RISK PROFILE OF TURKISH HOUSEHOLDS FOR CONSUMER CREDITS

Mehmet Baha KARAN
Hacettepe University,
Department of Business Administration,
Beytepe, Ankara 06800, Turkey
Phone: 0090 312 297 8700 ext: 145
Fax: 0090 312 297 2055
e-mail:mbkaran@hacettepe.edu.tr

Özgür ARSLAN
Hacettepe University,
Department of Business Administration,
Beytepe, Ankara 06800, Turkey
Phone: 0090 312 297 8700 ext: 129
Fax: 0090 312 297 2055
e-mail:arslan@hacettepe.edu.tr

This document has been produced with the financial assistance of the Scientific and Technological Research Council of Turkey (TUBITAK). The authors would like to grant their special thanks to Aydin Ozkan, Chrisostomos Florackis and Elisabeth O'Dowd for their valuable comments.

RISK PROFILE OF TURKISH HOUSEHOLDS FOR CONSUMER CREDITS

Abstract

The objective of this paper is to point out the roles of socioeconomic and demographic characteristics of Turkish households on their credit risk. For this reason, we use Budget Survey of Turkish Households conducted by Turkish Statistical Institute for the year 2003. Our sample is comprised by 25,586 households. We allocate households to the credit risk categories according to their income-expense balance. Our results obtained through neural networks, decision tree and logit analyses confirm that we are able to accurately allocate most of the households into two different credit risk categories. Our binomial logit estimations provide wide range of results on the direction of relationship between the credit risk categories of the Turkish households and their assets, savings, job characteristics and internal features.

Key Words: Credit risks, Turkish households, Logistic regression
Jel-Codes: G14,G20

RISK PROFILE OF TURKISH HOUSEHOLDS FOR CONSUMER CREDITS

Introduction

Being one of the world's leading emerging economies and a candidate for the European Union membership, Turkey's consumer credit market offers vast opportunities to foreign financial institutions. Besides, Turkey's consumer finance industry has been rapidly growing since the end of the financial crisis in the year 2001. This outcome is mainly owed to the rising disposable income on back of robust economic growth and rising employment since the end of the crisis. Moreover since the last quarter of the year 2002, foreign banks have been intensely merging, taking over or acquiring Turkish domestic banks. Specifically, the share of foreign banks in assets of Turkish banking system rose from less than 4 % in 2002 to 38.8 % in 2006. The entry of foreign banks has escalated competition with domestic banks particularly in consumer lending. Thus, the quantity of increase in Turkish consumer credits from the last quarter of the year 2002 to the year 2006 aligns with the rise of the share of foreign banks in the Turkish banking system (The Banks Association of Turkey, 2007).

It is emphasized in the International Monetary Fund's country report (November, 2007) that, since household liabilities to GDP in Turkey is merely less than 20 % of that of the Euro area, entering Turkish consumer credit market is highly advantageous. However the issue that how well is the performance of the Turkish lending market is an essential one for the financial institutions to take advantage of all the offered opportunities. As indicated by Getter (2006), a well functioning credit market is constituted once borrowers are treated differently in terms of credit availability and charged rates. Consequently, the well functioning of the consumer credit market in Turkey depends on the amount of information that enables the financial institutions to differentiate good quality borrowers from bad quality ones. Nevertheless, in the theoretical framework of Cutts, Van Order and Zorn (2000) it is

shown that expansion of consumer credit markets depends on the availability of information about borrowers.

The objective of our paper is to increase the information for financial institutions that aim to be lenders for Turkish households, hence contribute to the well functioning of the Turkey's credit market. In other words, through the results of the publicly available survey in Turkey, we aim to determine the risk profile of Turkish consumers. Therefore, this paper informs banks, building societies, credit card institutions and large retailers about the risk characteristics of their potential borrowers. In spesific, the institutions are acknowledged about particular socioeconomic and demographic features of Turkish households whom they must either target or avoid. Moreover, our work also contributes to efforts for maximizing wealth of Turkish consumers because borrowers face credit constraints when lenders are not able to obtain sufficient information to evaluate their credit risks correctly. Nevertheless, this improvement in the demand side of the credit market is called as the democratization of credit lending by Lyons (2003).

