



A Note on Building a Database on Russian Banks: Fieldwork Against the Odds

KOEN SCHOORS

Abstract

This note describes how I constructed a reliable database on Russian banks in early transition. I describe in detail data sourcing, data verification and the representativeness of the resulting database. The note makes available for research a database that captures the genesis of a commercial banking system and is therefore very interesting. Moreover, the note suggests that the availability of data in Russia might in practice be better than generally perceived. Often the data do exist, but it takes a lot of time and effort to find them and turn them into a user-friendly format. Mastering the Russian language and a fair portion of perseverance are a great help in this respect.

‘Nel’zya, no esli ochen’ khochetsya to mozno’
(It is impossible, but if you really want it then it is possible)

Prologue

In 1995 I was in Moscow, desperately looking for data on Russian banks. I was supposed to write a doctoral dissertation about the emergence of Russian commercial banks in 1992–94, but the data did not seem to be available at all. Bank balances were difficult to find and detailed profit and loss accounts were as rare as white elephants. According to many colleagues, I would fail to put my hands on this information. In fact, I should not even hope to find it, since it did not exist in the first place. Indeed so it seemed. But then again, in Russia nothing is quite what it seems to be. This note reports my personal quest for Russian bank data. The process of data gathering in itself revealed a lot of information about Russian banks.

I found out that many data do exist. As a matter of fact, for many banks there are better data than anyone could have expected. On the other hand, the data are hard to locate and often in a state of disorder. I first describe how I proceeded to find the data. Next I discuss how I used the raw data to construct bank balances and profit and loss accounts structured and clean enough to serve research purposes. Then I evaluate the representativeness of the resulting database and offer some concluding remarks.

Dr Koen Schoors, Faculty of Economics and Management, Ghent University, Hoveniersberg 24, B-9000 Ghent, Belgium, LICOS, RECEP.

Sourcing

In May 1995, as a first stage, I started to interview a number of Russian bankers. In the process I learned that there existed a directory of Russian commercial banks known as ‘the white bible’. The provider of this directory was Intelbridge, a medium-sized firm specialising in financial information. One of its basic activities is to collect information about commercial banks on the basis of voluntary cooperation on their part. I purchased the directory both in printed form and on diskettes in May 1995. The Intelbridge directory contains the answers of about 1500 commercial banks to 35 questions (the latter are listed in full in Appendix 1).

This directory was useful, since it contained the names of the bank directors and their phone numbers, fax numbers and mailing addresses. Between September and November 1995 I faxed requests for information across Russia to more than 200 bank directors, randomly chosen from the Intelbridge directory. In these faxes, directed to the bank director and written in Russian, I explained my academic mission, asked for the bank’s balance and profit and loss account of 1994 and promised not to disclose any confidential data. Reception of the faxes was checked and confirmed by personal phone calls—a thrilling experience by itself. Nevertheless, I received only three answers, namely from Inkombank, International Financial Company and Promstroibank. These three banks provided relevant information, and I was able to interview senior officers of Inkombank and Promstroibank. Even so, a response rate of 1.5% came as a severe disappointment. This dismally thin feedback probably reflected the closed character of Russian banking and the distrust pervading social relations in Russia.

