***SERIES:*** *ECONOMICS*

*Dmitry I. Malakhov[[1]](#footnote-1), Nikolay P. Pilnik [[2]](#footnote-2), Igor G. Pospelov [[3]](#footnote-3)*

**Stability of distribution of relative sizes of banks as an argument to usage of concept of a representative agent[[4]](#footnote-4)**

We propose new theoretical model of 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 volume of deposits. This result provides an additional argument in favor of usage of representative agent concept in banking sphere modeling.

Empirical test 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 in this distribution, thus distribution of general population of banks is stable over time.

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

JEL classification: E10, L11, G21

**Introduction**

Microfounded macroeconomic models appeared as a result of Lucas critique (Lucas (1976)). Modern studies, which analyzes banking sphere in DSGE models, typically use 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 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 absence of conventional approach to modeling.

In this paper we show, how stability of distribution of relative sizes of banks can be used as an argument to usage of concept of representative agents in general equilibrium models. As proxy for relative banks’ sizes we use proportion of assets, because for banks, as financial firms, assets are good approximations of scale. Moreover, volume of bank’s assets is rather sensitive to requirements for banks’ products and financial policy (Jagtiani, Khanthavit (1996), James, Houston (1996)). In our paper we pay attention to specific activity of banks, such as deposits accumulation and interbank money transactions to give an answer to raised question and structure of 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 of 2007-2009 years revealed, that structure of the financial system is crucial for stability, forecasting of its development and characteristics analysis. So attentions from famous 4 «L», leverage, liquidity, losses, and linkages, now shifted to last «L», because risk measures for the first 3 «L» are rather investigated, but not for the last one. Modern studies, which analyze structure of financial industry (Acemoglu, Ozdaglar, Tahbaz-Salehi (2015), Billio, Getmansky, Pelizzon (2012), Iori, De Masi, Precup, Gabbi, Caldarelli (2008)) show, that number of linkages and their characteristics are very important for good risk resistance, but these models in fact don’t pay attention to reasons of developing such structures in banking system. Moreover, analysis of dynamics or evolution of industry is very important, because it reveals mechanisms of changing of industry structure and helps make more precise forecasts.

Plan of the paper is following. In the first section we discuss theoretical model of banking system. In the second section we provide empirical testing of our model using data from Russian banks. And finally we make conclusions.

**1.** **Model of interbank transactions**

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

We show how mechanisms of money distribution affect structure of banking industry. We don’t impose any restrictions to banks heterogeneity, but get that distribution of proportion of assets of banks is stable over time. If empirical tests of these results show, that it will correct, than it can be used as an additional argument to concept of representative agent. Because if banks are really heterogeneous, than empirical distribution of proportion 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 use below is the backward Kolmogorov equation for Markov processes. In our case, in fact, it is correctly recorded total probability formula. This approach is used in economic problems related to the description of the dynamics of systems consisting of small agents, for example, see discussion of dynamic of industry with monopolistic competition in Melitz (2003), Hopenhayn, H. (1992a), Radionov, Pospelov (2014). In this article, we use it to model dynamic of banking system in order to analyze evolution of distribution of overall set of banks, rather than behavior of individual banks. Due to this approach we are able to avoid assumptions about rational behavior of individual agents and existence of equilibrium, for example, as in Ericson, Pakes (1995).

**1.1. General description of the model**

New money appears in the economy through the two channels:

1. Loans from outside of banking system. Residents and nonresidents can put their money into national banks. National banks can give credits to residents and nonresidents or can use other financial instruments, such as debt emissions.
2. Credit emission. Bank can give credit to client with corresponding creation/changing of his/her current account.

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

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

Withdrawals from bank occur only when bank repays for its’ debts or 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.

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

Assume that all credits are repaid with frequency, which is proportional to duration of credits. All assets are identical to each other in the sense of duration (only moments of creation of bank’s assets are changing). Also we consider, that all emissions are equal in size and value of assets of banks depend only on the number of 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 developed banking system with great amount of highly competitive banks and identical clients during rather long period of time, then assumptions about durations and deposits sizes are relevant, because we can average all transactions.