The first attempt of a study like ours would be to acquire the risk profiles of borrowers through database of financial institutions which specify the households who do not systematically pay back their loans. However in Turkey, even still in the most advanced economies in the world[1], availability of credit history information is a proprietary and difficult for academics[2]. Besides it is highlighted by Hsieh (2006) that even assuming that bank databases are available, they are too difficult to be analyzed due to their multi-dimensional structure. Therefore, the data of our work is obtained from the annual *Budget Survey of Turkish Households (Hanehalkı Bütçe Anketi)*[3], which is conducted by Turkish Statistical

[1] See, Getter (2006)
[2] One of the studies that tackle the availability problem is that of Carling, Jacobson, Linde and Roszbach (2007) since the authors build their data on a major Swedish bank's business customers.
[3] The data is collected using the approach of Nomenclature of Territorial Units for Statistics (NUTS) which is the European Union territorial classification system. By this system the indicators of consumption expenditures

Institute (*TURKSTAT*) for the year 2003. *TURKSTAT* is a government organization responsible for periodically collecting, evaluating, analyzing and publishing statistics about Turkey's economic, social, demographic, cultural, environmental and scientific fields. Therefore, this survey reflects the socioeconomic and demographic characteristics of Turkish households in detail. We derive risk concepts that may emerge during consumer credit usage of Turkish households through both the individual responses to the survey and their observed annual spending behaviors documented by employed surveyors of the *TURKSTAT*. Households, who are respondents of the survey, are chosen by *TURKSTAT* under a scientific framework to ensure that the sample group most accurately represents the entire Turkish households. Furthermore, for an entire month, all the revenues and expenditures of the selected households are closely monitored and documented by the survey conductors. Even, the expenditures and revenues of the householders are asked to be evidenced by shopping receipts and payrolls or wage slips. For this reason, the responses to the survey are not only comprised by the personal statements of the sample households but also through personal observations by conductors. For this reason, biasedness in the survey responses is perfectly minimized. Another strength of our paper is, we conduct our analyses on an unusually high number of observations, namely 25,586 households, that responded to the *TURKSTAT* survey. The large sample size in our analyses helps to increase reliability of our empirical tests.

Furthermore the year 2003, for which we conduct our analyses, points out a strategic turning point for the Turkish economy. As it is indicated in the *Balance of Payment Report* (December 2004, page 30) prepared by Central Bank of the Republic of Turkey, after the crisis period in the years 2001 and 2002, Turkish economy promptly entered into the expansion stage in the year 2003. In that year, macro economic fundamentals, such as inflation and interest rates fell back to standard levels thus economic stability is reached. For

---

and income distribution are obtained at national level, at rural-urban division and at provincial level by using NUTS1 and NUTS2 classification.

this reason, investments to Turkish credit market by foreign institutions initiated and intensified in the year 2003.

Before testing our regression model we use logit analysis, neural network and decision tree methods to verify how correctly we assign each households into respective credit risk categories through a binary classification.

Our binomial logit estimation provide wide span on findings on the risk profile of Turkish households. We specifically highlight which socioeconomic and demographic features are directly associated with being a low credit risk household. In other words, our results show the relationship between being highly likely to be qualified for consumer credits and assets, savings, job characteristics and internal features of Turkish households.

The next section of our study presents the background of our work followed by the section involving the description of our data and our methodology. Afterwards we evaluate the results of the logit, decision tree and neural networks analyses and present our univariate test. We demonstrate the empirical findings from our binary logistic estimation in the next section. Finally the last section includes the concluding remarks.

Background

It is shown by King and Levin (1993) that a country's financial development is directly related to its economic growth and improvement in economic efficiency. Right after recovering the adverse effects of the 2001 crisis, since the year 2003 Turkey has been enjoying a high economic growth and welcoming foreign originated financial institutions and banks into the domestic financial system. Two of the leading factors that contribute to the positive progress in Turkey's macro economic and institutional structure are successful application of International Monetary Fund (IMF) packages and the initiation of the accession negotiations with European Union (EU). Therefore, foreign financial institutions and domestic banks, which have pulled through the damaging impact of the last financial Turkish

crisis expeditiously, concentrated their activities particularly on the consumer financing of Turkish households. In accord with this, consumer credit has been becoming a fundamental financial source for households in Turkey. Moreover, between the years 2003 to the date Turkey has been experiencing an annual average GNP growth rate of 7 %, which is mainly thanks to consumer spending. Consequently, a well functioning domestic consumer credit market is essential for sustainability of the high GNP growth rate.

