Brazauskas V. (2003) "Information matrix for Pareto (IV), Burr, and related distributions". *Communications in Statistics-Theory and Methods*, 32(2), pp. 315-325.

Buldyrev S. V., Growiec J., Pammolli F., Riccaboni M., Stanley H. E. (2007) "The growth of business firms: Facts and theory". *Journal of the European Economic Association*.5(2-3), pp. 574-584.

Cabral L.M.B., Mata J. (2003) "On the evolution of size distributions: facts and theory". *The American Economic Review*, volume 93, №4, pp.1075-1090.

Carpenter, S., Demiralp, S. (2012). "Money, reserves, and the transmission of monetary policy: Does the money multiplier exist?". *Journal of macroeconomics*, *34*(1), 59-75.

Chang Y., Kim S. B., Schorfheide F. (2013) "Labor-market heterogeneity, aggregation, and police (in) variance of DSGE model parameters". *Journal of the European Economic Association*. 11(s1), 193-220.

Cont, R., Moussa, A. (2010) "*Network structure and systemic risk in banking systems". Edson Bastos e, Network Structure and Systemic Risk in Banking Systems* (December 1, 2010).

Crosato L., Ganugi P. (2007) "Statistical regularity of firm size distribution: the Pareto IV and truncated Yule for Italian SCI manufacturing". *Statistical Methods and Applications*, 16(1), pp. 85-115.

Dawid H., Gemkow S., Harting P., VanderHoog S., Neugart M. (2012) "The eurace@ unibi model: An agent-based macroeconomic model for economic policy analysis". Universitat Bielefeld, *Working Papers in Economic and Management*, №05-12

Dedova M. S., Pilnik N. P., Pospelov I. G., (2014) "Description of liquidity demand of Russian banking system using turnover statistics". *Journal of New Economic Association*, №4(24), pp. 87-110 (in Russian).

Eeckhout, J. (2004) “Gibrat's Law for (All) Cities”. *The American Economic Review*, Vol. 94, №5, pp.1429-1451

Ericson, R., & Pakes, A. (1995). “Markov-perfect industry dynamics: A framework for empirical work”. *The Review of Economic Studies*, *62*(1), 53-82.

Evans D. S. (1987) “The relationship between firm growth, size, and age: estimates for 100 manufacturing industries”. *Journal of Industrial Economics*, vol. 35(4), pp. 567-581.

Farmer J. D., Foley D. (2009) “The economy needs agent-based modeling”. *Nature*. 460(7256), 685-686.

Fishlow A. (1972) “Brazilian size distribution of income”. *The American Economic Review*. Vol 62. № ½, Pp. 391-402.

Galor O., Zeira J. (1987) “Income distribution and macroeconomics”. *The Review of Economic Studies*. 60(1), 35-52.

Gertler M., Karadi P. (2011) “A model of unconventional monetary policy”. *Journal of Monetary Economics*. 58(1), 17-34

Gertler M., Kiyotaki N. (2010) “Financial intermediation and credit policy in business cycle analysis”. *Handbook of Monetary Economics*. 3(3), 547-599

Gertler M., Kiyotaki, N. “Banking, liquidity and bank runs in an infinite-horizon economy”. *NBER WP № w19129*

Gibrat R. (1931) “Les Inégalités économiques”. *Recueil Sirey.*

Gonzales-Val R., Lanaspa L., Sanz F. (2010) “Gibrat’s Law for Cities Revisited”. *MRPA Paper № 20169*

Gordy, M. B. (1998) “*A generalization of generalized beta distributions*”. Division of Research and Statistics, Division of Monetary Affairs, Federal Reserve Board.

Greenwood J., Jovanovic, B. (1989) “Financial development, growth, and the distribution of income”. *NBER WP № w3189*.

Hall B. N. (1986) “The relationship between firm size and firm growth in the U.S. manufacturing center”. *NBER WP № 1965*.

Hopenhayn, H. (1992a) "Entry, Exit, and Firm Dynamics in Long Run Equilibrium", *Econometrica,*vol. 60, №. 5, pp. 1127-1150

Hosking, J. R. M. (2015). L-moments. R package.

Iori, G., De Masi, G., Precup, O. V., Gabbi, G., & Caldarelli, G. (2008). “A network analysis of the Italian overnight money market.” *Journal of Economic Dynamics and Control*, *32*(1), 259-278.

Jagtiani, J., Khanthavit, A. (1996) “Scale and scope economies at large banks: Including off-balance sheet products and regulatory effects (1984–1991)”. *Journal of Banking & Finance*, *20*(7), 1271-1287.

James, C., Houston, J. (1996) “Evolution or Extinction: Where are banks headed?”. *Journal of Applied Corporate Finance*, 9(2), 8-23.

Javonovich, B. (1982) “Selection and the evolution of industry”. *Econometrica*, volume 50, issue 3, pp. 649-670.

Johannes, J. M., Rasche, R. H. (1979). “Predicting the money multiplier”. *Journal of Monetary Economics*, *5*(3), 301-325.

Kechedzhi K. E., Usatenko O. V., Yampol’skii V. A. (2005) “Rank distributions of words in correlated symbolic systems and the Zipf law”. *Physical Review E*, 72. №. 4, С. 046138.

Kleiber, C., Kotz, S. *Statistical size distributions in economics and actuarial sciences.* 2003. Vol. 470. John Wiley & Sons

Lotti F., Santarelli E., Vivarelli M. (2003) “Does Gibrat’s Law Hold Among Young, Small Firms”. *Journal of Evolutionary Economics,* volume 13, pp. 213-235.

Lucas R. E. (1976) “Econometric policy evaluation: A critique” *In Carnegie-Rochester Conference Series on Public Policy*. 1976, December. Vol. 1, pp. 19-46. North-Holland.

Malakhov D.I., (2015) “Risks of structural changes in industry and their consequence”. *Journal of Risk Management in Credit Organization,* Vol. 20, №4, pp. 72-88 (in Russian).

Malakhov D.I., Pilnik N.P., Radionov S.A. (2015) “Stability of distribution of banks as an argument to concept of aggregate agent”. *HSE Economics Journal*, Vol. 19, №4 (in Russain).

Malakhov, D. I., Pospelov I. G. (2014) “Evolution of distribution of Russian banks on key aggregate variables”. *Proceeding of MIPTs,* vol.6, №4(24), pp. 57-66 (in Russian).

Mansfield E. (1962) “Entry, Gibrat's Law, Innovation, and the Growth of Firms”. *American Economics Review,* vol. 52, № 5, Pp. 1023-1051.

McDonald J.B. (1984) “Some generalized functions for the size distribution of income”. *Econometrica: Journal of the Econometric Society*. Pp. 647-663.

McDonald J. B., Xu Y.J. (1995) “A generalization of the beta distribution with applications”. *Journal of Econometrics*, vol. 66, № 1, pp. 133–152.

Melitz M. (2003) “The impact of trade on intraindustry reallocations and aggregate industry productivity”. *Econometrica,* vol. 71(6), pp.1695-1725.

Miao J. (2005) “Optimal capital structure and industry dynamics”. *The Journal of Finance.* Vol. 60(6), 2621-2659.

Negrea B. (2014) “A statistical measure of financial crises magnitude”. *Physica A: Statistical Mechanics and its Applications*. Vol. 397, pp. 54-75.

Prescott E.S., Janicki H.P. (2006) “Changes in the Size of U.S. Banks: 1960-2005”. *Economic Quarterly*. issue Fall, pp. 291-316

Radionov S., Pospelov I. G. (2014) “Firm Dynamics in Melitz-Type Models”. *Mathematical Models and Computer Simulations.* Vol. 6. No. 5. P. 445-455.

Sepanski J. H., Kong L. (2007) “A family of generalized beta distributions for income”. *arXiv Preprint, arXiv:0710.4614*.

Sandholm W. (2010) Population games and Evolutionary dynamics. MIT Press. 567 p.

Sutton J. (1991) “Sunk costs and market structure : price competition, advertising, and the evolution of concentration”. *MIT Press,* Cambridge, Mass.

Wang Z. (2008) “Income distribution, market size and the evolution of Industry”. *Review of Economic Dynamics,* vol.11, issue 3, pp. 542-565.

Yee T. W., Wild C. J. (1996). “Vector Generalized Additive Models”. *Journal of Royal Statistical Society,* series B, vol 58. №3, pp.481-493.