Propose, that there are set of banks ****. Proportion of individual bank’s assets in the overall amount is , . We assume that set **** is stable over time.

Bank  induced initial emission , transaction is needed with probability. We can assume, that each emission induces chain of emissions (like in banking multiplier models), if amount of banks is great and  is small, than corresponding series converge. Moreover, this assumption does not decrease an explanatory power of model. So let’s assume that with probability  transaction  is needed. Average assets change after transaction will be

.  

**1.2. Stochastic process of assets change**

Let’s assume, that at moment  bank  has assets 

,    
 During period  one client of randomly chosen bank , independently on the others initiates the emission with probability  with size , which is much smaller, than . Note, that  is “real” demand and  is proxy for inflation and society welfare. Thus for simplicity we will further consider, that  is constant.

With probability  induced emissions are independently covered. Furthermore, we propose, that  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  can change during repayment time.

**1.3. Dynamic of generalized moments**

Now discuss averaged value of some function  over realization of stochastic process :

,   
where – mathematical expectation of  over . Calculate ,  using chain rule for mathematical expectation:

.   
 During the period :

1. with probability  assets are not changed
2. with probability  assets of bank  decrease by 
3. with probability  initial emission  occurs at bank  and with probability bank  continues this transaction. Here we assume, that probability of continuation of emission is proportional to amount of assets.

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





 (5)

.

Than we use (4):





. (6)

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



, (7)

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

We could find solution to this equation:

. (8)

Notice, that density function  can be represented as a quotient of two typically different factors. Numerator  depends only on proportions of assets, but does not depend on time or absolute value of assets. Denominator depends on absolute values of assets of individual banks are only in the expression . If this expression has constant value over time, then it makes sense only to pay attention to aggregate value of assets in the denominator. We will call proportion of assets as relative size of banks.

Since the function  is a function of density, of course, that the integration of all the values of  at any time  gives 1. After the change of variables in the integral, we can go to the  indexes of proportion and the total absolute amount of assets, as is easily seen, the possibility of integrating separately the numerator and denominator. Therefore, the function  can be regarded as up to a constant a function of density, depending upon only from a fraction of 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 unchanging monetary regime and absence of structural shocks (such as risk requirements or Central Bank regulation policy) distribution of relative sizes of banks is stable over time, so economy growth (which could increase number of banks clients and/or volume of their deposits) does not affect structure of relative sizes of banks. It means, that system has some sort of homeostasis property[[5]](#footnote-5), so usage of macroagents could lead to correct results without any loss of generality due to stability and homogeneity of banking system.

**2. Empirical testing**

**2.1. Model validation**

To validate the model we provide empirical tests. We decide to use financial statement of banks as source of information, because financial statement is really informative for bank analysis[[6]](#footnote-6). Moreover, we decide to use information from Russian banks, because Russian banking system is rather developed and competitive, especially for the last 10 years and the data about Russian banks is open and very detailed.

**2.1.1. Data**

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

Information about 101 turnover balance sheet is collected from the official website of the Central Bank of Russian Federation <http://www.cbr.ru/>. In our sample in average for period 2009-2015 we have approximately 99% of overall number of banks and about 99% of 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).

Subaccounts are rather minor, so they are noisy and are not very representative indicators of financial health of individual banks. We use aggregate variables, because they are very informative, are not so noisy and number of these variables is not very high. All values are given in thousands of rubles.

We decided to use following variables:

1. Total amount of assets.
2. Fixed date deposits of banks and other credit organizations, including overdraft (further we will use abbreviations for financial variables, so this variable is Db)
3. Fixed date deposits of non-residents (Df),
4. Fixed date deposits of natural persons – residents (Dh),
5. Fixed date deposits of nonfinancial organizations (Da),
6. Fixed date credits to commercial nonbank organizations-residents, including overdraft (La),
7. Fixed date credits to natural persons – residents (Lh),
8. Fixed date credits to foreign organizations (Lf).