Besides a country's financial development is measured by some characteristics of financial intermediaries such as; amount of issued credits, size of financial institutions to GDP and quantity of financial services (Ge and Qiu. 2007). However the availability of information in credit markets is crucial for triggering the financial development in a country. In order to encourage the supply side of the consumer credit market, financial institutions must be able to have sufficient information concerning borrowers to find out how much and at which rate to lend.

As it is highlighted by Thomas (2000) financial institutions must be able to differentiate profitable customers from the risky ones. The rationale behind this argument is the objective of the institutions to maximize their profits and minimize their credit risks. Therefore this study aims to increase the information symmetry in the market through differentiating low risk customers from the high risk ones[4].

Swain (2007) highlights the importance of considering socioeconomic and demographic aspects while elaborating the supply side of consumer credits. Hazembuller, Lombardi and Hogarth (2007) point out the *Three C*'s of credit, which are respectively; capacity, collateral and character of borrowers representing income, assets and demographics of householders. The authors find that older, house owner, single, male and working people and the ones having higher financial assets are less risky borrowers for the US credit market.

[4] Following Getter (2006) we refrain using the term "credit constrained" for high risk borrowers, wheras this term is more appropriate for situations when borrowers face credit constraints since lenders do no have enough information to correctly evaluate the risk of a borrower due to asymmetric information.

Besides, Hogarth and Hilgert (2002) emphasizes the third C of the credit, which is the character, and the result of their survey show that older, minority and limited-educated respondents are relatively riskier borrowers in US credit market.

Budget Survey of Turkish Households, conducted by *TURKSTAT*, provides us wide range of information about Turkish borrowers concerning the *Three C*'s of credit. The implementation of this survey has commenced in the year 2003 and ever since has been continuing to be conducted in the following years. However, only the results of the year 2003 are publicly available. Moreover, the questions in the survey and the number and identity of respondents are different in the each year. Thus the method of the survey ensures the objectivity of our results.

The database of our study is emanated from the information on both each of individuals in a household and household itself. In other words, while the financial data on households encompass entire family, demographic data is only comprised from the household head, usually father of family. This classification is perfectly appropriate for the Turkish family structure given that solidarity and collaboration in Turkish families are very strong and the household head has the most influential role in household decisions[5].

On the one hand Avery, Calem and Canner (2004) point out that excessive spending of a household may be triggered by some extreme cases in one year such as a medical emergency. For this reason, these extreme cases may create an inconvenience for its sufferers since they will be considered in the same risk category as the ones displaying chronic excessive spending for consecutive years. Conducting our analyses for a single year may conjure up such a drawback for interpreting our results. On the other hand, given that we have a very large sample size for our analyses, we assume that the households facing the extreme cases for the year 2003 are exceptions among the high risk category of consumers.

[5] Moreover, in the study of collectivism and individualism by Hofstede (1980) Turkey ranks as the 28th in individualism index, thus Turkish culture has repeatedly been described as a moderate collectivistic culture, in which family ties are characterised as being very strong.

It is acknowledged by Rosenberg and Gleit (1994) that quantitative techniques are very useful for assisting consumer risk determination. Besides, in the study of Joos, Vanhoof. Ooghe and Sierens (1998) it is found that decision tree method is superior for the qualitative and short scheme data whereas logit models are more consistent in the credit risk decision process. Moreover, since the beginning of the decade, researchers show that neural networks perform very well for the complex and unstructured problem of credit risk classification (West, 2000). Finally so as to verify the accuracy of our credit risk classification we employ neural networks methodology along with decision tree and logistic analyses.

Data and Methodology

Our analyses are built on the Budget Survey of Turkish Householders, conducted by *TURKSTAT* for the year 2003. Initially the survey included 25,764 respondents however after excluding 178 observations displaying invalid data we are left with the replies representing 25,586 households in Turkey. The questions in the survey reflect the information on both households and their members. On the one hand, sum of revenues, expenditures and assets of the each member is accepted as a single data of a household. On the other hand, demographic information such as age, education and employment of household head is accepted to encompass the entire household.

While estimating our model we use the *Clementine 11.1* software package, which is a licensed data mining program by *SPSS*.