Zhu D., Zinde-Walsh V. (2009) "Properties and estimation of asymmetric exponential power distribution". *Journal of Econometrics,* vol. 148, №1. Pp. 86-99

Appendix 1. Theoretical model

In subsection 1.3. we get, that (6) is equal to:

$$\frac{d}{dt}\mathrm{X}(t)=E_{\mathrm{a}(t)}[-\left(\alpha(t)A(t)+\beta A(t)\right)\chi\left(\mathrm{a}(t)\right)+\beta\sum_{b\in B}\mathrm{a}_b(t)\chi\left(\left\{\mathrm{a}_i(t)-m(t)\delta_i^b\right\}_{b\in B}\right)+$$

$$+\alpha(t)\sum_{b\in B}\mathrm{a}_b(t)\cdot\left(\left(1-\tau\right)\cdot\chi\left(\left\{\mathrm{a}_i(t)+m(t)\delta_i^b\right\}_{b\in B}\right)+\right.$$

$$\left.+\tau\cdot\sum_{c\in B\setminus\{b\}}\frac{\mathrm{a}_c(t)}{A(t)-\mathrm{a}_b(t)}\chi\left(\left\{\mathrm{a}_i(t)+m(t)\delta_i^b+m(t)\delta_i^c\right\}_{b\in \mathrm{a}}\right)\right)]\cdot$$

We assume with probability equal to 1, that $\mathrm{a}_i(t)\square\; m(t)$, but $\mathrm{a}_i(t)\cdot m(t)$ is finite. Then:

$$\chi\left(\left\{\mathrm{a}_i(t)+m(t)\delta_i^b\right\}_{b\in \mathrm{B}}\right)=\chi(\mathrm{a}(t))+m(t)\cdot\partial_b\chi(\mathrm{a}(t)), \tag{8}$$

$$\chi\left(\left\{\mathrm{a}_i(t)-m(t)\delta_i^b\right\}_{b\in \mathrm{B}}\right)=\chi(\mathrm{a}(t))-m(t)\cdot\partial_b\chi(\mathrm{a}(t)), \tag{9}$$

$$\chi\left(\left\{\mathrm{a}_i(t)+m(t)\delta_i^b+m(t)\delta_i^c\right\}_{b\in \mathrm{B}}\right)=\chi(\mathrm{a}(t))+m(t)\cdot\partial_b\chi(\mathrm{a}(t))+m(t)\cdot\partial_c\chi(\mathrm{a}(t)), \tag{10}$$

Substitute the expressions (8), (9), (10) into (6) and using (2):

$$\frac{d}{dt}\mathrm{X}(t)=m(t)\cdot\tau\cdot E_{\mathrm{a}(t)}[\left(\alpha(t)-\beta\right)\sum_{b\in B}\mathrm{a}_b(t)\partial_b\chi(\mathrm{a}(t))+$$

$$+\alpha(t)\sum_{b\in B}\frac{\sum_{c\in B\setminus\{b\}}\mathrm{a}_b(t)\mathrm{a}_c(t)\partial_c\chi(\mathrm{a}(t))}{\sum_{c\in B\setminus\{b\}}\mathrm{a}_c(t)}]\;. \tag{11}$$

We know, that for each smooth function $\varphi:\mathrm{R}^1\to\mathrm{R}^1$:

$$\int_{-\infty}^{\infty}\delta(x-y)\cdot\varphi(y)dy=\int_{-\infty}^{\infty}\delta(y-x)\cdot\varphi(y)dy=\varphi(x),$$

$$\int_{-\infty}^{\infty}\delta'(x-y)\cdot\varphi(y)dy=-\int_{-\infty}^{\infty}\delta'(y-x)\cdot\varphi(y)dy=-\varphi'(x),$$ where $\delta(\cdot)$ - Dirac function.

If $\chi^{\delta}(\mathrm{a})=\prod_{b\in B}\delta(\mathrm{a}_b-x_b)$, x_b - constant parameters, and distribution of random variable $\mathrm{a}(t)$ near x has smooth density $f(t,\mathrm{a})\cdot d\,\mathrm{a}$, then:

$\mathrm{X}(t)\square\; E_t\left\{\chi^{\delta}(\mathrm{a}(t))\right\}=f(t,\mathrm{x})$ - density of joint distribution.

t is parameter, so:

$$\frac{d}{dt}\mathrm{X}(t)=\frac{\partial}{\partial t}f(t,\mathrm{x}), \tag{12}$$

$$\partial_b\chi^\delta(\mathrm{a})=\prod_{i\in B\setminus\{b\}}\delta(\mathrm{x}_i-\mathrm{a}_i)\cdot\delta'(\mathrm{x}_i-\mathrm{a}_i),$$

Using expression given above:

$$E_{\mathrm{a}(t)}\left\{\mathrm{a}_b(t)\partial_b\chi^\delta(\mathrm{a}(t))\right\}=$$

$$=\int_0^\infty d\,\mathrm{a}_1\underset{\neq b}{\ldots}\int_0^\infty d\,\mathrm{a}_{|B|}\prod_{i\neq b}\delta(\mathrm{a}_b-x_b)\int_0^\infty d\,\mathrm{a}_b\ f(t,\mathrm{a})\cdot\mathrm{a}_b\cdot\delta'(\mathrm{a}_b-x_b)=-\frac{\partial}{\partial x_b}\left(x_b\cdot f(t,x)\right).$$

Simplifying expression above, we get:

$$E_{\mathrm{a}(t)}\left\{\sum_{b\in B}\tau\frac{\sum_{c\in B\setminus\{b\}}\mathrm{a}_c(t)\mathrm{a}_b(t)\partial_c\chi(\mathrm{a}(t))}{\sum_{c\in B\setminus\{b\}}\mathrm{a}_c(t)}\right\}=$$

$$=\sum_{b\in B}E_{\mathrm{a}(t)}\left\{\sum_{c\in B\setminus\{b\}}\frac{\tau\cdot\mathrm{a}_c(t)\mathrm{a}_b(t)\partial_c\chi(\mathrm{a}(t))}{\sum_{d\in B\setminus\{b\}}\mathrm{a}_d(t)}\right\}=$$

$$=\sum_{b\in B}\sum_{c\in B\setminus\{b\}}\int_0^\infty d\,\mathrm{a}_1\underset{\neq b,c}{\ldots}\int_0^\infty d\,\mathrm{a}_{|B|}\prod_{i\neq c}\delta(\mathrm{a}_i-x_i)\int_0^\infty d\,\mathrm{a}_c\ f(t,\mathrm{a})\cdot\frac{\tau\cdot\mathrm{a}_c\mathrm{a}_b\,\delta'(\mathrm{a}_c-x_c)}{\sum_{d\in B\setminus\{b\}}\mathrm{a}_d}=$$

$$=\sum_{b\in B}\sum_{c\in B\setminus\{b\}}\int_0^\infty d\,\mathrm{a}_c\ f(t,\mathrm{x}_1,\ldots.,\mathrm{a}_c,\ldots,\mathrm{x}_{|B|})\cdot\frac{\tau\cdot\mathrm{a}_c\,x_b\delta'(\mathrm{a}_c-x_c)}{\sum_{d\in B}x_d-x_b-x_c+\mathrm{a}_c}=$$

$$=-\sum_{b\in B}\sum_{c\in B\setminus\{b\}}\left(\frac{\partial}{\partial x_c}\left(f(t,\mathrm{x})\cdot\tau\cdot\frac{x_cx_b}{\sum_{d\in B\setminus\{b\}}x_d}\right)\right). \tag{13}$$

Substitute the expressions (12) and (13) into (11):

$$\frac{1}{\mathrm{m}(t)}\frac{\partial}{\partial t}f(t,\mathrm{x})+\sum_{b\in B}\left(\alpha(t)-\beta\right)\frac{\partial}{\partial x_b}\left(x_b\cdot f(t,\mathrm{x})\right)+$$

$$+\alpha(t)-\tau\cdot\sum_{b\in B}\sum_{c\in B\setminus\{b\}}\left(\frac{\partial}{\partial x_c}\left(f(t,x)\cdot\frac{x_c x_b}{\sum\limits_{d\in B\setminus\{b\}} x_d}\right)\right)=0. \tag{14}$$

