

MPRA

Munich Personal RePEc Archive

The significance of socioeconomic factors on personal loan decision a study of consumer banking local private banks in Pakistan

Rehan Azam and Danish Muhammad and Suleman Syed
Akbar

IQRA UNIVERSITY

10. October 2012

Online at <https://mpa.ub.uni-muenchen.de/42322/>

MPRA Paper No. 42322, posted 1. November 2012 05:38 UTC

The significance of socioeconomic factors on personal loan decision a study of consumer banking local private banks in Pakistan

Rehan Azam

(Ph.D Scholar)

IQRA University-IU,

Defence View, Shaheed-e-Millat Road (Ext.), Karachi-75500, Pakistan.

E-mail: rehan_azam@hotmail.com

Contact No. +92 (300) 8294839

Muhammad Danish

(Ph.D Scholar)

IQRA University-IU,

Defence View, Shaheed-e-Millat Road (Ext.), Karachi-75500, Pakistan.

E-mail: mr_dn_mpa@yahoo.com

Contact No. +92 (332) 7860520

Syed Suleman Akbar

(Ph.D Scholar)

IQRA University-IU,

Defence View, Shaheed-e-Millat Road (Ext.), Karachi-75500, Pakistan.

E-mail: suleman_akber@hotmail.com

Contact No. +92 (334) 2413288

Abstract

This paper explores the influences of the approved results of loans cases, the loan applicants' socioeconomic attributes in the decision of personal loan. The results can improve the credit quality and avoid the misjudgment of screening personal loan customers and also establish a better personal loan risk management forecasting model. The main purpose of the present paper was to evaluate significance of loan applicant socioeconomic attributes on personal loan decision in the local private commercial banks of Pakistan. The statistical techniques, descriptive and logistic regression were used. The model identified that out of six independent variables, region, residence status and year with the current organization have significant impact on personal loan decision.

Keywords: personal loan; socio-economic; consumer banking; logistic regression model

1. Introduction

The primary problem of any lender is to differentiate between "good" and "bad" debtors prior to granting credit. Lately, credit risks have become one of the most important financial topics of interest, especially in the banking sector. The role of credit risks has changed dramatically over the last ten decades, from passive automation to a strategic device. Personal Loan is easy and

convenient cash with no equity. The utilization of a personal loan is at the discretion of the person taking the loan. Unlike a home or a Car Financing loan its utilization is not bound for a specific purpose. According to the State Bank of Pakistan's Prudential Regulations for Consumer Financing Part-A, (3) Consumer Financing (iv) Personal Loans mean the loans to individuals for the payment of goods, services and expenses and include Running Finance/Revolving Credit to individuals.

Personal Loan Decision refers approval or rejection of loan application. The Eight Common Attributes of Socioeconomic (1) region, (2) age, (3) gender, (4) household income, (5) residence status, and (6) years with current organization on the personal loan decision. Once the lender completes its review of the application, it will make a decision. If the loan is approved for the amount requested, there will be some additional paperwork required to finalize the loan before the process is complete. The lender also may approve a loan for a lesser amount than you requested or decline to make a loan.

The socio-economic variables establish the identity of the borrower for the purpose of the loan and looks at legal aspects. These variables have the highest importance and they capture various regional, age and other relevant differences. For example, it is often found that old man is less risky than young men. In general, the risk of default decreases with age. Home owners also represent a less risky category due to a house as collateral. Occupation, Employment or self-business Indicators Type of organization - Public sector / Private sector, Government / Non Government, length of employment. The occupation indicator is used for fixing repayment period and for the purpose of documentations. The other financial indicators are used to determine the quantum of loan.

2. Previous Research

Steenackers, and Goovaerts, (1989) analyze the significant variables of credit rating model that impact the personal loans of lending banks in Belgium. Their results show that the significant variable belonging to the credit criteria is the loan period, while those not belonging to the credit criteria include a total of 11 variables: the borrower's age, whether the phone number is provided, the permanent residence address, the current residence address, the duration of the job, the location area, occupation, whether the job is at the government sector, the monthly income, the house ownership, and where he (or she) has previous loans. Tor and Kasper 2003) show whether the house is the borrower's own property, occupation, income, debt ratio, and guarantor are significantly positively correlated with the borrower's credit risk, while age, region or city of residence, and whether there are other loans are significantly negatively correlated with the borrower's credit risk.