[Insert Table 1 about here]

Table 1 provides the broad definitions of the data we use for our analyses. The definitions in the table are directly organized from the *TURKSTAT*'s *2003 Guided Handbook*

for Household Data-set Definitions of Budget Survey of Turkish Households. As can be seen from the table, majority of the variables are included as dummy variables in our analyses.

Table 2 demonstrates descriptive statistics of the variables used in our analyses. Therefore the table gives detailed information concerning the socioeconomic and demographic aspects of the budgetary characteristics of Turkish households.

[Insert Table 2 about here]

Furthermore, we present the results of our correlation matrix in Table 3. The table provides evidence that the correlation does not exceed the level of 0.50 among the variables included in our analyses. Initially the survey allowed us to include around 40 independent variables to our analyses. However except the ones presented in the Table 3, the correlation coefficient between the variables is found to be always more than 50 % . For instance the respective correlation coefficients for some variables are as follows; being a blue collar worker and education, 0.71; living in a city and education, 0.65; income per person and marital status, 0.59. For this reason the total number of such variables we eliminated from our analyses is 26.

[Insert Table 3 about here]

While investigating credit risk profile of potential consumer credit customers; our binomial logistic estimation is based on dummy variables differentiated as the ones taking the value of unity if they are low risk households and those taking the value of zero if they are not found to be low risk households. In other words, low risk households are the ones that financial institutions must primarily target for consumer credits. Whereas, the other group is situated after the low risk group in pecking order of consumer credit supply.

However we do not have any information on households who pay back their debts punctually or those either default or delay their expiring outstanding debts. For this reason, while investigating risk profiles of households, we follow a different method. Accordingly, we firstly use their annual revenue and their annual expenditure items. Annual revenue is obtained by adding total annual cash and non cash revenues of a household. Annual expenditure is constituted by total annual expenditures realized by a household. Therefore we reach to the income and expense balance of households through difference between their annual revenues and expenditures. On the one hand we believe that households having a positive income-expense balance have low default risk given that their total revenues exceed their total expenditures. On the other hand the household group, which is identified with the dummy variable taking the value of zero if households have more expenditures than their total incomes, must not be primarily targeted for consumer credits. We demonstrate our household classification as follows;

Income Dummy for = 1 if (Annual Revenue – Expenditures) > 0 and Income Dummy = 0 if (Annual Revenue – Expenditures) < 0

We find that among the total sample of 25,586 households, the number of households falling to the low risk category is 14,132 and that of the other category is 11,454[6]. We compare the performances of logit analysis, neural network and decision tree methods to find out to what extend our diversifying methodology is successful in assigning householders to the correct categories. It is an imperative issue that interpretation of our multivariate

[6] We are not suprised with the high number of non-low risk category of households. Most of the households in this group is highly likely to be receiving irregular informal additional payments from their family members outside their households. However it is almost impossible to reflect this additional income to the survey since it is not documented, hence does not have an official character.

estimation results must be reflecting the accurate categories. Table 4 presents the results of the three classification methodologies.[7]

[Insert Table 4 about here]

The respective panels in the table are apportioned for three different methodologies and each of them demonstrates what percentages of householders were accurately assigned to the each category. Overall results show that we have apportioned the householders with more than 70 % accuracy. The overall accuracy level of decision tree (72.68 %) outruns those of the neural networks (71.54 %) and the logistic analysis (71.53 %). Moreover for the low risk category of households, the prediction percentage of decision tree (66.35 %) is higher than those of the neural networks (63.90 %) and logit analysis (62.91 %). Finally, for non-low risk category of households the accuracy level of logit analysis (78.51 %) is superior to those of the decision tree (77.81 %) and neural networks (77.73 %). Consequently the results demonstrated in Table 4 convince us with the accuracy of our classification.

Before moving to our multivariate analysis we compare the mean differences of the key variables between low risk households and the ones having a negative income-expense balance. Table 5 provides univariate mean comparisons of the variables.

[Insert Table 5 about here]

Our univariate results show us that most of the mean values of the two credit risk categories of the households display significant differences. Considering the mean values of the variables, relative to the non-low risk group, low risk households are significantly found to be on average; older, more likely to own a car, less likely to own a credit card, more

[7] We base on logistic regressions for our multivariate analyses however we include neural networks and decision tree methods to enrich our findings. Nevertheless, the results obtained through the latter methods align with that of the logistic analysis.

educated, living in a smaller house, more likely to be working in a executive or managerial level, have higher income per person and more likely to be a house owner. Besides when compared to the low risk group, non-low risk group of households are significantly found to be on average, more likely to have a second house but less likely to own a dwelling and have dwellings with relatively higher market value. However it must be noted that, no significant differences are obtained between the two groups of households for the mean values of the number of people in a household and the savings dummy.