Simplify the derivatives in (14) and divide both sides of equation by $f(t,x)$:

$$\frac{1}{\mathrm{m}(t)}\frac{\partial}{\partial t}\ln f(t,\mathbf{x})+\sum_{b\in\mathrm{B}}\left(\left(\alpha(t)-\beta\right)\right)+\sum_{b\in\mathrm{B}}\left(\mathbf{x}_b\left(\left(\alpha(t)-\beta\right)\right)\cdot\frac{\partial}{\partial\mathbf{x}_b}\ln f(t,\mathbf{x})\right)=$$

$$=-\sum_{b\in\mathrm{B}}\sum_{c\in\mathrm{B}\setminus\{b\}}\left(\tau\cdot\frac{\mathbf{x}_c\mathbf{x}_b}{\sum\limits_{d\in\mathrm{B}\setminus\{b\}}\mathbf{x}_d}\frac{\partial}{\partial\mathbf{x}_c}\ln f(t,\mathbf{x})+\right.$$

$$\left.+\frac{\mathbf{x}_c\mathbf{x}_b}{\sum\limits_{d\in\mathrm{B}\setminus\{b\}}\mathbf{x}_d}\frac{\partial}{\partial\mathbf{x}_c}\tau+\tau\cdot\frac{\partial}{\partial\mathbf{x}_c}\left(\frac{\mathbf{x}_c\mathbf{x}_b}{\sum\limits_{d\in\mathrm{B}\setminus\{b\}}\mathbf{x}_d}\right)\right)=$$

$$=-\sum_{b\in\mathrm{B}}\left(\tau\cdot\mathbf{x}_b\frac{\partial}{\partial\mathbf{x}_c}\ln f(t,\mathbf{x})+\mathbf{x}_b\frac{\partial}{\partial\mathbf{x}_c}\tau+\tau\cdot\sum_{c\in\mathrm{B}\setminus\{b\}}\left(\frac{\mathbf{x}_b}{\sum\limits_{d\in\mathrm{B}\setminus\{b\}}\mathbf{x}_d}-\frac{\mathbf{x}_c\mathbf{x}_b}{\left(\sum\limits_{d\in\mathrm{B}\setminus\{b\}}\mathbf{x}_d\right)^2}\right)\right)$$

Than

$$\frac{1}{\mathrm{m}(t)}\frac{\partial}{\partial t}\ln f(t,\mathrm{x})+\sum_{b\in B}\left(\left(\alpha(t)-\beta\right)\right)+\sum_{b\in B}\left(x_b\left(\left(\alpha(t)-\beta\right)+\tau\right)\cdot\frac{\partial}{\partial x_b}\ln f(t,\mathrm{x})\right)=$$

$$=-\sum_{b\in B}\left(\tau\cdot x_b\cdot\left(\frac{|\mathrm{B}|-2}{\sum\limits_{d\in B\setminus\{b\}} x_d}\right)\right),$$

$$\frac{1}{\mathrm{m}(t)}\frac{\partial}{\partial t}\ln f(t,x)+\sum_{b\in B}\left(\left(\alpha(t)-\beta\right)\right)+\sum_{b\in B}\left(x_b\left(\left(\alpha(t)-\beta\right)\right)\cdot\frac{\partial}{\partial x_b}\ln f(t,x)\right)=$$

$$= -\tau \cdot \sum_{b \in B} \left(x_b \frac{\partial}{\partial x_c} \ln f(t,x) + \sum_{c \in B \setminus \{b\}} \left(\frac{x_b}{\sum_{d \in B \setminus \{b\}} x_d} - \frac{x_c x_b}{\left(\sum_{d \in B \setminus \{b\}} x_d \right)^2} \right) \right). \tag{15}$$

If we integrate second and third summands over the whole space and in each element integrate over x_b, then this integral will be equal to 0, because $f(t,x)$ turns into 0 at the edges.

Because number of banks $|B| \square\ 2$, so we can write down:

$$\frac{1}{\mathrm{m}(t)} \frac{\partial}{\partial t} \ln f(t,\mathrm{x}) + \sum_{b \in B} \left(\left(\alpha(t) - \beta \right) \right) + \sum_{b \in B} \left(x_b \left(\left(\alpha(t) - \beta \right) + \tau \right) \cdot \frac{\partial}{\partial x_b} \ln f(t,\mathrm{x}) \right) =$$

$$= -|B| \cdot \tau \cdot \sum_{b \in B} \frac{x_b}{\sum_{c \in B \setminus \{b\}} x_c} \quad . \tag{16}$$

We can make in (16) following substitution:

$$\ln f(t,x) = \ln g \left(t, x e^{-\int_0^t \left(\mathrm{m}(\xi) \left(\alpha(\xi) - \beta \right) \right) d\xi} \right) - \sum_{b \in B} \int_0^t \mathrm{m}(\xi) \left(\alpha(\xi) - \beta \right) d\xi \tag{17}$$

Then we substitute (17) into (16):

$$\partial_t \ln g \left(t, x e^{-\int_0^t \left(\mathrm{m}(\xi) \left(\alpha(\xi) - \beta \right) \right) d\xi} \right) - \sum_{b \in B} \left(\alpha(t) - \beta \right) -$$

$$- \sum_{b \in B} \left(\alpha(t) - \beta \right) \cdot x_b e^{-\int_0^t \left(\mathrm{m}(\xi) \left(\alpha(\xi) - \beta \right) \right) d\xi} \cdot \partial_b \ln g \left(t, x e^{-\int_0^t \left(\mathrm{m}(\xi) \left(\alpha(\xi) - \beta \right) \right) d\xi} \right) + \sum_{b \in B} \left(\left(\alpha(t) - \beta \right) \right) +$$

$$+ \sum_{b \in B} \left(x_b e^{-\int_0^t \left(\mathrm{m}(\xi) \left(\alpha(\xi) - \beta \right) \right) d\xi} \left(\left(\alpha(t) - \beta \right) + \tau \right) \cdot \partial_b \ln g \left(t, x e^{-\int_0^t \left(\mathrm{m}(\xi) \left(\alpha(\xi) - \beta \right) \right) d\xi} \right) \right) =$$

$$= -|B| \cdot \tau \cdot \sum_{b \in B} \frac{x_b}{\sum_{c \in B \setminus \{b\}} x_c}.$$

We make the following change of variables:

$$y = xe^{-\int\limits_0^t \left(\mathrm{m}(\xi)\left(\alpha(\xi)-\beta\right)\right)d\xi}$$

After simplification we get:

$$\partial_t \ln g\left(t,\mathrm{y}\right)+\tau\cdot\sum_{b\in B} y_b\cdot\partial_b \ln g\left(t,\mathrm{y}\right) = -\left|B\right|\cdot\tau\cdot\sum_{b\in B}\frac{x_b}{\sum\limits_{c\in B\backslash\{b\}} x_c}.$$

Functions, which depend on x are homogenous of zero degree, so:

$$\frac{\partial}{\partial t}\ln g\left(t,\mathrm{y}\right)+\tau\cdot\sum_{b\in B} y_b\cdot\frac{\partial}{\partial y_b}\ln g\left(t,\mathrm{y}\right) = -\left|B\right|\cdot\tau\cdot\sum_{b\in B}\frac{y_b}{\sum\limits_{c\in B\backslash\{b\}} y_c}.$$

Notice, that this equation is independent on time, so we can find the stationary solution:

$$\ln f(t,\mathrm{x}) = \ln g\left(xe^{-\int\limits_0^t \left(\mathrm{m}(\xi)\left(\alpha(\xi)-\beta\right)\right)d\xi} \right) - \sum_{b\in B}\int\limits_0^t \mathrm{m}(\xi)\left(\alpha(\xi)-\beta\right)d\xi,$$

$$\sum_{b\in B} y_b\tau\cdot\frac{\partial}{\partial y_b}\ln g\left(y\right)+\left|B\right|\sum_{b\in B}\tau\cdot\frac{y_b}{\sum\limits_{c\in B\backslash\{b\}} y_c} = 0.$$

We can write the preceding relation in the following form:

$$\sum_{b\in B} y_b\frac{\partial}{\partial y_b}\ln g\left(y\right)+\left|B\right|\sum_{b\in B}\frac{y_b}{\sum\limits_{c\in \mathrm{B}\backslash\{b\}} y_c} = 0. \qquad (18)$$