All our variables are calculated by summing corresponding subaccounts of 101 turnover balance sheet. We choose these variables, because they are significant proportion of total amount of assets (liabilities). Final data are tables, where columns indicate time period and rows indicate bank id. We have separate tables for each financial variable. Period of observation begins at January 2004 and ends by Febrary 2015 (monthly data without any omissions). We have actual data for each time period.

We will use in our analysis relative size of individual banks in total amount of particular variable. So, for example, proportion of assets of Bank A is amount of assets of Bank A at the end of month i, divided by total amount of assets of all banks in the sample at the end of month i. We use proportions of assets instead of absolute assets because distribution of proportions is investigated in the first part of our work. Moreover, we need not to deflate them and they give relevant picture of banking system structure. As it was mentioned before, we would call proportion of assets as relative sizes of banks, due to the fact, that amount of assets is a good proxy for bank size.

Number of banks in Russia changed through time, also portion of banks which gave information to the Central Bank changed too, so we have different number of observations each month. Generally, number of banks didn’t vary greatly. Number of banks with nonzero values at the beginning of time period is approximately 700 and 1100 by the end. It is important to mention, that although we work only with banks, which provide information to the Central Bank.

We don’t drop any banks from our sample, so we estimate distributions including very big banks, such as Sberbank and VTB. Moreover, we don’t ignore very small banks, which form left tail of distribution. This section is mainly based on 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  for each month (vertical axes is value of corresponding parameter, horizontal axe is time (January 2004-Febrary 2015)).

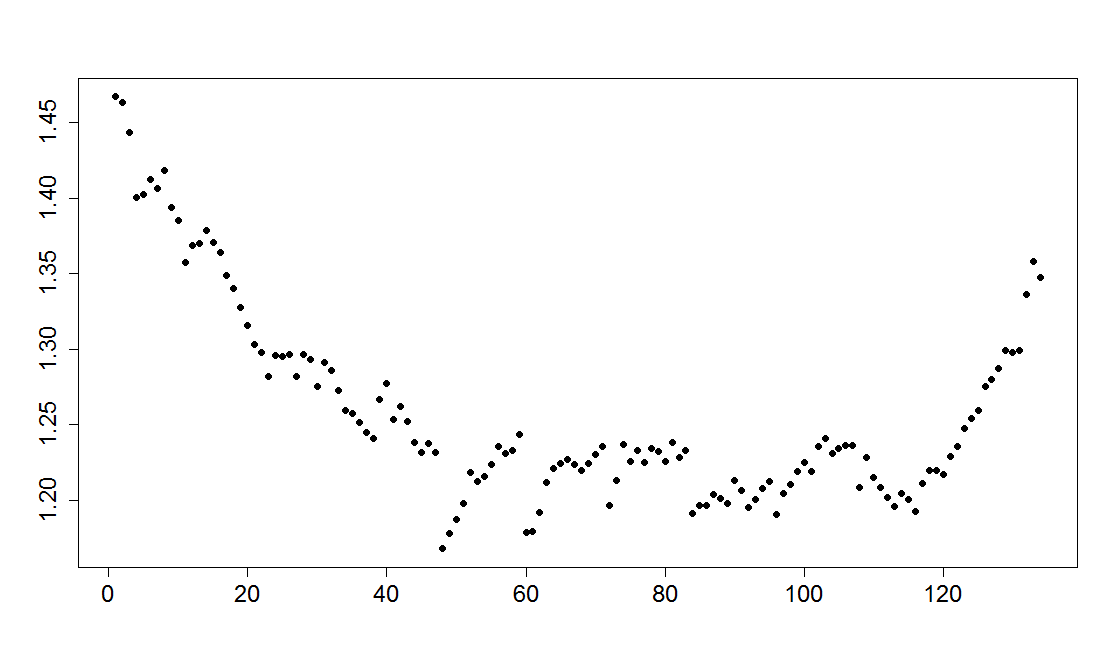


Figure 1. Values of the main factor of (22).