Results for Turkish Households

Table 6 presents results of our binomial logit estimation for the households[8]. Our dependent variable in the regression is the dummy variable which takes the value of unity if a householder is classified as a good credit household and zero otherwise.

[Insert Table 6 about here]

Before interpreting the results, the overall fit of the model is assessed with several goodness of fit tests, such as the Nagelkerke test, McFadden test and Cox & Snell test. The significance of the model is reflected in the result of the likelihood test. Overall tests clarify the success of our model. While interpreting the impact of the variables we consider the absolute strength of their respective odds ratios.

First of all, size of a dwelling is found to be negatively related with being a low risk household. An odds ratio of 0.996 implies that an increase by one in the house size of a household reduces the odds of being defined as a low risk household by less than 1 %. Size of a dwelling is regarded as a prestige for a Turkish family therefore households are likely to push beyond their budget limit to own a relatively larger house. Therefore a relatively larger

[8] See, Westgaard and Van der Wijst (2001) for the superiority of the logistic model approach in determining credit risk.

dwelling does not signal that its occupying household is more likely to be qualified for a consumer credit.

Our results show that number of persons in a household is positively related to being a low risk household. This result aligns with the expectation that as the number of people in a household increase, their contribution to total household income is highly likely demonstrate a raise. Besides, an odds ratio of 0.845 represents that one category of increase in the number of persons in a household improves the risk profile by almost 15 %. However in our previous univariate test, the mean differences of the variable representing the number of persons in a household was not found to be significantly different between the two groups of households. The siginificant result in our multivariate estimation must be based on the non-linear character of the logistic regression.

Age of a household head is also positively related with being a low risk household. This result is in accord with our expectations since it is easier to detect the risk profile of relatively older consumers given that they have a longer past which allows a wider assessment on their professional activities and consumption patterns. For this reason, information asymmetry between supplier of credits and households reduce as household heads become older people. However, younger consumers have a shorter credit history thus they are more likely to impose higher credit risks.

As the total income per person in a household increases then it is found to be more likely to be a qualified for a consumer credit. Nevertheless, the odds ratio of 0.801 highlights that one level of increase in the total income per person improves the credit risk profile of households by almost 20 %.

The coefficient of the variable “log rent 2” illustrates that there is a strongly inverse relationship between being a low risk household and the total revenue obtained if the currently owned and resided dwelling was rented. The stated variable stands for the market

value of the currently resided dwellings of our sample households. Even the odds ratio of 0.358 indicate that a one per cent increase in the market value of the currently resided dwelling decreases the independent variable by almost 65 %. With this result, Turkish households are found to have an inclination towards living in relatively more expensive houses than their affording capacity in order to increase their prestige in the society. For this reason dwellings of consumers lose their ability to be evaluated as collaterals against their default risk.

Having a second house is also found to be negatively associated with being a low risk household. This result aligns with our finding in our univariate analysis. An extra house purchases of Turkish households are not mostly derived from a saving motivation. Instead, it is very common among Turkish households to purchase another dwelling in order to guarantee a housing for their children right after they move out. Therefore it is fairly likely for a household to lose the qualification for consumer credits if it aims to opt for owning a second house other than the currently resided one.

Furthermore, owning a car is highly strongly associated with having a low credit risk. The odds ratio of the variable is 0.267 and suggests that one category of increase in car ownership increases the independent variable by almost 75 %. Most of the cars in Turkey are purchased by households through consumer credits offered by banks. A prior detailed investigation of households by banks for vehicle credits provides an additional intelligence for financial institutions concerning the households before granting them consumer credits. For this reason, a more symmetric information established between the institutions and the previously inspected households.

However, having a credit card is found to be inversely related with having a priority for consumer credit qualification. Therefore our results suggest that the institutions must primarily target households which do not have a credit card yet. This results can be explained

by the fact that having multiple credit cards may weaken the ability of households to fulfill their outstanding debt obligations.

