Interest Rate and Size of Credit – A Non-Parametric Exploratory Analysis

Ashok Kumar Nag [†]
Bodhisattva Sen[‡]
Dibyendu Bhaumik[⋆]

January 17, 2007

Abstract

In this paper we examine empirically how interest rate is related to size of credit limit made available by commercial banks of India and how the underlying relationship varies over well-defined geographical areas of the country. The methodology that we employ for this purpose is known as Fractile Graphical Analysis, which is one of the earliest non-parametric regression technique developed especially to compare two regression functions for two bivariate populations (X, Y) where the co-variate (X) for the two populations are not necessarily on comparable scales. Our empirical study leads to the conclusion that credit market imperfection exists in the Indian economy.

Keywords and Phrases: Interest Rate, Credit, Fractile Graphical Analysis, Fractile Graphs.

AMS Subject Classification: 62G08

[†] Riskraft Consulting Limited, Mumbai

[‡] Ph. D student, University of Michigan, Ann Arbor

[⋆] (Corresponding Author) Reserve Bank of India, Mumbai; **Mailing Address**: B1/207 Kailash Apartment, Reserve Bank Officers' Quarters, Quarry Road, Malad (East), Mumbai - 400097, India; **EMail**: dbhaumik@rbi.org.in; **Telephone#**: +91-22-22159330(Work); +91-22-28744322(Home);

The views expressed in this article are those of the authors and do not necessarily represent those of the organisations to which they belong.

1 Introduction

The monetary transmission mechanism is a complex phenomenon and the channels through which monetary policy impacts the real economic activities are many. Some of the most important and identified channels are - interest rate channel, wealth channel, exchange rate channel and credit channel. An alternative broad classification of these various transmission mechanisms is twofold, namely *money view* and *credit view*. While under the *money view* of the transmission mechanism the process is traced through impact of monetary policy on macro-economic aggregates at the level of whole economy or its sub-sectors, the *credit view* allows us to examine the impact at more disaggregated micro-level. In the presence of credit market imperfection and non-substitutability between bank credit and other sources of external finance of investing firms, cross-sectional studies using micro-data may provide more insight about the distributional aspect of the policy induced changes in the monetary conditions under which the real economy operates. While we have a number of studies using firm level data on financing pattern to assess the differential impact of monetary policy induced interest rate changes on firms with different characteristics, there hardly exists any study using disaggregated data of a bank's credit portfolio. Kashyap *et al.* (1997) have used a large panel data set that includes quarterly bank-level data for all insured commercial banks of USA for the period 1976–1993 to examine the lending behaviour of these banks in response to policy induced interest rate changes. The focus of their exercise is on a bank's total credit portfolio and not on individual credit accounts. In this respect, the Indian banking data is uniquely placed. The

Reserve Bank of India collects a very rich set of credit account wise information for all scheduled commercial banks of India on a yearly basis. These data allows us to study the distributional aspect of interest rate for a set of borrowers with given characteristics. In this paper we propose to examine how interest rate is related to size of credit limit made available by a bank and how the underlying relationship varies over well-defined geographical areas. The novelty of our exercise is that we consider the relationship in its entirety in a non-parametric set up. Any non-linearity in the underlying relationship becomes also visible through this exploratory analysis. The methodology that we employ for this purpose is known as *Fractile Graphical Analysis* (FGA), introduced originally by Prof. P. C. Mahalanobis, way back in 1960s (Mahalanobis, 1960). This non-parametric methodology allows us a good visual description of the underlying relationships, although no precise estimate of strength of relationship is obtained. Thus our exercise is largely in the nature of exploratory data analysis. The remaining part of the paper is divided into four sections. In the first section, we present some relevant details about our data set and the empirical exercise undertaken. The second section on methodology gives the details of our empirical strategy with a brief exposition of the FGA. The third section presents the results that we have obtained. The last and concluding section presents comparison of estimated smooth fractile graphs.