Holmes, Isham, Petersen, and Sommers, (2007) indicate that the community bank relies on credit scoring but not relationship lending. The low-income households with strong ties to the community development credit union (CDCU) are likely to receive loans despite poor credit histories.

2.1. Region: Impacts of the Loan Applicants' Region on personal loans decision

Region means the area of the country that borrower lives. As people of similar wealth tend to live in the same location, the geographic criterion can indicate a borrower's level of financial wealth. Some suburb might attract richer residents and this could result increase in housing and property prices. This also affects the collateral value and probability of default. The hypothesis H1 is established as follows:

2.2. Impacts of the Loan Applicants' Age on personal loans decision

Age measures the borrower's age in years. Boyle, J. N., Hamilton, and Thomas (1992) confirm that older borrowers are more risk adverse, and therefore the less likely to default. Thus banks are more hesitant to lend to younger borrowers who are more risk averse. The hypothesis H2 is established as follows:

2.3 Impacts of the Loan Applicants' Gender on personal loans decision

Gender is a fair discriminatory - base on the statistical default rates of men versus women. There are ample evidences that women default less frequently on loans because women are more risk adverse (Coval & Shumway 2000). The hypothesis H3 is established as follows:

2.4 Impacts of the Loan Applicants' Income on personal loans decision

(Attanasio, Goldberg, P. K., and Kyriazidou, 2008) find that, with the exception of high-income households, consumers are very responsive to maturity and less responsive to interest rate changes. (Caselli, Gatti, and Querci, 2008) indicate that the loss given default rate (LGDR) is more sensitive to the default-to-loan ratio, the unemployment rate, and household consumption for households. From a practical perspective, after interviewing the credit staffs of the case bank, we find whether the borrower has provided the proof of income is considerably correlated with the stability of the company he works for. If the borrower's company can provide the proof of his financial capability such as income tax withholding voucher, insurance card, the passbook of salary transfer, and tax disc, it will indicate that the borrower has a sure source of income and his company has a certain scale and system. Hence, there are fewer concerns about the borrower's loan repayment source and ability. Therefore, both the general banks and the financial institutions related to the loan think that the security and the reliability of the borrower who can provide the proof of income are higher than those of the borrower who cannot. Thus, the paper expects if the borrower can provide the proof of income, his overdue probability of automobile loan will be lower than that of the automobile loan of the borrower who cannot. Thus, the hypothesis H4 is established as follows:

2.5 Impacts of the Loan Applicants' Residence Status on personal loans decision

According to the practical loan experience of credit staffs of the case bank over the past years, customers are classified based on the status of the residence. The statistical analysis of customers' payment records shows that the borrowers own the residence belong to the low-risk group of the probability of personal loan, and the group has a considerable impact on the probability. Therefore, the paper expects if the borrowers own the house the low-risk list of the case bank, chances of approval of loan application will be higher. The hypothesis H5 is established as follows:

2.6 Impacts of the Loan Applicants' Year with current organization on personal loans decision

Steenackers and Goovaerts, M. J. (1989) point out that the job service years of the borrower is significantly negatively correlated with the credit rating of personal loans, that is, the longer the job service years of the borrower is, the more stable the income source and the ability of repayment will be. Therefore, the paper expects that the job service years of the borrower are negatively correlated with the overdue probability of automobile loan. Hence, the hypothesis H6 is established as follows:

3. Hypotheses

H1: Region has a significant impact on personal loan decision.

H2: Age has a significant impact on personal loan decision.

H3: Gender has a significant impact on personal loan decision.

H4: Household income has a significant impact on personal loan decision.

H5: Residence status has a significant impact on personal loan decision.

H6: Years with current organization has a significant impact on personal loan decision.

4. Research Method

4.1 Population, Sample and Data source

As the all banks are now fully equipped with MIS (Management Information System) or different accounting and operation software, in the paper has keyed in and scanned the data of personal loan applications data or cases at the beginning of application, these data are then filed after being analyzed. The advantage of the current way is that loan customers' attributes can be set up more completely and systematically in the file so as to be served for future analysis, cases screening, knowing the criteria, quality of assessment and the use of statistics on relevant information. This is also the competitive advantage of the case bank in the current personal loan business market.