As we can see, for a fairly long period of time from March 2008 (approximately 50th point) to January 2014 (120th point) value of  was rather stable. Consequently, during this period, it can be expected that the distribution of relative sizes banks did not change as much as in the remaining periods. Shocks of could be possibly connected with sample changes, monetary policy regime switching and other banking system regulation aspects. It is quite interesting, that “stability” period include part of global financial crises, but it might be connected with the fact, that in USA this crisis began at the end of 2007 and when shocks affected Russian economy Russian banks and Central Bank already knew how to react and provided some preparation activities.

Also we investigate dynamic of . The chart clearly shows three periods: before 50th point, from 50th to 120th and after 120th. Thus figure 1 represents some realistic trends, not “statistical artifacts”.

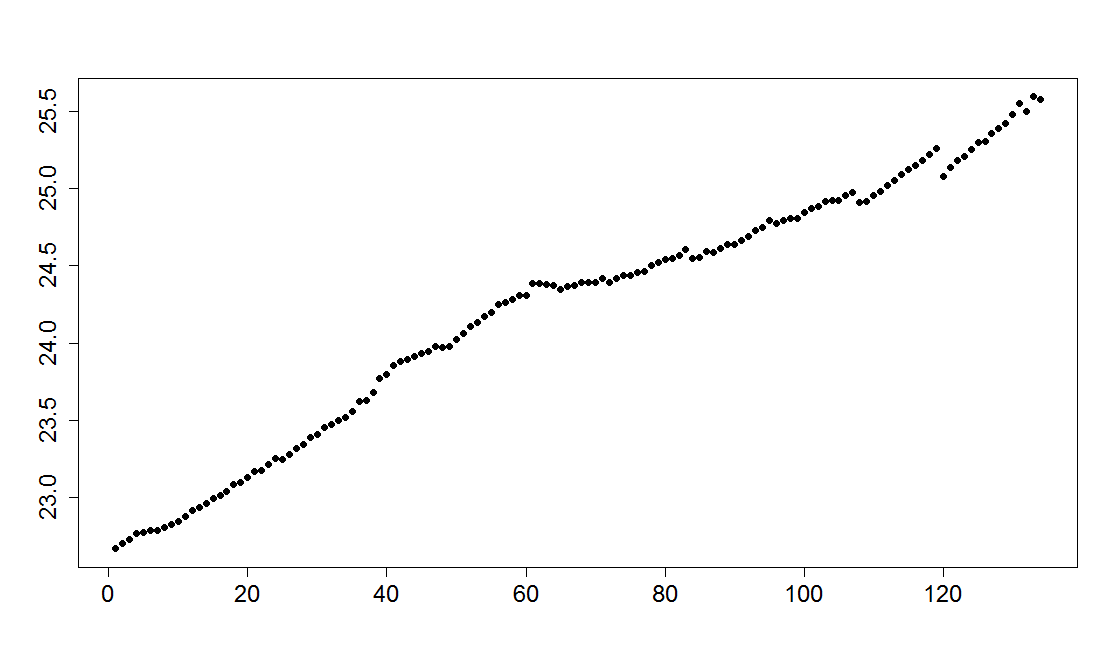


Figure 2. Logarithm of the total assets of the Russian banking system.

**2.2. Approaches to modeling distributions of firms’ sizes**

Theoretical model showed, that distribution of relative size of banks is stable over time, so these results is highly connected with industry evolution and development. Today there are many works connected with evolution of particular industries. Classical work is Gibrat (1931), in which the following hypothesis is formulated: firm size and its growth rate are independed. If this hypothesis is true, than firm growth rate is independent of its size, so big and small firms have approximately equal growth rates. Researchers tested this hypothesis for many different industries. Generally, it is difficult to say if this this Law correct for 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) provide empirical test of Gibrat Law for young firms. Authors postulate that there are three approaches to Gibrat Law testing. First approach is the most general: Gibrat Law is correct for all firms, independently of their bankruptcy fact during period of observation. Second approach: Gibrat Law is correct only for firms, which are functioning during period of observation. Third approach: Gibrat Law is correct only for firms which size is bigger than minimum efficient scale. Authors provide great literature survey (actual for 2003 year), they show that for some industries these Law is correct, for some - not.