2 Basic Statistical Return (BSR) Database

The Basic Statistical Returns (BSR) System introduced by Reserve Bank of India (RBI) way back in 1972, contains detailed account wise information on credit outstanding as on 31st March of every year. The BSR system covers all commercial bank branches in India. Account wise details for all credit accounts having credit commitments more than rupees two hundred thousands (called credit limits) by the bank are being collected through BSR system. For accounts having credit limit below this cut-off limit, only consolidated data are obtained. The enormity of scope and coverage of this database can be understood from the fact that currently data are collected from more than 65 thousand commercial bank branches with account wise details available for around 3 million credit accounts. For individual credit accounts, data are obtained with regard to following characteristics - account type (i.e. cash credit, overdraft, term loan, bill discounted etc.), organization of the borrower (i.e. individual, public sector undertaking, non-government public and private limited companies etc.), industrial classification of the borrower (i.e. manufacturing, agriculture etc.) and type of asset (i.e. standard, sub-standard, doubtful, lost). Interest rates are subject to regulatory ceiling for certain type of credits like export credit and for certain type of industries and nature of organization, and impaired credit accounts may have characteristics altogether different from the standard credit accounts. Similarly, the nature of facilities extended by the banks also determine the interest rate charged on them. Thus credit for working capital purposes is given on a revolving credit basis and may attract different rate of interest as compared to the rate charged for fixed duration loan called term loans in India. Moreover, management practices and labour relations prevailing in state owned banks differ substantially from those prevailing in privately owned banks. We have taken consideration of these facts in our study-design. Thus in our study we have considered accounts with the following characteristics:

1. Asset Type: Standard (i.e. unimpaired)
2. Industrial Classification: Manufacturing
3. Organization: Public and private limited non-government companies
4. Account Type: Term loan.
5. Location: Metropolitan centers of four states, namely Delhi, Maharashtra, Tamil Nadu and West Bengal

As borrowal accounts with the above characteristics are also classified according to whether they are in the Small Scale Industries (SSI)/Artisan sector or others, the analysis is done for all accounts with above characteristics as well as for SSI and non-SSI sectors separately. The accounts with the above characteristics can also be classified in respect of Bank Groups viz. *State Bank of India (SBI) and Associates*, *Nationalized Banks* and *Private & Foreign Banks*. In the study the analysis is also done for the above three bank groups separately. We have used one year's data - 2003. The four states have been chosen so as to broadly represent various geographical regions of the country. As corporate firms as well as lending banks in the metropolitan areas are expected to be similarly placed in terms of infrastructure availability, access to bank's decision makers and access to operational information of the firms on the part of lending banks, we have chosen credit accounts sanctioned from these branches only.

2.1 Some descriptive Statistics of the datasets analyzed

Selected descriptive statistics of our datasets are presented at the end. It may be observed that average size of credit vary substantially over the four regions. The distribution of credit is highly skewed in the positive direction for all regions and bank groups. The average credit size of private sector banks is generally higher than that of other bank groups in three regions out of four. Interest rates on the other hand are more range bound across regions and bank groups. This could be a reflection of lack of depth of the bank credit market in India, leading to concentration of credit over a very narrow range of risk profiles of the borrowers.

3 Methodology

A relationship between economic variables is generally postulated in real terms, though many a times the data are available in nominal terms. Although one routinely uses some price index to remove the price effect, a lot of uncertainty remains about the choice of price index. Prof. P. C. Mahalanobis (1960) encountered such a problem when he wanted to examine the impact of planning on the socio-economic conditions on a defined set of people over two time periods or to compare the conditions of two sets of people at a given point of time. In particular, he was interested to find out whether the economic condition of rural populations of India has changed for better due to the beneficial effect of India's planning process. For this, he wanted to use the *National Sample Survey*(NSS)[1] Consumption expenditure survey results for two rounds - 7th round (October 1953 to March

[1]Now *National Sample Survey Organisation*(NSSO)

1954) and 9th round (May to Novemember 1955). Following the standard Engel's law of demand, Prof. Mahalanobis hypothesized that improvement in living conditions should get reflected in the consumption basket of the people. Thus with rise in real income, the share of expenditure on food in total expenditure should decline. If X is taken as total per capita household consumption expenditure for a given period (say a month) and Y be the share of food articles in that total expenditure for the corresponding period, what is being sought to be examined is a comparison of two regression functions - Y on X for two time periods. If the economic conditions of people have registered an improvement, we would expect that to show up in a comparison of the regression curves for the two periods. But this comparison of two regression curves would not be appropriate as due to inflation, nominal expenditure classes (i.e. X) over two periods would become incomparable. To solve this problem, Prof. Mahalanobis suggested a comparison of mean of Ys' for each fractile or percentile groups of total expenditure between two periods. A formal description of his proposed method can be given in the following way:

Let (X, Y) and (X', Y') be two bivariate populations. We would like to compare the regression functions $\mu(x) = E(Y/X = x)$ and $\nu(x) = E(Y'/X' = x)$. If X and X' are not comparable because of location shift, we may consider the following functions called *Fractile Graphs* (FGs) or *fractile regressions*

$$\tilde{\mu}(t) = E(Y/X = F^{-1}(t)) \text{ and } \tilde{\nu}(t) = E(Y'/X' = F'^{-1}(t))$$

Where F and F' are the distribution functions of X and X' respectively. The comparison of the two relationships is now based on $\tilde{\mu}(t)$ and $\tilde{\nu}(t)$. It may be noted here that FGA has an apparent similarity to *Lorenz Curve* that has been a popular tool for analysis of income inequality.

In order to estimate $\tilde{\mu}(t)$ and $\tilde{\nu}(t)$ Prof. Mahalanobis adopted a heuristic method. After arranging the sample data points as per the values of X in an ascending order, he calculated mean of the Y-variable for each of the fractile groups which are formed by dividing the ordered sample observations into g groups of equal sizes. These means are labeled $y'_1, y'_2, y'_3 \ldots y'_g$ and the polynomial graph obtained by joining adjoining (i, y'_i) points by straight lines is the estimated FG. Since the estimated FG consists of continuously joined straight-line segments, the estimate is usually not smooth even though population FG may be smooth in nature. To obtain a smooth curve the modern kernel regression method may be used. The methodological details of this approach and the statistical properties of the estimates are given in (Sen, 2005). As it is well known the bandwidth used in smoothing is the most critical parameter that determines the degree of smoothness obtained by a kernel regression. The choice of kernel is of much less importance. In our exercise we have used Nadarya-Watson type weight function and standard normal density function as kernel. The optimal bandwidth is chosen by using the method of cross-validation carried out over a reasonable range of possible bandwidths.

Apart from estimating the FG for a particular sample, we also need to make statistical comparison between two estimated FGs. Prof. Mahalanobis in his characteristics way sought to solve the problem in a very novel way. In fact, what he was suggesting a kind of resampling method that has now blossomed into a very important arsenal in the toolkit of a statistician i.e. the method of Bootstrapping. Hall in his history of bootstrap methodology has pointed out in great detail the link between interpenetrating sub-sample method introduced by Mahalanobis and the modern bootstrap methodology. We describe below the main features of Mahalanobis' method:

Suppose we have data from two population (X_{1i}, Y_{1i}) for $i = 1 \ldots n_1$ and (X_{2j}, Y_{2j}) for $j = 1 \ldots n_2$ and we want to test the hypothesis that the corresponding FGs are not significantly different. The first sample of size n_1 is drawn from the first bivariate population in the form of two independent half samples, each size of $n_1/2$. The first half sample is then considered and the FG $G(1)$ is constructed. $G(2)$ is the second FG constructed from the second half sample. By construction, the two FGs have identical statistical properties. A combined FG $G(1,2)$ is then obtained by pooling the two half samples and constructing the FG based on this pooled sample. The error associated with this combined FG is measured by the area between the half-samples based FGs. By following a similar procedure, the corresponding FGs $G'(1)$, $G'(2)$ and $G'(1,2)$ are obtained for the second population using a sample size of n_2. Let $a(1,2)$ and $a'(1,2)$ be the errors associated and measured as described above with the graph $G(1,2)$ and $G'(1,2)$ respectively. Thus

$$a(1,2) = \int |(G(1) - G(2)| \; and \; a'(1,2) = \int |(G'(1) - G'(2))|$$

The area between two combined FGs is denoted by $S(1,2)$ and called *separation area* between the two graphs. The statistical error associated with $S(1,2)$ is given by the formula $E = \sqrt{a^2(1,2) + a'^2(1,2)}$. The significance of the observed value of $S(1,2)$ is tested by using the test statistic $\frac{S^2(1,2)}{E^2}$, which according to Mahalanobis would be distributed approximately as a chi-square variable.