The population includes all those Local Private Banks who are offering Personal Loans. There are 12 (twelve) local private Banks with 6,850 Branches in Pakistan are offering personal Loans to the customers (2010). The past three-year loan applicant data was collected. The paper considers the personal loan cases accepted, rejected and outstanding the bank as the research object, including the all cases which are in record according to the loan application format. The samples of the paper come from the loan cases provided by the major local private banks who are dealing in personal loans from July 2010 to June 2012. A total of 4,112 loan cases serve as the research object of the paper for analyzing the significance of applicants' socio-economic attributes on personal loan decision, which includes 3,664 approved or reject or 88.9 in term of percentage and 458 or 11.1 in term of percentage are outstanding cases.

4.2 Logistic regression model

According to the aforementioned hypotheses, the paper builds a logistic regression model as follows:

$$\text{DECISION} = \beta_0 + \beta_1 \text{REGION} + \beta_2 \text{AGE} + \beta_3 \text{GENDER} + \beta_4 \text{INCOME} + \beta_5 \text{RESIDENCE} + \beta_6 \text{ORG_YEAR} + e$$

where DECISION is the applicant's loan approval probability; REGION is the applicant's city; AGE is the applicant age; INCOME is the house hold income of applicant; RESIDENCE is the applicant residence status; ORG_YEAR is the applicant's year with the current organization where he or she employee or self-employed;; β_0 is the intercept; β_1, \dots, β_6 are parameters of regression model; e is the error term of regression model. Each variable in the above regression model will be introduced in the following sections.

4.3 Measurement of dependent variable

Applicant's loan approval probability (DECISION): the paper divides the dependent variables into two types: 1 for the approval case; 0 for the declined case where the borrower's application rejects.

4.4 Measurement of independent variables

The approved results of loan cases Applicant's Region (REGION): refers to the loan applicant's city are divided into 8 major cities of Pakistan, 1 for the Bahawalpur, 2 for Faisalabad, 3 for the Islamabad, 4 for the Karachi, 5 for the Lahore, 6 for the Multan, 7 for the Peshawar and 8 for the Sialkot.

Applicant's Age (AGE): defined as the age of loan applicant are divided into 6 types such as 1 for the 20-29years, 2 for the 30-39, 3 for the 40-49, 4 for the 50-59, 5 for the 60-69 and 6 for the 70 plus.

Applicant's Gender (GENDER): defined as the sex age of loan applicant is divided into 2 types such as 1 for the male and 2 for the female.

Applicant's Household Income (INCOME): the applicant's annual household income are divided into four types and that is 1 for the below 10,000, 2 for the 10,000 -19,000, 3 for the 20,000-29,000 and 4 for the 30,000 plus.

Applicant's Residence Status (RESIDENCE): this refers to applicant's current residence status and this divided into two types, 0 for the Rent and 1 for owner.

Year with current organization (ORG_YEAR): this refers to how long the borrower works for his current company; that is, the years of service is used to measure the degree of job stability. This divided into four categories; 1 for the 0-3years, 2 for the 3-6years, 3 for the 6-10years and 4 for the 10-Above years.

5. The Results of Hypotheses Testing

5.1 Descriptive statistic results

Table 1 summarizes the descriptive statistic of the samples and shows the relation between customer gender and decision of approval or rejection of loan. Which shows Male has more decline as well as approval ratios.

Table 2 shows that Karachi has high declined and at the same time approval ratio. This table basically shows the relationship between region of application and decision.

Table 3 shows the relationship between decision and income level. Here we can see that income is not might be a very important factor while making decision. It may be because Banks are dealing with their existing customer they might go for their previous history not their income.

Table 4 show that the age between 30-39 has high decline and approval ratio 107 decline and 1645 approved. As we move forward to age 40-49 here we can see that we have 282 declined and 338 approved there is not much difference but at the age 50-59 we have more declined and less approval ratio and so 60-69 and 70-79 because bank must have some criteria of age for approval of loan.

Table 5 shows that the applicants which have rental home has high declined and low approved ratios and at the same time the applicants who are not renters, approval ratio is high and decline ratio is less.