General solution of homogeneous equation:

$$\sum_{b\in B} y_b\frac{\partial}{\partial y_b}\ln g\left(y\right) = 0.$$

Free term of (18) can be expressed as series:

$$\sum_{b\in B}\frac{y_b}{\sum\limits_{c\in B\backslash\{b\}} y_c} = \sum_{b\in B}\frac{y_b}{\sum\limits_{c\in B} y_c - y_b} = \sum_{b\in B}\sum_{n=1}^{\infty}\left(\frac{y_b}{\sum\limits_{c\in B} y_c}\right)^n = \sum_{n=1}^{\infty}\frac{S_n}{S_1^{\,n}},\ S_n \,\square\, \sum_{b\in B} y_b^{\,n}.$$

We can find particular solution as series:

$$\ln g\left(y\right) = \sum_{n=1}^{\infty} K_n(S_1)\cdot S_n \ ,$$

$$\frac{\partial}{\partial y_b}\ln g\left(y\right) = \sum_{n=1}^{\infty} n\cdot K_n(S_1)\cdot y_b^{\,n-1} + \sum_{n=1}^{\infty} K_n'(S_1)\cdot S_n \,,$$

$$\sum_{b\in B} y_b \frac{\partial}{\partial y_b} \ln g(y) = \sum_{n=1}^{\infty} \left(n\cdot K_n(S_1) + K_n'(S_1)\cdot S_1 \right)\cdot S_n .$$

The series, which were mentioned above, can be solution if:

$$n\cdot K_n(S_1) + K_n'(S_1)\cdot S_1 = -\frac{1}{S_1{}^n} .$$

Homogeneous equation $n\cdot K_n(S_1) + K_n'(S_1)\cdot S_1 = 0 \Leftrightarrow n\cdot \frac{dS_1}{S_1} + \frac{dK_n}{K_n} = 0$ has solution $\tilde{K}_n(S_1) = C\cdot S_1{}^{-n}$. So we can use a variation of parameters:

$$n\cdot C(S_1)\cdot S_1{}^{-n} + C'(S_1)\cdot S_1{}^{-n+1} - n\cdot C(S_1)\cdot S_1{}^{-n-1}\cdot S_1 = -\frac{1}{S_1{}^n} ,$$

$$C'(S_1) = -\frac{1}{S_1} \Rightarrow C(S_1) = -\ln S_1 + C_0 \cdot$$

We need only a particular solution, so assume, that $C_0 = 0$, then $K_n(S_1) = \frac{-\ln S_1}{S_1{}^n}$ and so the partial solution will be:

$$\sum_{n=1}^{\infty} K_n(S_1)\cdot S_n = -\ln S_1 \sum_{n=1}^{\infty} \frac{S_n}{S_1{}^n} = -\ln\left(\sum_{c\in B} y_c \right) \sum_{b\in B} \frac{y_b}{\sum_{c\in B\setminus\{b\}} y_c} .$$

The general solution of the equation is following:

$$g(y) = \exp\left\{ h\left(\frac{y_1}{\sum_{b\in B} y_b}, \dots, \frac{y_{|B|}}{\sum_{b\in B} y_b} \right) - \ln\left(\sum_{c\in B} y_c \right) \sum_{b\in B} \frac{y_b}{\sum_{c\in B\setminus\{b\}} y_c} \right\} .$$

Now return to the original variables $y = xe^{-\int_0^t (m(\xi)(\alpha(\xi)-\beta))d\xi}$:

$$g(xe^{-\int_0^t (m(\xi)(\alpha(\xi)-\beta))d\xi}) = \exp\left[h\left(\frac{x_1}{\sum_{b\in B} x_b}, \dots, \frac{x_{|B|}}{\sum_{b\in B} x_b} \right) - \ln\left(\sum_{c\in B} x_c e^{-\int_0^t (m(\xi)(\alpha(\xi)-\beta))d\xi} \right) \sum_{b\in B} \frac{x_b}{\sum_{c\in B\setminus\{b\}} x_c} \right]$$

.

We notice again that $g(xe^{-\int_0^t \left(m(\xi)\left(\alpha(\xi)-\beta\right)\right)d\xi})$ does not depend on time, so, substituting

$$\ln f(t,x) = \ln g\left(t, xe^{-\int_0^t \left(m(\xi)\left(\alpha(\xi)-\beta\right)\right)d\xi}\right) - \sum_{b\in B}\int_0^t m(\xi)\left(\alpha(\xi)-\beta\right)d\xi$$

we get:

$$f(t,x) = \exp\left(\begin{array}{l} h\left(\frac{x_1}{\sum_{b\in B} x_b}, \ldots, \frac{x_{|B|}}{\sum_{b\in B} x_b}\right) - \ln\left(\sum_{c\in B} x_c e^{-\int_0^t \left(m(\xi)\left(\alpha(\xi)-\beta\right)\right)d\xi}\right)\sum_{b\in B}\frac{x_b}{\sum_{c\in B\setminus\{b\}} x_c} - \\ -\sum_{b\in B}\int_0^t m(\xi)\left(\alpha(\xi)-\beta\right)d\xi \end{array} \right),$$

And thus we get (8):

$$f(t,x) = \frac{\exp\left(h\left(\frac{x_1}{\sum_{b\in B} x_b}, \ldots, \frac{x_{|B|}}{\sum_{b\in B} x_b}\right)\right)}{\exp\left(\ln\left(\sum_{c\in B} x_c e^{-\int_0^t \left(m(\xi)\left(\alpha(\xi)-\beta\right)\right)d\xi}\right)\sum_{b\in B}\frac{x_b}{\sum_{c\in B\setminus\{b\}} x_c} + \sum_{b\in B}\int_0^t m(\xi)\left(\alpha(\xi)-\beta\right)d\xi\right)}.$$

Appendix 2. Used distributions

Pareto and Lognormal distributions

This part is mainly influenced by Kleiber, Kotz (2003) and McDonald (1984).

Random variable x has power distribution, if probability density function is $f(x) \sim x^{-(1+\alpha_0)}$, where α_0 - parameter. Often Pareto distribution is used as the basic power distribution:

$$F(x) = \begin{cases} 1-\left(x/x_0\right)^{-\alpha} & ,\text{if } \ x > x_0 \\ 0 & ,\text{if } \ x < x_0 \end{cases}$$

Much attention is paid to parameter value α. If $\alpha < 1$, then large values have a significant effect on the average value. If $\alpha > 1$, then observations with small values have a significant effect on the average. Pareto distribution is rather popular in modeling distribution of firm sizes (Axtell (2001), Crosato, Ganugi (2007)).

If Gibrat's Law is correct for the particular sample of firms, then we can use lognormal distribution for firms' sizes, as it is shown in Gibrat's original paper. So we pay a lot of attention to lognormal and Pareto distributions and their quality of approximation.

Presently there are a lot of generalizations of these distributions. Generalized distributions are very important tools, when data set is rather heterogeneous and it is difficult to use basic distributions, because of their small numbers of parameters. Larger number of parameters and much more flexible functional form help to get precise results. But when we estimate parameters of these distributions we face a lot of problems. Numerical procedures very often can't converge and estimation results are not very precise and robust. In our particular case we successfully avoid many estimation problems because dataset is homogenous and rather large. We use maximum likelihood and method of L-moments estimation procedures as the most reliable methods with good properties of estimators.

It is important to note that further distributions have been divided into two families rather roughly, sometimes there is no strict connection with the base distribution. This classification is made only for simplicity.

Family of Pareto-related distributions

Generalized Pareto distribution (Gen. Pareto) has three parameters: location, scale and shape. Including of shape parameter allows generalizing standard Pareto distribution:

$$F(x)=\begin{cases}1-(1+\xi z)^{-1/\xi} & ,if \quad \xi\neq 0\\ 1-e^{-z} & ,if \quad \xi=0\end{cases},$$

where $z=\dfrac{x-\mu}{\sigma}$, μ - location parameter, σ –scale parameter, ξ - shape parameter.

Wakeby distribution is a generalization of generalized Pareto distribution. In special cases this distribution is equivalent to Pareto, exponential and uniform distributions. Cumulative distribution function is very complicated and it is easier to use quantile function:

$$x(\mathrm{U})=\xi+\frac{\alpha}{\beta}(1-(1-U)^{\beta})-\frac{\gamma}{\delta}(1-(1-U))^{-\delta},$$

U - uniform random variable with support [0,1], ξ-location parameter, α,β-scale parameters, γ,δ - shape parameters.

So, Wakeby is a very general distribution with 5 parameters, it is widely used in financial application (for example, Negrea (2014)).