Sen (2005) has proposed two statistics for comparing the smoothed estimates of the two FGs. Let $\hat{G}_{1,n_1}$ and $\hat{G}_{2,n_2}$ be the two smoothed FGs based on n_1 and n_2 samples for two populations then we can compute the following two statistics:

$$T_{n_1,n_2} = \int_0^1 |\hat{G}_{1,n_1}(t) - \hat{G}_{2,n_2}(t)|dt \; and \; S_{n_1,n_2} = \int_0^1 (\hat{G}_{1,n_1}(t) - \hat{G}_{2,n_2}(t))^2 dt$$

The first one would be called *Separation Area* between the two smoothed graphs and the second statistics as *Squared Difference between the FGs*. Under the null hypothesis of two FGs representing same population, we would expect a small value of these two statistics. The main hurdle to such hypothesis testing is the analytical intractability of the sampling distribution of these two statistics. The standard practice in this regard is to resort to bootstrapping. To test the significance of these two test statistics we bootstrap not from empirical joint distribution but from the kernel estimate of joint density of (X, Y). This is because the empirical joint distribution is neither continuous, nor strictly increasing and bootstrapping from it would have led to repetitions in data points and would distort the fractile graphs. For further details, (Sen, 2005) may be seen.

4 Results

We discuss the results of our empirical exercise in two parts. First we discuss inter-state comparison of FGs for all banks taken together. Assuming that credit appraisal and delivery system may differ as between different types of banks we compare the inter-state FGs for each bank group separately. This would allow us to examine to what extent the different observed differences in interest rate distribution are due to differential presence of various bank-groups in different regions. We have estimated the separation area (T) statistics for various combinations of states and tested for their significance by the method described above. For estimation of FG, the choice of bandwidths is of critical importance. As data driven bandwidth may not always generate smooth graph, which is important to draw any conclusion based on visual representation of FG, we have used a broader bandwidth other than the optimal one obtained by cross-validation method to get a reasonably smooth graph. This has been done in addition to FGs obtained by data driven bandwidth. The values of T-statistics and the visual representation of FGs (both data-driven and smooth) under comparison are presented at the end.

The conclusion that can be drawn based on these results are:

1. There are significant differences in FGs between any two states for the selected type of credit accounts and for SSIs as well as non-SSIs. This leads us to following conclusion:

   The borrowers of different states with similar characteristics like size of credit, type of credit facility and nature of economic activity and organization, are not identically placed with respect to interest rates charged. It is possible that other important but excluded characteristics like credit rating, state wise composition of industries etc. might account for the observed differences in interest rate structure for loan accounts.

2. There appears to be a clear negative relationship between size of credit

and interest rate charged for term loans. This is true for all states, for all bank-groups and for all types of borrowal accounts (e.g. SSIs and Other than SSIs).

3. Data-driven bandwidths do not smooth the observational data adequately and much higher bandwidths are required to obtain a reasonably smooth regression graphs between the two variables under study.

5 Conclusion

This study in the nature of exploratory data analysis brings out certain interesting features of a large dataset on credit accounts of scheduled commercial banks. The technique of Fractile Graphical Analysis, originally proposed by Prof. Mahalanobis is employed in this study and is found to be a useful technique in comparing the two regression curves, when the explanatory is subject to location shift as between two datasets.

The relationship between credit size and interest rate for a specified set of accounts with given characteristics has been examined in this empirical exercise. As interest rate may be dependent on a large number of factors, and to keep the exercise restricted to a bi-variate regression framework, we have examined the relationship for accounts belonging to a specific geographical areas and specific type of credit accounts and also for a specified type of banks. All though this delineation of datasets control for certain conditioning factor, it can not be said that we have been able to condition on many other factors that may affect the bi-variate relationship under scrutiny. A multiple regression approach is required to address these deficiencies of a bivariate framework. Work is under progress in this area.

It is found that interest on term loan has a negative relationship between size of credit and interest rate. It is also interesting to observe that there is significant difference in the estimated regression relationship for different states. Since policy makers would like to ensure similar interest rate environment over the entire country for borrowers with similar characteristics, our result has important policy implications and therefore, we need to probe further to find out what other factors could be possible determinants of observed differences. Notwithstanding this caveat, the policy makers need to be concerned about the credit market imperfection prevailing in the economy that this study apparently points to.

References

[1] Kashyap, A. K. and Stein, J. C. (1997): What do a Million Banks have to say about the Transmission of Monetary Policy. *Working Paper No. 6056*. NBER.

[2] Mahalanobis, P. C. (1960): A Method of Fractile Graphical Analysis. *Econometrica*. **28** 325–351.

[3] Sen, B. (2005): Estimation and Comparison of Fractile Graphs Using Kernel Smoothing Techniques. *Sankhyā*. **67** 305–334.