Table 6 indicates that the applicants stay at current organization since long has less chance of defaulter that's why more years at current organization shows less declining ratio.

4.2 Logistic regression results

Table 7 Classification Table: The table shows that the total number of applicants is 3664 out of which 886 are declined and 2778 are approved

Table 8 Variable in the Equation: Under Variables in the Equation the intercept-only model is $\ln(\text{odds}) = 1.143$ If we exponentiate both sides of this expression we find that our predicted odds $[\text{Exp}(B)] = 3.135$. Since 2778 of our variables are approved and 886 are declined observed odds are $2778/886 = 3.135$

The probability of this model is $3.135 / (1 + 3.135)$ is 0.75 that's is 75%

Table 9 Hosmer and Lemeshow Test: The Hosmer-Lemeshow statistic indicates a Good fit if the significance value is greater than 0.05. Here, the model adequately fits the data.

Table 10 Variable in Equation: The test was conducted on .95 of confidence interval. Table shows the significant results of test here we can see region, residence status and year with current organization which have sig value 0.00 is less than 0.05 which makes our test significant hence age and income significant impact on approval and rejection of loan requests.

$$\text{Decision} = \text{Intercept}^{3.257} - \text{Region (3)}^{0.639} - \text{Region (4)}^{0.939} - \text{residence (1)}^{1.504} \\ - \text{org_year(1)}^{5.289} - \text{org_year(2)}^{4.742}$$

6. Summary and Concluding Remarks

The main objective of this study was to find out the loan applicant socio-economic attributes on personal loan decision. The 6 hypotheses were tested including Region has significant impact on personal loan decision, Age has significant impact on personal loan decision, Gender has significant impact on personal loan decision, income has significant impact on personal loan decision, residence status has significant impact on personal loan decision and year with current organization has significant impact on personal loan decision.

After studying the whole data file it can be concluded that we have region, age, gender, income, residence status and years with current organization have positive impact on decision of facilitating loan or not. These variables play a vital role to predict the decision for pending/missing values. According to test region like Karachi or having more population has great chance of acceptance of loan application, Age more than 40 has low chance of approval, hence income also very important in this matter to decide whether to approve or reject request and here we have two more variable residence status which would be rental or owned and since when he's residing there this makes the risk low of being defaulter. Though these are not only the reason for rejection of application there are also many reason in data file like bad credit history of customer (Electronic Credit Information Bureau data check), exceeds aggregate debt burden ratio and so on. So here we have find in data that region, residence status and year with organization are more significant to predict the dependent variable (decision).

We studied in this data file that the different variables are applied in the decision of approval or rejection of loan. The factors which are predicting the decision in the data file having the positive relationship with income, age, gender, region, residence status and year with organization. On the basis of our analysis it's suggest to the company that they must focus more on the variable which has significant impact on decision.

References

- Attanasio, O. P., Goldberg, P. K., & Kyriazidou, E. (2008). Credit constraints in the market for consumer durables: Evidence from micro data on car loans. *International Economic Review*, 49(2), 401-436.
- Boyle, M., J. N. Crook, R. Hamilton, and L. C. Thomas (1992), Methods for credit scoring applied to slow payers. In; Thomas, L. C., Crook, J. N., Edelman, D. B. (Eds), "Credit scoring and credit control", Oxford University Press, Oxford, Pg 75-90.
- Caouette, J.B., E.I. Altman, and P. Narayanan. (1998). *Managing Credit Risk: The Next Great Financial Challenge*. New York: John Wiley & Sons Inc.
- Caselli, S., Gatti, S., & Querci, F. (2008). The sensitivity of the loss given default rate to systematic risk: New empirical evidence on bank loans. *Journal of Financial Services Research*, 34(1), 1-34.

Coval, J., T. Shumway (2000). “Do behavioural biases affect prices?” University of Michigan, Working Paper.

Hand, D.J., & W.E. Henley. 1997. Statistical Classification Methods in Consumer Credit Scoring: A Review. Journal of the Royal Statistical Society: Series A (Statistics in Society), 160: 523-541.

Holmes, J., Isham, J., Petersen, R., & Sommers, P. M. (2007). Does relationship lending still matter in the consumer banking sector? Evidence from the automobile loan market. Social Science Quarterly, 88(2), 585-597.