Another way generalization of Pareto distribution is **Pareto IV type**. Pareto IV is a generalization of many different Pareto type distributions and has the following cumulative distribution function:

$$F(x)=\begin{cases}1-\left(1+\left(\dfrac{x-\mu}{\sigma}\right)^{\frac{1}{\gamma}}\right)^{-\alpha}, & if \;\; x>\mu\\ 0, & if \;\; x\leq\mu\end{cases},$$

where $\mu-$ location parameter, $\sigma-$ scale parameter, $\gamma-$ shape (inequality) parameter, $\alpha-$ tail parameter.

Pareto IV type is one of the most popular generalized Pareto distributions. Location parameter affects mathematical expectation and tail fatness, scale parameter stretches cumulative distribution and probability density functions along the OX axes. Gini coefficient is greatly affected by shape parameter. And tail parameter has significant impact on tail fatness. Due to its flexibility Pareto IV type is very often used in firm size modeling Crosato, Ganugi (2007).

Generalized Beta of the second kind (Beta prime distribution) is a generalization of many different distributions, including Pareto IV type distribution. Probability density function for Generalized Beta of the second kind has the following functional form:

$$f(x)=\frac{a\left(\frac{x}{b}\right)^{ap-1}}{b\cdot B(p,q)\cdot\left(1+\left(\frac{x}{b}\right)^{\alpha}\right)^{p+q}},\ x\in(0,+\infty),$$

where α, p, q - shape parameters and b - scale parameter.

Beta prime distribution is very often used in income distribution analysis and financial modeling, but there are many problems with the parameters' estimation of this distribution, because its probability density function is rather complicated.

Family of Normal-related distributions

Generalized normal distribution (Gen. normal distribution) has probability density function: $F(x)=1-\exp(-\mathrm{y})$,where

$$y=\begin{cases}-\frac{1}{\kappa}log[1-\frac{\kappa(x-\xi)}{\alpha}], \text{ if } \kappa\neq 0\\ \frac{(x-\xi)}{\alpha}, \qquad if\, \kappa=0\end{cases}$$

where ξ,α,κ - parameters of location, scale and shape, respectively. According to the three parameters, this distribution is much more flexible than typical normal distribution. Moreover, this distribution is skewed and it can be very important in the modeling of income or firm sizes distribution, because this data is typically asymmetrically distributed.

Skew normal distribution has the following probability density function:

$$f(x)=\frac{1}{\omega\pi}\cdot e^{-\frac{(x-\xi)^2}{2\omega^2}}\int_{-\infty}^{\alpha\left(\frac{x-\xi}{\omega}\right)}e^{\frac{t^2}{2}}dt,$$

where ξ,ω,α - parameters of location, scale and shape, respectively. This distribution has a very important feature – its distribution function is asymmetric, which is very relevant for our dataset. During parameter estimation problems may occur, for example, probability density functions are calculated using simulation methods, so accuracy and robustness of parameters' estimates are not very high.

We also used **asymmetric exponential power distribution (asymmetric generalized error distribution or simply AEP)** with cumulative probability function:

$$F(x)=\begin{cases}\dfrac{\kappa^2}{1+\kappa^2}\gamma\left(\left[\dfrac{\xi-x}{\alpha\kappa}\right]^h,\dfrac{1}{h}\right), x<\xi\\ 1-\dfrac{\kappa^2}{1+\kappa^2}\gamma\left(\left[\dfrac{\xi-x}{\alpha\kappa}\right]^h,\dfrac{1}{h}\right), x\geq\xi\end{cases},$$

where $\gamma(\mathrm{Z},\mathrm{a})=\dfrac{\int_{Z}^{\infty} y^{\alpha-1}\cdot\exp(-\mathrm{y})\,\mathrm{dy}}{\Gamma(\mathrm{a})}$, $\Gamma(\mathrm{a})-$ incomplete gamma function, ξ - location parameter, α - scale parameter, κ, h - shape parameters.

Asymmetric exponential power distribution was developed as an asymmetric generalization of exponential power distribution (also known as generalized error distribution), which in turn is a generalization of normal distribution with kurtosis parameter. It is important to note that asymmetric exponential power distribution has maximum entropy in the very wide class of distributions with support $(-\infty;+\infty)$ (Zhu, Zinde-Walsh (2009)). Moreover, tails of this distribution can potentially have different fatness and they are much fatter than normal ones. In the paper Buldyrev, Growiec, Pammolli, Riccaboni, Stanley (2007) asymmetric exponential power distribution is used for modeling firms sizes.

Generalized lambda distribution (Gen. lambda distribution) is also used for modeling data. Quantile function as follows:

$x(\mathrm{U})=\xi+\alpha(\mathrm{U}^{\kappa}-(1-U)^h)$,

U - uniform distribution with support [0,1], ξ - location parameter, α - scale parameter, κ, h - shape parameters.

Generalized lambda distribution is asymmetric distribution with power law tails, so it is often used for financial modeling. Generalized lambda distribution was developed as an approximation of many standard distributions, because it was originally used in Monte Carlo modeling, so flexibility of this distribution should be very high. In Beena, Kumara (2010) generalized lambda distribution is used for modeling inequality of income distribution.

Appendix 3. Distribution figures

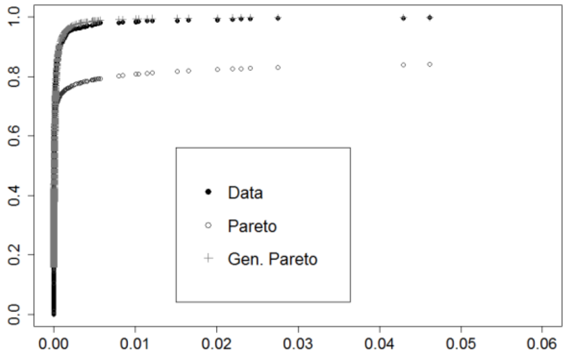


Figure 21. Relative sizes, family of Pareto distribution (Pareto, Gen. Pareto)

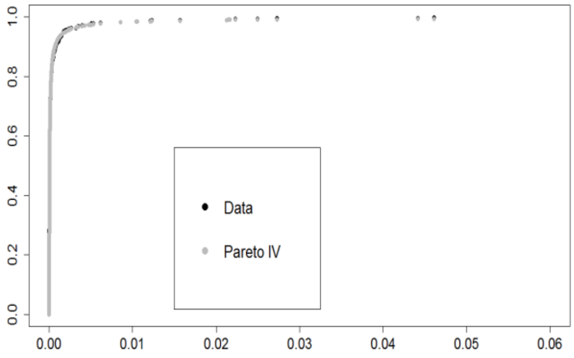


Figure 22. Relative sizes, family of Pareto distribution (Pareto IV)

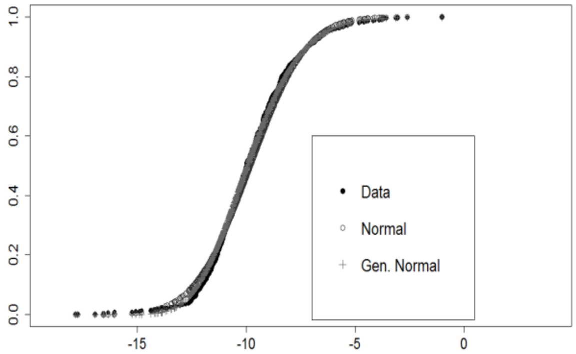


Figure 23. Logarithm of relative sizes, family of normal-related distributions (Normal, Gen. Normal)

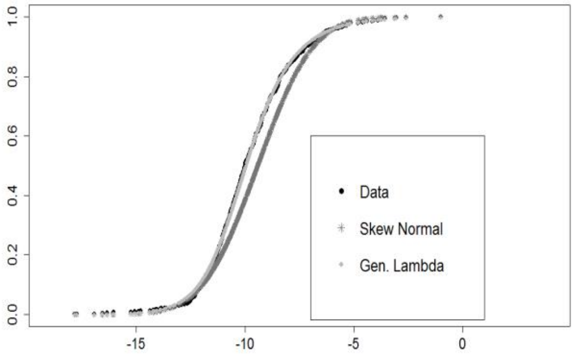


Figure 24. Logarithm of relative sizes, family of normal distributions (Skew Normal, Gen. Lambda)