A Descriptive Statistics of the datasets analyzed

Table 1: Standard Assets[2]

State	#Obs	Credit Limit(in Lakhs)			Interest Rate (%)		
		Mean	C.V	IQR/Mean	Mean	C.V	IQR/Mean
Delhi	1467	361	3	0.5	14.3	0.1	0.1
Maharashtra	4400	1067.1	2.6	0.8	14.4	0.1	0.1
Tamil Nadu	2420	201.8	3.2	0.7	14.4	0.1	0.1
West Bengal	789	400.5	2.9	0.6	14.1	0.1	0.1

Table 2: Standard Assets and Small Scale Industries

State	#Obs	Credit Limit(in Lakhs)			Interest Rate (%)		
		Mean	C.V	IQR/Mean	Mean	C.V	IQR/Mean
Delhi	594	53.4	2	0.8	14.3	0.1	0.1
Maharashtra	1113	57.2	1.9	0.7	14.6	0.1	0.1
Tamil Nadu	742	44.9	1.9	1	14.4	0.1	0.1
West Bengal	290	54.8	2.4	0.6	14.2	0.1	0.1

Table 3: Standard Assets and Other than Small Scale Industries

State	#Obs	Credit Limit(in Lakhs)			Interest Rate (%)		
		Mean	C.V	IQR/Mean	Mean	C.V	IQR/Mean
Delhi	873	570.3	2.4	1	14.3	0.1	0.1
Maharashtra	3287	1409	2.2	0.9	14.4	0.1	0.1
Tamil Nadu	1678	271.1	2.8	0.8	14.4	0.1	0.1
West Bengal	499	601.4	2.4	0.8	14.1	0.1	0.1

Table 4: Standard Assets and State Bank of India & Associates

State	#Obs	Credit Limit(in Lakhs)			Interest Rate (%)		
		Mean	C.V	IQR/Mean	Mean	C.V	IQR/Mean
Delhi	278	338.7	2.8	0.7	13.4	0.1	0.1
Maharashtra	542	539.7	3.1	0.4	13.6	0.1	0.1
Tamil Nadu	515	251.6	2.2	0.8	13.6	0.1	0.1
West Bengal	239	564.2	3	0.4	13.2	0.1	0.1

Table 5: Standard Assets and Nationalised Bank

State	#Obs	Credit Limit(in Lakhs)			Interest Rate (%)		
		Mean	C.V	IQR/Mean	Mean	C.V	IQR/Mean
Delhi	892	266.6	3.6	0.4	14.4	0.1	0.1
Maharashtra	1864	290.1	3.6	0.4	14.5	0.1	0.2
Tamil Nadu	929	210.2	3.5	0.6	14.3	0.1	0.1
West Bengal	470	284.1	3	0.6	14.6	0.1	0.1

Table 6: Standard Assets and Private & Foreign Bank

[2] C.V–Coefficient of Variation; IQR–Inter-Quartile Range

State	#Obs	Credit Limit(in Lakhs)			Interest Rate (%)		
		Mean	C.V	IQR/Mean	Mean	C.V	IQR/Mean
Delhi	151	761.1	2.4	1	15.6	0.2	0.3
Maharashtra	218	1087.5	1.7	1	14.5	0.2	0.2
Tamil Nadu	211	204.6	5.3	0.1	15.8	0.3	0.4
West Bengal	19	861.6	1.1	1.5	15.2	0.2	0

B Comparison of smooth Fractile Graphs

Table 7: Comparison of smooth Fractile Graphs having Bandwidth 0.1

Population-1	Population-2	T_1	T_2	T_3	T_4	T_5	T_6
Delhi	Maharashtra	0.2616	0.2831	0.5215	0.433	0.1444	1.5308
Delhi	Tamil Nadu	**0.1243**	**0.1218**	0.2367	0.1667	0.1775	1.7628
Delhi	West Bengal	0.2488	**0.2011**	0.2787	0.5232	0.203	**0.9514**
Maharashtra	Tamil Nadu	0.2213	0.2075	0.3742	0.2743	**0.0832**	2.1539
Maharashtra	West Bengal	0.3179	0.4703	0.2796	0.2693	0.2023	**0.8558**
Tamil Nadu	West Bengal	0.2987	0.3147	0.3102	0.4135	0.2757	1.7163

The figures in **bold** are non-signifacant.