Steenackers, A., & Goovaerts, M. J. (1989). A credit scoring model for personal loans. Insurance Mathematics Economics, 8(1), 31-34.

Tor, J., & Kasper, R. (2003). Bank lending policy, credit scoring and value-at-risk. Journal Banking and Finance, 27(4), 615-633.

TABLE 1:
customer gender * Decision Cross tabulation

Count

	Decision		Total
	Declined	Approved	
customer gender Male	789	2514	3303
Female	97	264	361
Total	886	2778	3664

TABLE 2:
Region * Decision Cross tabulation

	Decision		Total
	Declined	Approved	
Region BAHAWAL PUR	0	209	209
FAISALABAD	64	253	317
ISLAMABAD	135	379	514
KARACHI	428	764	1192
LAHORE	113	391	504

MULTAN	55	284	339
PESHAWAR	20	155	175
SIALKOT	71	343	414
Total	886	2778	3664

TABLE 3:
Decision * income Crosstabulation
Count

		Income				Total
		1000-9000	10000-19000	20000-29000	30000-above	
Decision	Declined	257	369	260	0	886
	Approved	504	741	1224	309	2778
Total		761	1110	1484	309	3664

TABLE 4:
Decision * age of applicant Crosstabulation
Count

		age of applicant					Total	
		20-29	30-39	40-49	50-59	60-69		70-79
Decision	Declined	71	107	282	247	107	72	886
	Approved	901	1538	338	1	0	0	2778
Total		972	1645	620	248	107	72	3664

TABLE 5:
Decision * residence status Crosstabulation
Count

		residence status		Total
		rental	owned	
Decision	Declined	587	299	886
	Approved	780	1998	2778
Total		1367	2297	3664

TABLE 6:

Decision * year with current organization Crosstabulation
Count

	year with current organization	Total
--	--------------------------------	-------

		0-9	10-19	20-above	
Decision	Declined	843	42	1	886
	Approved	2334	228	216	2778
Total		3177	270	217	3664

TABLE 7:

Classification Table^{a,b}

Observed			Predicted		
			Decision		Percentage Correct
			Declined	Approved	
Step 0	Decision	Declined	0	886	.0
		Approved	0	2778	100.0
		Overall Percentage			75.8

a. Constant is included in the model.

b. The cut value is .500

TABLE 8:

Variables in the Equation

	B	S.E.	Wald	df	Sig.	Exp(B)
Step 0 Constant	1.143	.039	877.259	1	.000	3.135

TABLE 9:

Hosmer and Lemeshow Test

Step	Chi-square	df	Sig.
1	6.727	8	.566

TABLE 10:

Variables in the Equation

	B	S.E.	Wald	Df	Sig.	Exp(B)
Step 1 ^a Region			49.655	7	.000	
Region(1)	18.839	2485.440	.000	1	.994	1.519E8

Region(2)	-.137	.296	.215	1	.643	.872
Region(3)	-.639	.255	6.288	1	.012	.528
Region(4)	-.939	.225	17.353	1	.000	.391
Region(5)	-.270	.266	1.025	1	.311	.764
Region(6)	.322	.313	1.057	1	.304	1.380
Region(7)	.390	.385	1.026	1	.311	1.476
Age			409.190	5	.000	
age(1)	23.318	4479.363	.000	1	.996	1.339E10
age(2)	23.374	4479.363	.000	1	.996	1.417E10
age(3)	20.819	4479.363	.000	1	.996	1.100E9
age(4)	15.244	4479.363	.000	1	.997	4171478.770
age(5)	-.125	5795.088	.000	1	1.000	.883
Income			49.710	3	.000	
Income(1)	-19.625	2041.807	.000	1	.992	.000
Income(2)	-19.751	2041.807	.000	1	.992	.000
Income(3)	-18.764	2041.807	.000	1	.993	.000
residence(1)	1.504	.128	137.719	1	.000	4.498
org_year			14.940	2	.001	
org_year(1)	-5.289	1.605	10.861	1	.001	.005
org_year(2)	-4.742	1.623	8.533	1	.003	.009
Constant	3.257	4922.772	.000	1	.999	25.964

a. Variable(s) entered on step 1: Region, age, salary, residence, residenceyr.