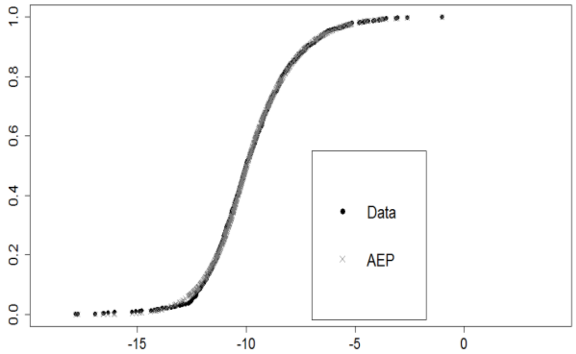


Figure 25. Logarithm of relative sizes, family of normal-related distributions (AEP)

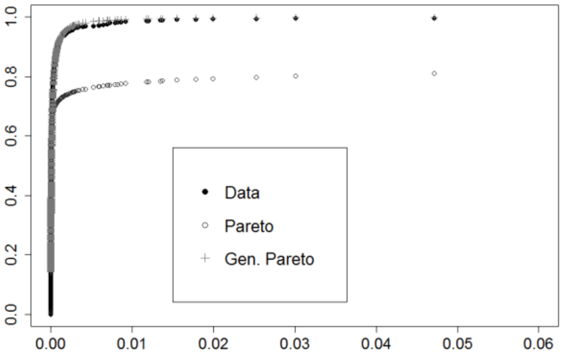


Figure 26. Fraction of credits to firms (La), family of Pareto distributions (Pareto, Gen. Pareto)

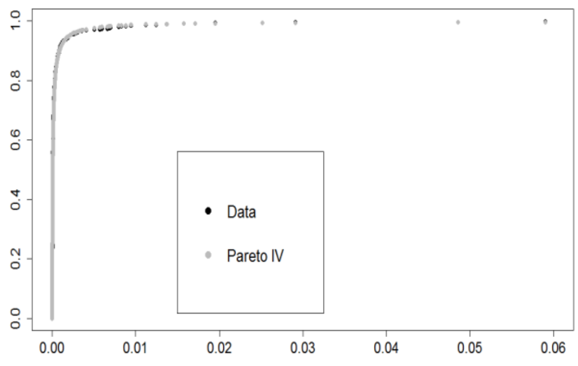


Figure 27. Fraction of credits to firms (La), family of Pareto distributions (Pareto IV)

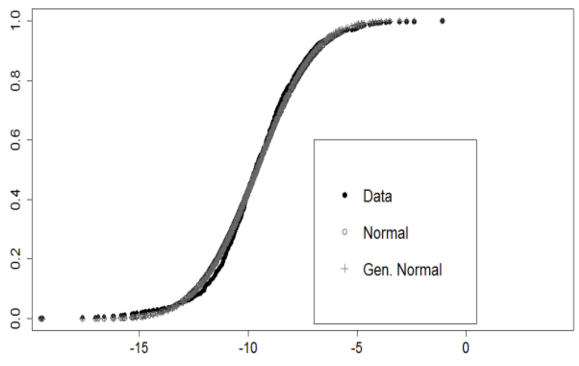


Figure 28. Logarithm of fraction of credits to firms (La), family of normal-related distributions (Normal, Gen. Normal)

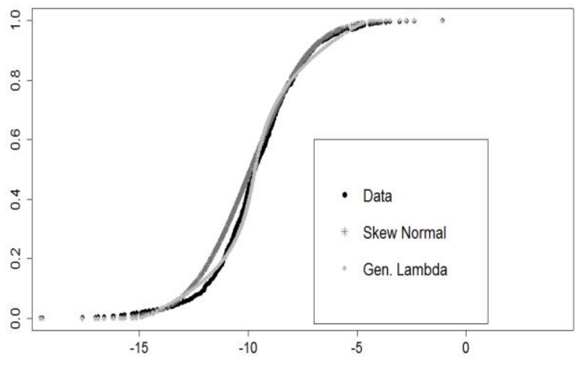


Figure 29. Logarithm of fraction of credits to firms (La), family of normal distributions (Skew Normal, Gen. Lamda)

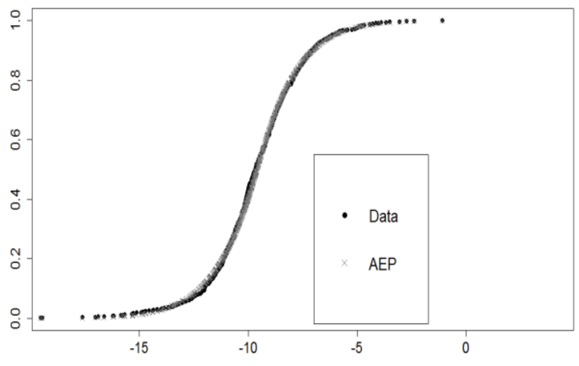


Figure 30. Logarithm of fraction of credits to firms (La), family of normal distributions (AEP)

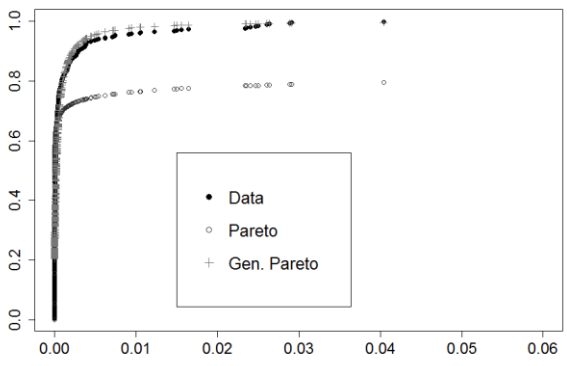


Figure 31. Fraction of interbank deposits (Db), family of Pareto distributions (Pareto, Gen. Pareto)

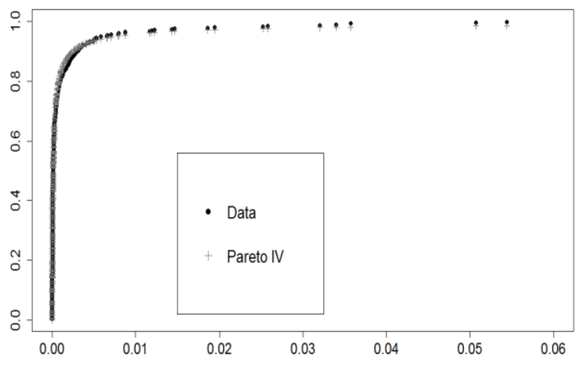


Figure 32. Fraction of interbank deposits (Db), family of Pareto distributions (Pareto IV)

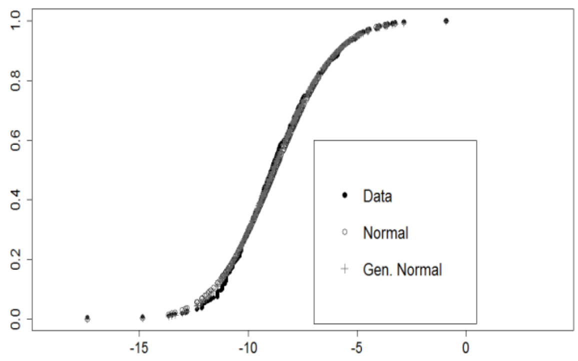


Figure 33. Logarithm of fraction of interbank deposits (Db), family of normal-related distributions (Normal, Gen. Normal)

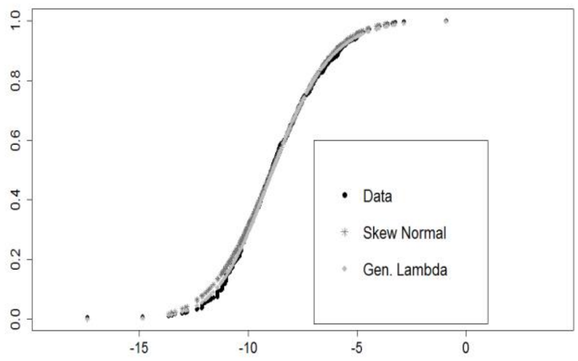


Figure 34. Logarithm of fraction of interbank deposits (Db), family of normal-related distributions (Skew Normal, Gen. Lambda)

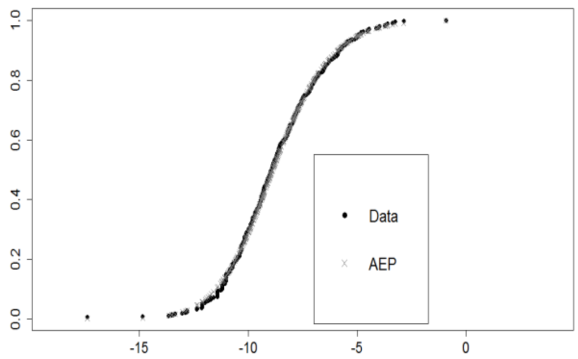


Figure 35. Logarithm of fraction of interbank deposits (Db), family of normal-related distributions (AEP)