Meanwhile, I kept on searching for other sources that might help to build my database. Several business information companies were contacted, but none could provide me with what I needed. Balance sheets were generally not a problem, but income statements and profit and loss accounts were usually unavailable. At one point, however, I read an article in a Russian business journal (*Kommersant*) suggesting that the journalist knew more about Russian banks than I did.¹ I contacted the journalist, who was willing to reveal his source. The ‘source’ referred me to another source, which transferred me to yet another source, and so on. I soon felt trapped in a Kafkaesque procedure, and as I went from one source to the next I started to fear that I would end up back with the journalist I began with. At some point, however, I was lucky enough to be transferred to X (this letter does not exist in the Russian alphabet and is chosen for that very reason). X listened to my story and agreed to a meeting in person in the Irish Pub on the old Arbat. This time it sounded as if I was on to something real. I asked X how I would recognise him and he said that he would wear a black coat and hold a newspaper under his arm. What in a different context might have come across as a worn-out cliché was under the circumstances a bit scary, but my desperate need for bank data dominated my fear and I went to see X in the Irish pub. There he stood, duly wearing a black coat and holding a newspaper. After a short conversation, he invited me to follow him to a small office (one small room in the middle of a federal ministry). I asked him whether he was working in that ministry. He answered that he used to work for the state, but he had lately gone private. When I insisted and asked him what ministry he worked for—I thought the Ministry of Finance or something like that—he answered very calmly: the KGB. Then he added my address and phone number in Russia, which I had not yet told him. My heartbeat stopped for some seconds, but recovered fast and reached a peak when he showed me his exciting database. It

contained information on more than 2500 banks, i.e. on about all the banks that existed at that time. For all of them, the database contained the name, date of registration with the CBR, address and phone/fax number. For most banks, the database included extremely detailed information on assets and liabilities, although not always on consistent dates. The information on profit and loss accounts was, however, less complete. Unfortunately, the data were not arranged in a user-friendly format, and hence of little value to non-specialists. Actually they were just an enormous pile of bookkeeping accounts, stating the number of the account and the corresponding amount of rubles. X, not being a trained economist, bookkeeper or banker, could not sort out the data by himself and probably did not realise the value of his pile of numbers to me. This certainly facilitated my 'purchase'. After some negotiations, I could purchase the data very cheaply, on the promise that I would try to establish contacts with interested buyers in the West, which I did. From the database I selected all the banks for which it supplied the balance sheet and annual profit and loss account at the end of December 1994 (technically 1 January 1995). This yielded data on 154 banks. I refer to this source as source A.

On the occasion of my second visit to Intelbridge, in September 1995, the company was working on a new directory with the balance sheets and profit and loss accounts of commercial banks. I managed to buy the data prior to publication, subject to a secrecy agreement on my part not to disclose any bank-specific information. These accounts were on paper, in rough form, in several different formats and for different dates. This was due to the fact that their data collection was based on voluntary cooperation of the banks concerned and aimed at assembling the most recent data rather than comparable data. This source, containing initially 261 banks, is referred to as source B.

My own on-site visits were the last and most revealing source. After checking sources A and B, I realised that some important banks were missing. I really wanted the information for these banks in order to increase the representativeness of my database. My earlier experience, however, had taught me that the usual questionnaire approach was unlikely to deliver much. Therefore, I devised a new action plan: I would personally go to visit a number of Moscow-based bank headquarters or Moscow branches of regional banks in order to try to obtain their annual report. To this effect, I invented the following story: I would introduce myself as the young representative of a Western trade firm wanting to open a dollar account in a Russian bank. I made up a name for this fictional firm (Improd) and to boost my credibility I had a business card printed, with my name and title, the name of the firm, and the coordinates of the firm (address, fax number, phone number and so on).

Armed with only this fable and a fair dose of inborn perseverance, I attempted to enter all the banks on my hit-list, coming up with more stories along the way to obtain their accounts. The first barrier was physically entering the bank. Access to Russian banks was (and still is) impeded by severe entry barriers, ranging from a closed gate to armed guards and metal detectors. My efforts to enter met with uneven success. In some cases, I did not even manage to pass the first guard ('We don't work with clients'). In others, however, I got so far as to be invited to have a drink with the director.