If Gibrat Law is correct, than firms sizes 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 be also helpful.

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

In 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 concentration of counterparty exposures to a given institution in explaining its systemic importance. In this paper, which analyzes interbank sector, it is shown, that right tail of distribution can be approximated by Pareto distribution. For testing this hypothesis Cont, Moussa (2010) shows, that linear regression models are good approximations of the logarithmic data.

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

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

Selection of specific functional form of distribution helps understand properties of random variable, forecast more precisely its dynamics and calculate more accurately indexes of inequality (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 evolution of industry and connection of firm characteristics with their sizes. But authors don’t know such works, which discuss banking system.

**2.3. Preliminary analysis of data**

In this part we will focus only on relative sizes of Russian banks[[7]](#footnote-7). We calculate descriptive statistics for all time periods, but for better visualizing 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 difference between the mean smallest bank and the mean biggest bank[[8]](#footnote-8) is rather significant. So we have some clusters of big banks, such as Sberbank, VTB, and cluster of very small banks. It is important to mention, that mean of relative sizes is , n – number of banks. Moreover, we plot 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)).

|  |  |
| --- | --- |
|  |  |
| 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 can be easily seen, that since October 2009 (70th point) our sample is approximately equal to the amount of Russian banks. Inequality (measured as standard error) has nontrivial dynamic, but it has rising dynamics since 70th point. Skewness and kurtosis are from normal ones, so we expect, that distribution functions will be nontrivial.

**2.4. Distribution approximation**

In our research we use two families of distributions: 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 details description of distributions and motivation for their selection see Appendix 3.

We use R software for our analysis.[[9]](#footnote-9) Maximum likelihood and L-moments approach are used, because these methods give estimators with “good” properties and they don’t depend on chosen distaces (Asquith (2015), Hosking (2015), Yee, Wild (1996)).

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

|  |  |
| --- | --- |
|  |  |
| Figure 7. Relative sizes, Pareto distribution. | Figure 8. Rang distribution of relative sizes in both logarithmic axes. |

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

We estimate normal distribution for the log data (for motivation see Appendix 2). As we can see quality of approximation is moderate (but higher than for Pareto distribution), because the left tail is approximated (quantiles less than -12) not very good and there is noticeable bias of the middle part (quantiles is around -8), which is connected with calibration of the right tail. So we move to analysis of the rest distributions.

|  |  |
| --- | --- |
|  |  |
| Figure 9. Logarithm of relative sizes, normal distribution. |  |

Further we will discuss graphs for 4 variables: proportion of assets, proportion of credits to firms (La), proprotion of deposits of households (Dh), proportion of interbank deposits (Db) (see Appendix 3). We show graphs only for May 2012, because distribution is rather stable over time and patterns will be the same. We decided not to show results for generalized beta of the second kind, because its results is equal to Pareto IV ones and graphs become more comlicated. Wakeby distribution is turned into Pareto IV for our data, so we exlude this distribution too. Also we assume, that location parameter for Pareto IV type is 0. This assumption is recociled with data and does not lead to any limitations (Brazauskas (2003)).

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

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

1. 
2. , where n – number of observations in k-th month.

We show maximum, minimum, average and standard error of distances. 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 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 better visualization, we show graphs for these distributions for November 2004 and November 2014. November is one of months with small seasonality effect. Also it is interesting to investigate differences in distributions which occurred during 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 last 10 years. Both distributions has become more sloping, so in general banking system has become more homogenous. Recent crisis did not affect (in terms of relative sizes) banking system greatly, because there are no significant changes in the form of distribution. Quality of approximation is very high. Thus we can conclude, that functional form of distributions of relative sizes is stable over time.