Appendix 4. Maximum and mean distances for other financial variables

		Maximum	Minimum	Average	Stand.Err.
Db	Max, Pareto IV	0.0561	0.0217	0.0381	0.0068
	Max, AEP	0.0478	0.0163	0.0293	0.0068
	Mean, Pareto IV	0.0211	0.0073	0.0135	0.0029
	Mean, AEP	0.0182	0.0043	0.0090	0.0028
Df	Max, Pareto IV	0.0942	0.0318	0.0546	0.0165
	Max, AEP	0.0574	0.0173	0.0296	0.0078
	Mean, Pareto IV	0.0377	0.0093	0.0196	0.0078
	Mean, AEP	0.0213	0.0048	0.0099	0.0039
Dh	Max, Pareto IV	0.0601	0.0171	0.0325	0.0100
	Max, AEP	0.0472	0.0168	0.0312	0.0081
	Mean, Pareto IV	0.0200	0.0065	0.0121	0.0038
	Mean, AEP	0.0209	0.0056	0.0117	0.0041
Da	Max, Pareto IV	0.0401	0.0131	0.0260	0.0059
	Max, AEP	0.0429	0.0145	0.0269	0.0062
	Mean, Pareto IV	0.0149	0.0040	0.0087	0.0025
	Mean, AEP	0.0162	0.0044	0.0089	0.0026
Lf	Max, Pareto IV	0.0531	0.0163	0.0318	0.0065
	Max, AEP	0.0441	0.0142	0.0252	0.0053
	Mean, Pareto	0.0172	0.0054	0.0102	0.0027

	IV				
	Mean, AEP	0.0127	0.0042	0.0072	0.0017
Lh	Max, Pareto IV	0.0355	0.0128	0.0222	0.0043
	Max, AEP	0.0326	0.0093	0.0158	0.0048
	Mean, Pareto IV	0.0113	0.0041	0.0073	0.0019
	Mean, AEP	0.0097	0.0029	0.0047	0.0014
La	Max, Pareto IV	0.0335	0.0151	0.0240	0.0042
	Max, AEP	0.0387	0.0162	0.0262	0.0051
	Mean, Pareto IV	0.0116	0.0044	0.0080	0.0019
	Mean, AEP	0.0118	0.0052	0.0079	0.0016

Appendix 5. Dynamics of ranks of individual banks

The conclusion of the stability of the distribution of relative sizes of banks does not mean that the position (rank) of banks in this distribution remains constant. In this section we show how much some banks may change their positions in the distribution by the example of banks' assets.[10] For this we order all the banks in accordance with the rank of their assets at the end of the observation period (February 2015) and calculate the difference between the highest and lowest rank for the entire period of observation. Those banks, which at the end of the period had revoked license, we place on the right side of the x-axis (after position 814). For these banks the difference was calculated only for the period prior to license revocation. As we can see in Figure 36, the vast majority of banks shifted more than 100 positions, with nearly 250 of them shifting more than 400 positions. This means that individual banks are constantly changing their position along quite a stable distribution.

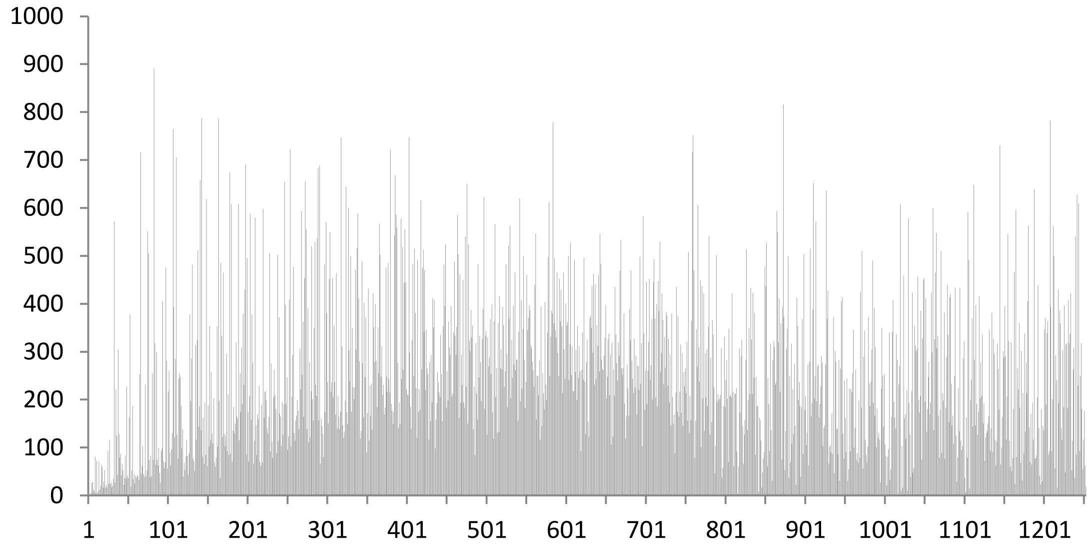


Figure 36. The difference between the highest and lowest ranked bank by assets during the period of observation. The abscissa is the rank of the bank at the end of the observation period (February 2015). Banks, whose license was revoked, are located after position 814.

Figures 37 and 38 describe the same process over the past three years and the last year, correspondingly. We see that there is a sufficient number of banks, whose ranks over the year

[10] Obviously, it is equivalent to analyze assets' rang of banks or rang of relative sizes of banks.

changed by more than 200 points. We could not find any patterns in banks' movements; moreover, we can't say that any cluster of banks is more active, than others.

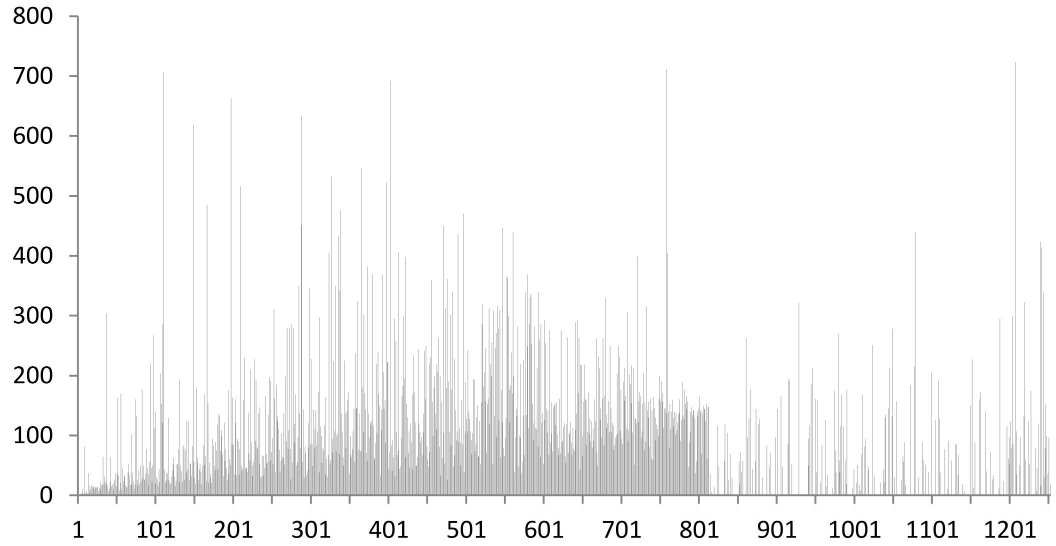


Figure 37. The difference between the highest and lowest ranked bank by assets during the past three years. The abscissa is the rank of the bank at the end of the observation period (February 2015). Banks, whose license was revoked, are located after position 814.