In many a bank my intrusion created a stir, as I was the first foreign visitor wishing to open a dollar account. Then I had to find the person who was responsible for opening dollar accounts. Usually I ended up sitting somewhere in the middle of the building, talking with a nice lady (indeed the large majority of bank officers turned out to be female). In most of the smaller banks I was some kind of a curiosity,

with half of the employees coming to have a look at me. The bank officer would ask me all kinds of questions about Improd and its line of business, and would give me a long list of administrative requirements that had to be fulfilled before I could open an account. After having talked to these persons for about half an hour, I would somewhat casually mention that ‘As there had been a severe crisis in the Russian interbank market in August 1995, my firm needed some kind of guarantee that the bank of our choice was secure. And yes, I would also like to see the annual report for 1994’. This trick worked fine in most cases. The main problem was that in some banks the annual report covering 1994 had not yet been completed (in October–November 1995!). These site visits were very well organised. As a matter of fact, I indicated all targeted banks on a map and organised the visits to minimise travelling time between banks. My determination to do the visits as fast as possible was not only motivated by time and budget constraints but most of all by the fear that bankers would somehow discover the hoax and that the rumour would spread that a young Belgian was hunting for bank accounts.

I thus succeeded in collecting information on the accounts of 30 banks, 20 of which were Moscow-based top-100 banks, absent from both source A and source B. The trustworthiness of the data collected through this channel was checked by comparing the balance sheet totals with the totals provided by source A. All accounts matched perfectly, except for one, which was accordingly dropped. This test was also reassuring with respect to the quality of source A. I refer to this source as source C.

Cleaning the Data

After the data collection, I walked the extra mile to check and clean my data. This was a time-consuming and particularly painful process, since many of the observations I had worked so hard to secure were lost.

I performed several tests on the correctness and consistency of each account. First I tested a number of accounting identities:

$$\begin{aligned} \sum \text{items of a category} \times &= \text{subtotal} \\ \sum \text{subtotals} &= \text{total} \\ \sum \text{liabilities} &= \sum \text{assets} \\ \sum \text{revenues} - \sum \text{costs} &= \text{profit} \end{aligned}$$

Then I tested whether some accounting entries were different from zero:

$$\begin{aligned} \text{Statutory capital} &> 0 \\ \text{Reserves at the CBR} &> 0 \end{aligned}$$

If an account fails one of these tests, something is badly wrong with the account concerned. Such accounts were left out of the sample.

Next, I tested for whether the banks concerned were operational. The banks should be functioning. I tested this by checking some structural ratios:

$$\begin{aligned} \text{Equity/total assets} &\leq 0.6 \\ \text{Reserves at the CBR/total assets} &\leq 0.6 \\ \text{Fixed assets/total assets} &\leq 0.6 \\ \text{Total deposits/total assets} &> 0 \\ \text{Total credits/total assets} &> 0 \end{aligned}$$

Banks that do not comply with these criteria are not operational or at least are not

Table 1. Description of the samples according to data sources

Sample	Source A	Source B	Source C	Total
D (1994)	126	–	–	126
G (1994)	126	84	20	230

operating as banks. These banks are excluded from the samples. In this way I avoided empty shell banks and facade banks that hide operations other than banking.

Lastly, I mentioned that some accounts looked so strange that one could doubt their trustworthiness and accuracy. Therefore I performed the following checks:

Other assets/total assets	≤ 0.6
Other revenues/total revenues	≤ 0.6
Other costs/total costs	≤ 0.6

Banks that do not comply with these criteria can be expected to cheat on their reporting obligations or are not functioning as genuine banks and are excluded.

After performing these tests and dropping overlapping accounts, I kept 126 accounts from source A, 84 accounts from source B and 20 accounts from source C (see Table 1).