We decided to provide Kolmogorov-Smirnov test for compering two nearest empirical distributions. Selection of empirical distribution as a subject of analysis is caused by fact, that empirical distributions is consistent estimates for true ones and it is difficult to conclude which theoretical distribution (Pareto IV or AEP) is better. Mean p-value for comparing two nearest distributions is 0.94, which means that hypothesis “both empirical distributions come from the same continuous theoretical distribution” can’t be rejected at 5% confidence level. If difference between distributions is more than 8 months, than for some pairs the null hypothesis can be rejected at 5% confidence level. Thus distribution of relative sizes of banks is rather stable over time.

For modeling other variables (different types of deposits and credits, as it was mentioned before) we use AEP and Pareto IV distributions, because they are the best approximation for these variable too (for results see Appendix 4, for space economy we provide results only for these distributions). It is difficult to make a choice which distribution is better. Different financial variables are approximated better by different distributions (generally, AEP distribution is better for most cases). Distances between empirical and theoretical distributions did not change during our time period (for all our financial variables), so our results are stable. We decide to select both, AEP and Pareto IV, as most appropriate distributions.

Also we analyzed graphs of density functions for relative sizes of banks, build with kernel functions and found out, that these function were rather typical, unimodal without any “statistical artifacts”.[[10]](#footnote-10) Thus this fact can be used as an argument to concept of representative agent.

Stability of distribution of relative sizes of particular banks is rather unrealistic result. So we guest, that not distribution of particular banks is stable, but distribution of banks of banking system. So we checked how individual banks could change their position on the sorted list. We got, that there were really a lot of changes in banks’ positions (for detailed results see Appendix 5). We could not find any patterns in banks’ movements. But in overall middle sized banks had in 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 in average had negative monthly rang trend, so banks tends to become relatively bigger, thus inequality could increase, this is highly connected with results Malakhov(2015). But despite of these changes overall distribution is rather stable. Maybe it is some sort of homeostasis property, thus if some bank lost clients/assets other banks accumulate them due to high competition. This fact also can be used an argument to representative agent concept, because if system is stable and rather homogeneous, macrolevel variables can be forecasted using only previous values of macrolevel variables, if frequency of data is rather high 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 our data. Also according to formal and graphical analysis functional form of data is table over time, so now it is important to define stability of estimates of parameters of Pareto IV type distribution and AEP distribution. Figures 15-17 shows dynamic of values of parameters of Pareto IV distribution (the first point is January 2004 and the last point is February 2015).

|  |  |
| --- | --- |
|  |  |
| 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 od parameter of shape of Pareto IV distribution |  |

After the 60th point values of estimates of parameters become stable (the 60th point is December 2008). We can see only a slight downward movement for estimates of shape and scale parameter and a slight upward trend for estimates of inequality parameter at the end of period. Also it is important to compare dynamics of estimates of parameters with dynamic of  (Figure 1). In the beginning of the paper we mentioned, that if is stable over time, than distributions are stable over time. Thus it can be easily seen, that periods, when is rather stable (for example, 50-120th point), estimates of parameters of Pareto IV is rather stable too. While during periods of instability (before 50th point or after 120th point), estimates of parameters are instable too.

Analysis of dynamic of estimates of parameters of AEP distribution (Figures 18-21) gives similar results. Location and first shape parameters are rather stable since 60th point and after 90th point have a slight downward trend possibly connected with recent crisis. But estimates of scale and second shape parameters are not stable since October 2009 and have very nontrivial dynamic for the last year. But this fact can be potentially explained by procedure of estimating of parameters of AEP distribution. Estimates of parameters of Pareto IV are get by method of L-moments, but estimates of AEP distribution are get by maximum likelihood estimates. Due to very nontrivial functional form maximization procedure for AEP is not very robust, thus some proportion of deviation of estimates 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 period of observation, but in fact distributions for 2004 and 2014 years are identical, so estimates changes are minor in absolute value (except scale parameter for Pareto IV). Also it is important to notice again, that our sample was changing even since October 2009, because number of banks in Russia was decreasing, so some part of dynamic of estimates can be explained not only by institutional changes of Russian banking industry, but also by sample changing (see Figure 3), but of course these processes are connected.