Education level of the household head is found to be an important concept for the risk structure of Turkish households. In other words our results show that as the education level of household head increases their risk profile recovers.

In accord with our expectations households that are able to make savings are found to be likely to have lower risk profile given that the savings dummy is negatively related with the dependent variable. Mean differences for the savings dummy between the two groups was not found to be significantly different in our previous univariate tests. However, the non-linear nature of our estimation justifies our significant finding. Nonetheless, the odds ratio of 0.615 means that one category increase in the savings dummy, which takes the value of unity if a household is not making any saving at all and zero otherwise, exacerbates the risk profile of a household by almost than 30 %.

Our next results are concerned with the job characteristics of the household heads. On the one hand being hired in an executive or managerial level, on the other hand being a monthly paid worker rather than a daily or seasonal one is found to be positively related with being a low credit risk household. The latter result highlights that receiving a salary per each month secures a stable income stream for a household whereas the other options display a unstable pattern of payment.

Last but not least, owning the currently occupied house is found to be positively affecting the independent variable. Monthly rent payments are expected to bring an extra burden in an event of a consumer credit therefore households that do not already own their house are not expected to be qualified for consumer credit at the first stage.

Conclusion

Our paper mainly aims to identify risk profile of Turkish householders so as to improve information symmetry in Turkish consumer credit market. Therefore we determine which risk factors contribute adversely or positively to the credit riskiness of Turkish households. We use the replies to *Budget Survey of Turkish Households,* which is conducted by *TURKSTAT* for the year 2003. The survey encompasses in depth questions concerning the socioeconomic and demographic characteristics of Turkish households. Our sample is comprised by 25,586 replies which are purposely chosen by TURKSTAT to reflect the general characteristics of Turkish households.

Through binomial logit estimations we separately test the risk factors affecting the households. We assign households that have a positive income-expense balance to low credit risk category totaling to the number of 14,132. These households are the ones which must be primarily targeted primarily by financial institutions for consumer credits. However, households that have a negative income–expense balance are classified as non-low credit risk category and they total to 11,454. This group of households are in the second line of the pecking order of the consumer credit supply, hence situated after the low risk group. Our neural networks, decision tree and logistic analyses confirm that each of the accuracy levels in allocating the households to the risk categories are around 70 % on average.

The following variables highlight the particularities of low risk households which should be initially targeted by financial institutions for their supply of consumer credits. Households with relatively larger dwellings or the ones having a second house or a credit card are not the ideal ones that should be attracted to consumer credits. Whereas, households including relatively more number of members, own their own dwelling, having higher level of income per person or owning a car is likely to increase the chance of being qualified for consumer credits at the first stage. However if, dwelling of a household has a relatively high

market value, households maintain less savings or the education level of household is relatively lower, then these are highly likely to be a non – low risk households. Finally if a household head is relatively older, works either in an executive or managerial position or receives a monthly wage rather than daily or seasonal salaries then that household is highly likely to be classified as a low risk one.

References

Avery, R. B., Calem, P. S. and Canner, G. B. (2004) "Consumer credit scoring: do situational circumstances matter? *Journal of Banking and Finance*, 28, 835-856.

Carling, K., Jacobson, T., Lindé, J. and Roszbach (2007) "Corporate credit risk modeling and the macroeconomy", *Journal of Banking and Finance*, 31, 845-868.

Central Bank of the Republic of Turkey (December, 2004) *Balance of payments report,* http://www.tcmb.gov.tr.

Cutts, A. C. Van Order, R. A. and Zor,. P. M. (2000) "Lemons with a twist: the role of the secondary market in market evolution." Paper presented at the 2000 annual ASSA/AREUEA Meetings, Boston. MA, January.

Ge. Y. and Qiu. J. (2007) "Financial development. bank discrimination and trade credit." *Journal of Banking and Finance*, 31, 513-530.

Getter, D. E. (2006) "Consumer credit risk and pricing", *Journal of Consumer Affairs*, 40 (1), 41-63.

Hazembuller, A., Lombardi, B. J., Hogarth, J. M. (2007) "Unlocking the risk based pricing puzzle: five keys to cutting credit card costs", *Consumer Interests Annual*, 53, 73-85.

Hofstede, G. (1980) "Culture's consequences". Newsbury Park. Sage: New York. 1st ed.