Conversion of the Data into Comparable Formats: Sample D and Sample G

The accounts must be in the same format or complementary formats that can be translated into a common format. Unfortunately, our different sources supplied accounts in different formats. Source A delivered very detailed information. The information was not really in a format but simply reported all the accounts and sub-accounts of bookkeeping. I aggregated the information to an interpretable balance sheet and profit and loss account according to the rules established by the August 1993 CBR instruction on the establishment of a common financial accounting system for commercial banks.²

In December 1994 the CBR issued a letter laying out a common format for commercial banks to fulfil their publication obligations.³ This format for obligatory publication provides a balance sheet and a profit and loss account that are far more concise than the internal accounting format of source A. Unfortunately, sources B and C were in this more condensed format. I could translate source A into the format of source B and C, but the opposite transformation was not feasible, because of the lack of detail in sources B and C. The transformation of source A to the condensed format of B and C implied a severe loss of information. Therefore I thought it useful to construct two samples. Sample D (D for detailed) is in the most detailed format and only contains data from source A. Sample G (G for general) contains accounts from sources A, B and C and is in the more condensed format, instructed as the official publication format by the CBR in December 1994. The sample has therefore a lower information value but its size is larger and its representativeness is superior. Table 2 gives an overview. The transformation from source A data to the Sample G format was accomplished according to the method proposed by Androsov (1995).

Table 2. Representativeness of the samples (1 January 1995)

	Number of banks		Total assets (billion rubles)	
	Genuinely operating banks	Samples as % of the population	All registered banks	Samples as % of the population
Bank population	± 1000		322445	
Sample G (1994)	230	23.0%	94695	29.4%
Sample D (1994)	126	12.6%	35102	10.9%

Representativeness and Structure of the Samples

It is generally accepted that, out of the more than 2500 banks that were officially registered by the end of 1994, only around 1000 were genuinely operating as banks. Since I selected only banks that can be considered as genuine banks, Sample D (1994) represents about 12.6% of the number of active banks, while Sample G (1994) represents about 23% of the active bank population. If I compare total assets in our sample with total assets in the population (with the exclusion of Sberbank) I observe that Sample D (1994) represents 10.9% of total bank assets and Sample G (1994) represents 29.4% of total bank assets. Again, these are strong underestimations of the underlying representativeness, since I only selected genuine banks for our samples, while the population contains the assets of all registered banks. Table 2 gives an overview of these checks on the representativeness of the samples.

To check the representativeness of the samples in more detail, banks were classified according to three criteria. (1) Is the bank Moscow-based or not? (2) Is the bank local, or rather a regional or even national player? (3) Is the bank a successor of a (part of a) former state bank or not? The criteria were defined as follows:

Moscow-based banks: These are banks with the official address of their headquarters in Moscow according to the register of the CBR. This category is important because Moscow has developed into the financial capital of the country and therefore has special characteristics.

National banks: These are banks with branches in at least three Russian regions, other than the Central Moscow region. Their scale and scope of operation might give them some competitive advantage over the others.

Regional banks: These are banks with at least five branches in a particular region. They are large in their region, but often not important outside it.

Former state banks: These banks are the successors of a branch, a local department, a regional department or a sectoral department of one of the specialised state banks. They were in large part founded in the process of decentralised spontaneous privatisation of 1990–92. For convenience I will refer to these banks as state banks or former state banks. The predicate ‘state’ refers to their history (and to certain common characteristics that follow from it) and not to current ownership.

The Intelbridge directory was the main data source for the classification of banks. I use three dummy variables to control the bank’s type, namely MOSCOW, STATE and NATREG. It is interesting to know the structure of the samples with respect to bank type. The three criteria are not mutually exclusive. The combination of three dummy variables allows eight different combinations, ranging from private, small,

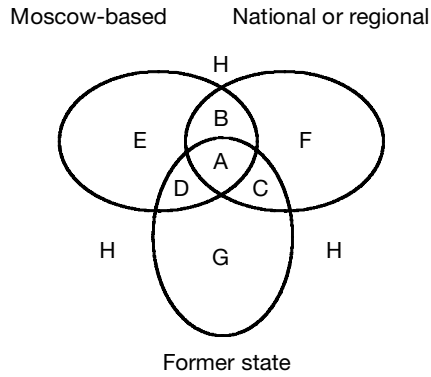


Figure 1. The structure of the samples.