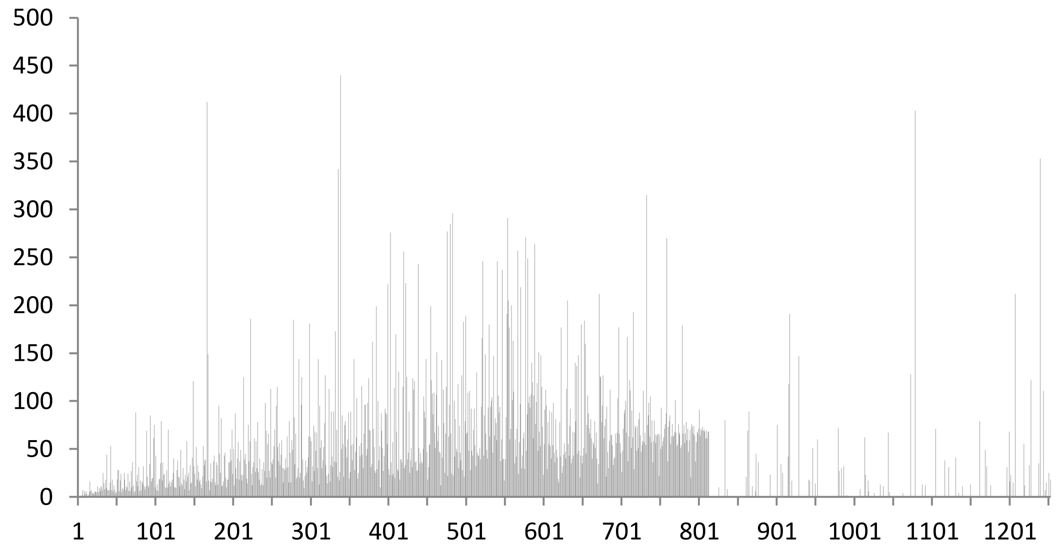


Figure 38. The difference between the highest and lowest ranked bank by assets during the last year. The abscissa is the rank of the bank at the end of the observation period (February 2015). Banks, whose license was revoked, are located after position 814.

Obviously, the question of direction of movement of a bank on the distribution occurs. To answer this question we calculate the average monthly change in the ranks of the assets of individual banks. The relevant information is shown in Figures 39 - 41 for the entire period of observation, for the last 3 years and over the last year. We found the banks which survived had on average a decreasing trend to their rang position. In turn, the banks whose license was revoked, shifted faster toward the bottom of the rang list.

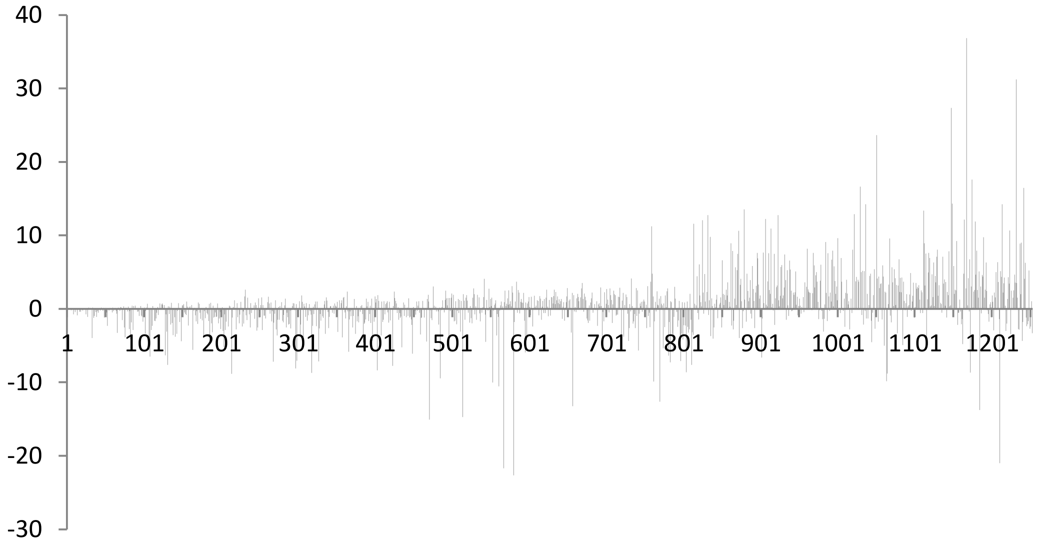


Figure 39. Average monthly change in rank of the bank's assets during the period of observation. The abscissa is the rank of the bank at the end of the observation period (February 2015). Banks, whose license was revoked, are located after position 814.

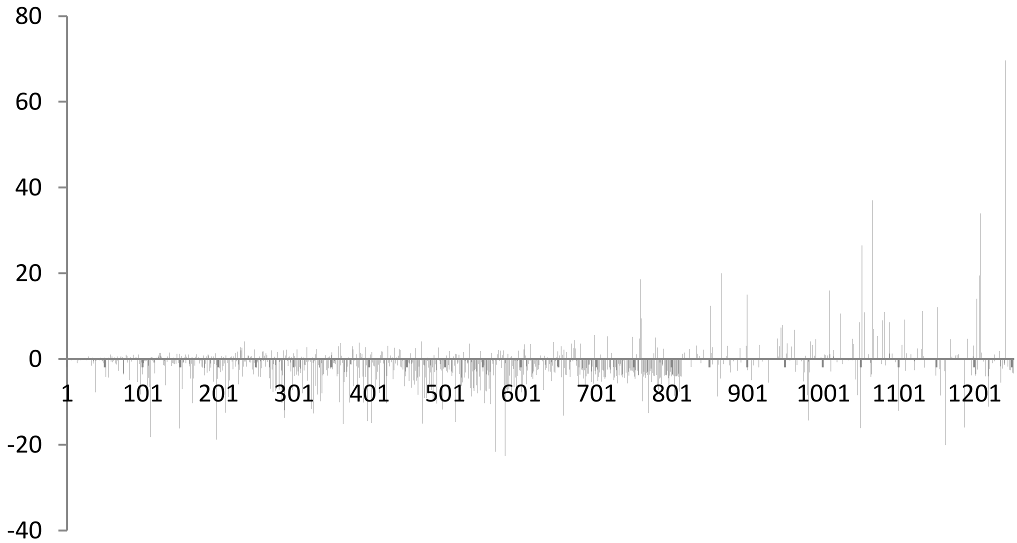


Figure 40. Average monthly change in rank of the bank's assets in the past three years. The abscissa is the rank of the bank at the end of the observation period (February 2015). Banks, whose license was revoked, are located after position 814.

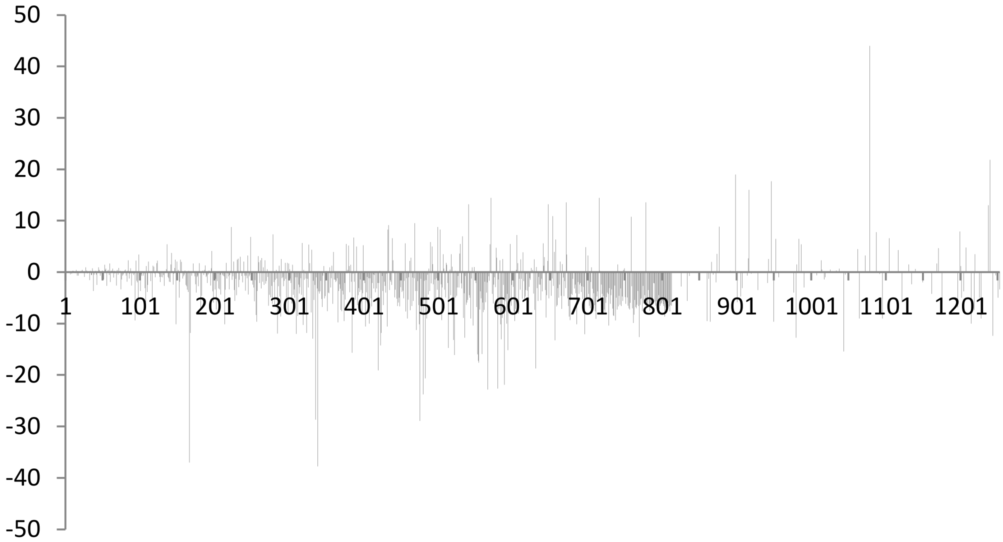


Figure 41. Average monthly change in rank of the bank's assets in the last year. The abscissa is the rank of the bank at the end of the observation period (February 2015). Banks, whose license was revoked, are located after position 814.

Thus, we can conclude that the overall stability of the distribution of bank assets is accompanied by the constant mixing of banks in a given distribution.

Dmitry I. Malakhov
National Research University Higher School of Economics. Laboratory for Macrostructure Modeling of Russian Economy, Research Assistant; Department of Applied Economics, Assistant Professor; E-mail: dmalakhov@hse.ru