Hogarth, J. and Hilgert, M. A. (2002) " A profile of consumers with high-rate home loans", *Financial Counseling and Planning*, 13 (1), 15-24.

Hsieh, N. (2004) "An integrated data mining and behavioral scoring model for analyzing bank customers", *Expert Systems with Applications*, 27 (4), 623-633.

International Monetary Fund (November, 2007) *IMF Country Report*, No:07/364.

King, R. G. and Levine, R. (1993) "Finance and growth: Schumpeter might be right", *Quarterly Journal of Economics*, 108, 717-737.

Lyons, A. C. (2003) "How credit access has changed over time for US households", *Journal of Consumer Affairs*, 37 (2), 231-255.

Joos. P., Vanhoof. K., Ooghe, H. and Sierens, N. (1998) "Credit classification: a comparison of logit models and decision trees", in Nakharizadeh. G. et al (eds.) "Application of machine learning and data mining in finance" European *Conference on Machine Learning (ECML'98), Workshop Note, 24 April 1998*, 59- 70.

Rosenberg, E. and Gleit, A. (1994) "Quantitative methods in credit management: a survey" *Operations Research*, 42, 589-613.

Swain, R. B. (2007) "The Demand and Supply of Credit for Households", *Applied Economics*, (August), 1 – 12.

The Banks Association of Turkey (2007), http://www.tbb.org.tr

Thomas, L. C. (2000) "A survey of credit and behavioral scoring: forecasting financial risk of lending to consumers", *International Journal of Forecasting*, 16, 149-172.

West, D. (2000) "Neural network credit scoring models," *Computers and Operations Research*, 27, 1131-1152.

Westgaard, S. and Van der Wijst, N. (2001) "Default probabilities in a corporate bank portfolio: A logistic model approach", *European Journal of Operational Research*, 135, 338 – 349.

TABLE 1: Definitions of the Variables

VARIABLE	SHORT NAME	MEASUREMENT EXPLANATION
Age	**Age**	Age of the leader of household head
Car owner	**Car owner**	Dummy Variable (Yes=1, No=0)
Number of individuals in household	**Person**	Value
Credit card possession of household head	**Credit card**	Dummy Variable (Yes=1, No=0)
Education level of household head	**Edu**	Dummy Variable =1 if; no education, only literacy, five year primary school, eight year primary school, middle school, high school, vocational school. Dummy Variable = 0 if; bachelors degree and graduate degree.
Owning the currently resided dwelling	**House owner**	Dummy Variable (Owner of the dwelling=1, Rented the dwelling= 0)
Space of the house in meter square (m^2)	**Housesize**	Value
Ratio of total household income to the number of individuals in household	**log income per person**	Natural logarithm if the total household income per person
Owning a second house	**Sechouse**	Dummy Variable (Yes=0, No=1)
Job position of household head	**Poscode**	Dummy Variable= 1 if; executive level, mid-level manager and low-level manager. Dummy Variable= 0 if; clerical works in service sector, sales, agricultural sector, art sector, engineering, international organizations and any unskilled work.
Income position of household head	**Position**	Dummy Variable=1 if monthly paid worker. Dummy Variable=0 if; daily paid worker, seasonally paid worker and non wage family worker.
Household head's savings	**Savings**	Dummy Variable=1 if no savings. Dummy Variable=0 if; savings on real estate, residential estate, precious metals, foreign currency, saving deposits, stocks, bonds, investment funds, venture capitals, corporate lendings.
Total value of the revenue obtained if household had rented the currently owned and occupied dwelling	**log rent2**	Natural logarithm of the total revenue obtained if household had rented the currently owned and occupied dwelling

TABLE 2: Descriptive Statistics

Variables	Mean	Median	Standard Deviation	Minimum	Maximum
Age	46.82	45	13.66	16	90
Car owner	0.239	0	0.427	0	1
Person	4.156	4	1.966	1	15
Credit card	0.765	1	0.224	0	1
Edu	0.736	1	0.441	0	1
House size	100.101	100	25.287	25	650
log income per person	21.214	21.223	0.837	16.811	24.530
log rent2	14.207	17.910	7.548	0	21.416
Sechouse	0.944	1	0.229	0	1
Savings	0.184	0	0.388	0	1
Poscode	0.310	0	0.298	0	1
Position	0.342	1	0.474	0	1
House owner	0.219	0	0.413	0	1

Notes: This table presents descriptive statistics for 25,507 households. Descriptions of variables are presented in Table 1.