Table 3. Structure of Sample D and Sample G according to the three criteria

	Sample D 1994	Sample G 1994
A	3	7
B	6	20
C	21	27
D	3	4
E	26	93
F	9	11
G	23	25
H	35	43
Former state banks (A + C + D + G)	50	63
Moscow-based banks (A + B + D + E)	38	124
National/regional banks (A + B + C + F)	39	65
Private small local banks (H)	35	43
Total	126	230

local banks (all dummies are 0) to Moscow-based large state banks (all dummies are 1). Figure 1 gives an overview of the eight different classes of banks and Table 3 gives the structure of our samples according to these eight classes. Comparing the sample distribution with the population distribution is difficult because there is no detailed information on the population distribution. The CBR (1994) notes that at the beginning of 1994, 609 of the 2041 registered banks were actually successors of the former SB, which amounts to 29.8%.⁴ To the best of my knowledge this is the last time the CBR disclosed data on the history of banks. Since many of the 2041 banks were not operational, the actual presence of state banks may even be larger. This shows that the strong presence of former state banks in our sample is not necessarily an over-representation. Both big banks and small banks are present in our study. Small local banks may seem to be under-represented if their number is considered, but given their small impact on the banking system as a whole, this is justified. Lastly, the share of Moscow banks in our sample is similar to the population characteristics (about one-third).

Concluding Remarks

At first sight, Russian data often seem to be absent. However, if one invests the needed time, effort and money, and if one learns to think like a Russian, one often finds that the availability of data is not so bad after all. Activism pays off in Russia. Mastering the Russian language also greatly helps. After carefully checking the reliability of Russian data, I was able to construct a decent database of Russian banks in 1994. This database captures the genesis of a commercial banking system and is therefore very interesting. I will be happy to share these data with other researchers. Also I hope that this note on data problems in connection with empirical work on Russia holds some interesting lessons for other researchers.

Notes

1. 'Sh'yutsya po odnoi vykroike. Rasshivayutsya po raznym', *Kommersant*, 22 August 1995.
2. Instruction of the CBR, No. 17, 24 August 1993, with addenda No. 1 and No. 2.
3. Letter of the CBR, No. 132, 22 December 1994.
4. The CBR (1994) notes that at the beginning of 1994, 609 of the 2041 registered banks were actually successors of the former SB, which amounts to 29.8%. Of these 609 banks, 42.7% were successors of Agroprombank, 28.2% were successors of Promstroibank and 20.2% were successors of Zhilsotsbank (p. 17).

References

Androsov, A.M., *Finansovaya otchetnost' banka* (Moscow, Menatep-inform, 1995), p. 459.
 CBR, *Tekushchie tendentsii v denezhno-kreditnoi sfere*, 1994, 4.

Appendix 1. List of questions in the Intelbridge directory

1. Full name of the bank
2. Abbreviated or reduced name of the bank
3. Registration date and number
4. Year of establishment
5. Legal address and/or mailing address of the headquarters
6. Phones
7. Fax
8. Telex
9. Teletype
10. E-mail
11. SWIFT code
12. Reuter Dealing code
13. Former type of ownership
14. Initial owners and founder of the bank
15. Current type of ownership
16. Principal owners and shareholders
17. Former belonging to the structures of Promstroibank or Zhilsotsbank or other banks of the USSR
18. Number of branches
19. Co-ordinates and managers of the branches

20. Major correspondents in Russia
21. Membership in interbank or other consortia, associations, alliances etc.
22. Basic services offered
23. Main areas and objects of investment
24. Financial description on 1 January 1994
25. Structure of credit investments
26. Structure of the authorised capital
27. Structure of income
28. Issue (1) of shares (2) of deposit and savings certificates and bills of exchange
29. Founding activities
30. International business activities
31. Number of permanent staff
32. Supervisory Council
33. Board of Directors
34. Charity activity
35. Additional data and considerations pertaining to future activities