T_1 : Estimated Separation Area for Standard Assets
T_2 : Estimated Separation Area for Standard Assets and Small Scale Industries
T_3 : Estimated Separation Area for Standard Assets and Other than Small Scale Industries
T_4 : Estimated Separation Area for Standard Assets and SBI & Associates
T_5 : Estimated Separation Area for Standard Assets and Nationalized Banks
T_6 : Estimated Separation Area for Standard Assets and Private & Foreign Banks

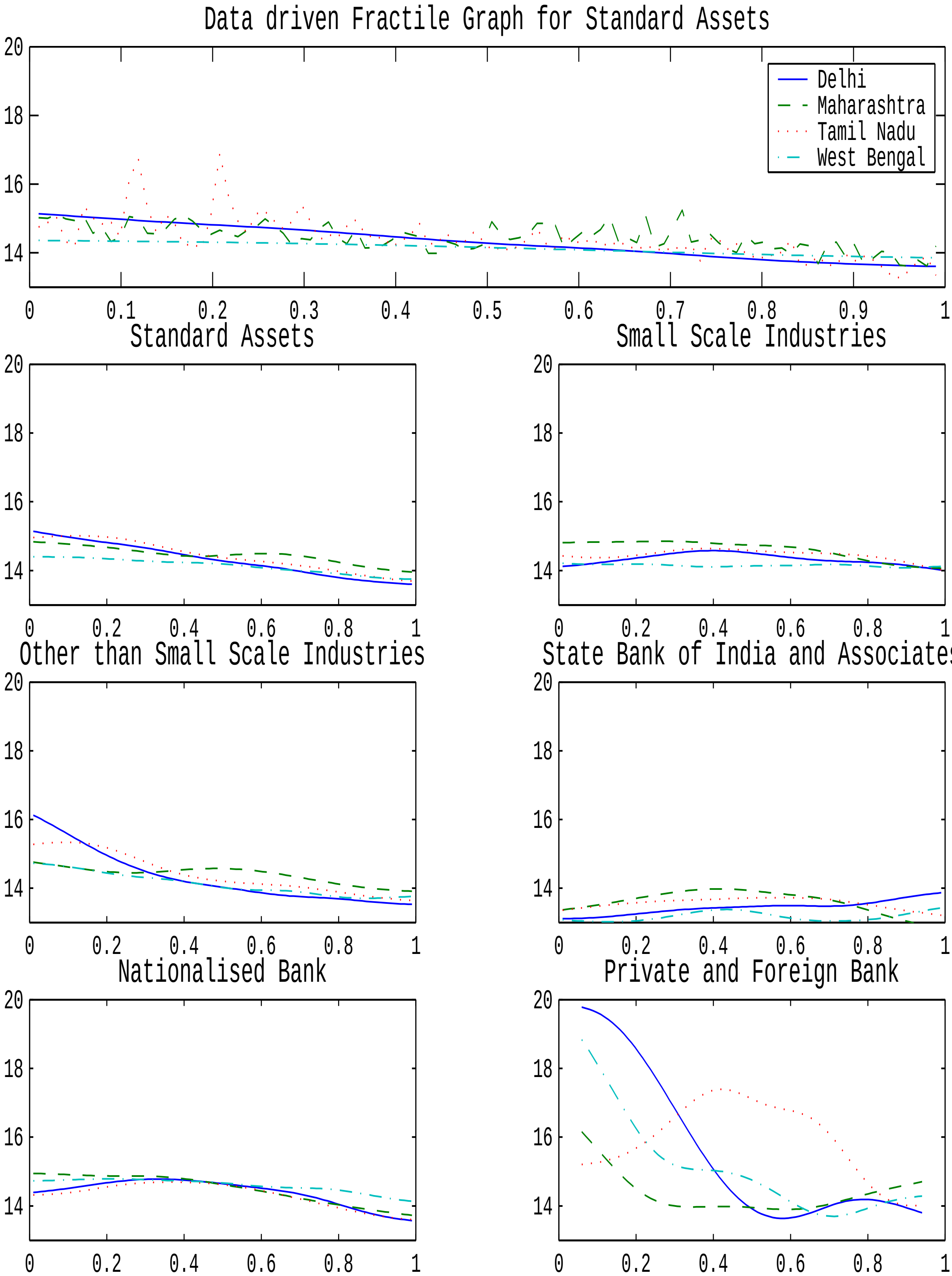

Figure 1: *Data Driven (topmost) and Smooth Fractile Graphs with bandwidth 0.1 (rest) for four states of India where X-axes represent fractiles based on Credit Limits and Y-axes represent Interest Rates*