So in fact these slow changes in estimates of parameters of distributions can show, that this banking system is affected by gradualism property. 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 use tools, described in Melitz(2003), Hopenhayn (1992a) to provide an argument to usage of concept of representative agents for banking sector modeling.

We show, that for large-scale open banking system under assumptions, that monetary policy and regulation aspects distribution of relative sizes of individual banks is stable over time. Thus economic growth, which can influence number of banks clients and/or volume of deposits, does not affect distribution of relative sizes of banks, measured as proportion of overall assets.

Data from Russian banking system is used to validating theoretical model. We discuss among assets other key variables, such as deposits of households, credits 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, quality of approximation is stable over time, thus we could say, that functional form of distributions is stable too.

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

Estimates of parameters values of Pareto IV distributions are mainly stable over time, 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 rather nontrivial dynamic. This result can be explained by usage of different estimations procedures. Moreover, absolute changes in estimates of parameters values are not very significant (except parameter of scale). Question about strength of connection of estimates dynamic with sample size and industry shocks, such as monetary policy regime switching and banking regulation, is open.

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

In future research project it is important to investigate influence of changes of monetary policy and regulation requirements on distribution of relative sizes of banks and analyze in details dynamics of estimates of parameters. It will be also useful to develop forecasting technique for predicting evolution of banking system. Moreover, it is better to use data of banking system of developed country, such as USA, to test the theoretical model, because differences between banking sector 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*.

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:





.   
 We assume with probability equal to 1, that , but  is finite. Then:

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

   
 . (11)

We know, that for each smooth function :

,

, where  - Dirac function.

If ,  - constant parameters, and distribution of random variable  near  has smooth density , then:

 - density of joint distribution.

 is parameter, so:

, (12)  
****,

Using expression given above:



.

Simplifying expression above, we get:





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



. (14)  
 Simplify the derivatives in (14) and divide both sides of equation by:







Than



**,



. (15)  
 If we integrate second and third summands over the whole space and in each element integrate over , then this integral will be equal to 0, because  turns into 0 at the edges.

Because number of banks , so we can write down:



 . (16)

We can make in (16) following substitution:

 (17)  
 Then we substitute (17) into (16):



.

We make the following change of variables:

   
 After simplification we get:

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

.

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

,

.

We can write the preceding relation in the following form:

. (18)  
 General solution of homogeneous equation:

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

, .   
 We can find particular solution as series:

,

,

.

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

.

Homogeneous equation  ⇔  has solution . So we can use variation of parameters:

,

 ⇒ .

We need only particular solution, so assume, that , then  and so partial solution will be:

.   
 Then general solution of equation is following:

.

Now return to original variables :

.

We notice again, that  does not depend on time, so, substituting 

we get:

,

And thus we get (8):

.

**Appendix 2. Used distributions**

**Pareto and Lognormal distributions**

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

Random variable x has power distribution, if probability density function is, where  - parameter. Often Pareto distribution is used as the basic power distribution:



Much attention is paid to parameter value . If , than big values have significant effect on the average value. If , than observation with small values have significant effect on the average. Pareto distribution is rather popular in modeling distribution of firm sizes (Axtell (2001), Crosato, Ganugi (2007)).

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

Today 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 big. 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 notice, that further distributions have been divided into two families rather roughly, sometimes there are 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:

,

where ,- location parameter, –scale parameter,  - shape parameter.

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

,

- uniform random variable with support [0,1], -location parameter, -scale parameters,-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 of generalization of Pareto distribution is **Pareto IV type**. Pareto IV is a generalization of many different Pareto type distributions and has following cumulative distribution function:

,

where  location parameter, scale parameter,  shape (inequality) parameter, 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 generalization of many different distributions, including Pareto IV type distribution. Probability density function for Generalized Beta of the second kind has following functional form:

,

where - shape parameters and -scale parameter.

Beta prime distribution is very often used in income distribution analyses and financial modelling, but there are many problems with 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: ,where



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

**Skew normal distribution** has following probability density function:

,

where  - parameters of location, scale and shape, respectively. This distribution has 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:

 ,

where  incomplete gamma function, -location parameter, - scale parameter, ,- 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 notice that asymmetric exponential power distribution has maximum entropy in the very wide class of distributions with support  (Zhu, Zinde-Walsh (2009)). Moreover, tails of this distribution can potentially have different fatness and they are much fatter, than normal ones. In paper Buldyrev, Growiec, Pammolli, Riccaboni, Stanley (2007) asymmetric exponential power distribution is used for modelling firms sizes.

**Generalized lambda distribution (Gen. lambda distribution)** is also used for modelling data. Quantile function is following:

,

- uniform distribution with support [0,1], -location parameter, -scale parameter,-shape parameters.

Generalized lambda distribution is asymmetric distribution with power law tails, so it is often used for financial modelling. Generalized lambda distribution was developed as an approximation of many standard distributions, because it was originally used in Monte Carlo modelling, so flexibility of this distribution should be very high. In Beena, Kumara (2010) generalized lambda distribution is used for modelling 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. Proportion of credits to firms (La), family of Pareto distributions (Pareto, Gen. Pareto) |
|  |  |
| Figure 27. Proportion of credits to firms (La), family of Pareto distributions (Pareto IV) | Figure 28. Logarithm of proportion of credits to firms (La), family of normal-related distributions (Normal, Gen. Normal) |
|  |  |
| Figure 29. Logarithm of proportion of credits to firms (La), family of normal distributions (Skew Normal, Gen. Lamda) | Figure 30. Logarithm of proportion of credits to firms (La), family of normal distributions (AEP) |
|  |  |
| Figure 31. Proportion of interbank deposits (Db), family of Pareto distributions (Pareto, Gen. Pareto) | Figure 32. Proportion of interbank deposits (Db), family of Pareto distributions (Pareto IV) |
|  |  |
| Figure 33. Logarithm of proportion of interbank deposits (Db), family of normal-related distributions (Normal, Gen. Normal) | Figure 34. Logarithm of proportion of interbank deposits (Db), family of normal-related distributions (Skew Normal, Gen. Lambda) |
|  |  |
| Figure 35. Logarithm of proportion 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 IV | 0.0172 | 0.0054 | 0.0102 | 0.0027 |
| 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.[[11]](#footnote-11) 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 will place 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 shifted more than 400 positions. This means that individual banks are constantly changing their position along quite 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, which ranks over the year changed more than by 200 points. We could not find any patterns in banks’ movements; moreover, we can’t say that some 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 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 on Figures 39 - 41 for the entire period of observation, the last 3 years, over the last year. We could find, that survived banks had in average tendency to rang position decreasing. In turn, the banks, whose license was revoked, faster shifting toward the bottom of 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](mailto:dmalakhov@hse.ru), Tel.: +79032206238

**Any opinions or claims contained in this Working Paper do not necessarily reflect the views of HSE.**

**© Malakhov, Pilnik, Pospelov, 2015**

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 [↑](#footnote-ref-1)
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 [↑](#footnote-ref-2)
3. Federal Research Center “Informatics and Control” of Russian Academy of Science, professor, corresponding member, E-mail: [pospeli@yandex.ru](mailto:pospeli@yandex.ru) [↑](#footnote-ref-3)
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. [↑](#footnote-ref-4)
5. Homeostasis - selfregulation, it means that open system tends to return to equilibrium and optimal structure, despite of any external effect. [↑](#footnote-ref-5)
6. It is not correct for real sector firms (physical indicators (output, etc.) for firms sometimes are very useful, but rather subjective). [↑](#footnote-ref-6)
7. Due to space restrictions. [↑](#footnote-ref-7)
8. Values are stable over time. [↑](#footnote-ref-8)
9. Computer code and dataset could send for request. [↑](#footnote-ref-9)
10. Corresponding graphs can be send for a request. [↑](#footnote-ref-10)
11. Obviously, it is equivalent to analyze assets’ rang of banks or rang of relative sizes of banks. [↑](#footnote-ref-11)