TABLE 3: Correlation Matrix

	Age	Car owner	Person	Credit card	Edu	House size	l. i. p. p.	log rent2	Sechouse	Savings	Poscode	Position	House owner
Age	1												
Car owner	-0.041	1											
Person	-0.143	-0.017	1										
Credit card	0.132	-0.343	0.081	1									
Edu	0.227	-0.265	0.147	0.438	1								
House size	0.001	0.206	0.096	-0.193	-0.186	1							
l. i. p. p.	0.138	0.344	-0.143	-0.397	-0.385	0.247	1						
log rent2	0.280	0.083	0.039	0.045	0.105	0.075	0.023	1					
Sechouse	-0.107	-0.163	0.044	0.115	0.095	-0.101	-0.209	-0.070	1				
Savings	0.023	0.184	-0.018	-0.151	-0.139	0.153	0.321	0.063	-0.119	1			
Poscode	-0.347	0.141	0.069	-0.243	-0.280	0.052	0.081	-0.234	0.038	0.054	1		
Position	-0.431	0.092	-0.021	-0.266	-0.032	0.009	0.077	-0.226	0.055	-0.020	0.421	1	
House owner	-0.283	-0.060	-0.057	-0.077	-0.136	-0.049	0.018	-0.416	0.055	-0.054	0.257	0.245	1

Notes: This table correlation coefficients across variables. Descriptions of variables are presented in Table 1. *l.l.p.i* stands for *log income per person*.

l

Table 4: Detailed Analyses for Good Credit Households

Panel A: Prediction Percentages with Decision Tree Analysis			
	Dummy Variable =0	Dummy Variable = 1	Prediction Percentage (%)
Dummy Variable = 0	10,996	3,136	77,81
Dummy Variable = 1	3,854	7,600	66.35
Overall Prediction (%)			72.68
Panel B: Prediction Percentages with Logit Analysis			
	Dummy Variable =0	Dummy Variable = 1	Prediction Percentage (%)
Dummy Variable = 0	11,095	3,037	78.51
Dummy Variable = 1	4,248	7,206	62.91
Overall Prediction (%)			71.53
Panel C: Prediction Percentages with Neural Networks			
	Dummy Variable =0	Dummy Variable = 1	Prediction Percentage (%)
Dummy Variable = 0	10,985	3,147	77.73
Dummy Variable = 1	4,135	7,319	63.90
Overall Prediction (%)			71.54

Notes: This table presents results for decision tree. neural networks and logistic analyses. Dummy Variable = 1 denotes low risk customers and Dummy Variable = 0 if otherwise.

TABLE 5: Univariate Mean Comparisons

Variables	Non – Low Risk Households	Low Risk Households	t-test
Age	46.593	47.067	-2.65**
	n = 11,454	n = 14,132	
Car owner	0.218	0.268	-8.964***
	n = 11,454	n = 14,132	
Person	4.156	4.175	-0.707
	n = 11,454	n = 14,132	
Credit card	1.798	1.725	12.959***
	n = 11,454	n = 14,132	
Edu	0.701	0.763	-10.725***
	n = 11,454	n = 14,132	
House size	10.363	9.181	6.504***
	n = 11,454	n = 14,132	
log income per person	20.947	21.529	-6.110***
	n = 11,454	n = 14,132	
log rent2	14.620	13.686	9.344***
	n = 11,454	n = 14,132	
Sechouse	0.952	0.933	5.973***
	n = 11,454	n = 14,132	
Savings	0.278	0.116	3.484
	n = 11,454	n = 14,132	
Poscode	0.030	0.590	-6.283***
	n = 11,454	n = 14,132	
Position	0.226	0.510	0.981
	n = 11,454	n = 14,132	
House owner	0.198	0.244	-3.365***
	n = 11,454	n = 14,132	

Notes: This table provides univariate mean comparisons of the variables. Non – Low Risk households are the ones that have negative income-expense balance. Whereas Low-Risk Households are the ones that have a positive income-expense balance stands for the number of observations. The *t*-statistic is for the difference of means between the first and the second credit risk classifications. Definitions for all the variables are provided in Table 1. *** and ** indicate coefficient is significant at the 1%, 5% and 10% level respectively.

Table 6: Binomial Logit Estimation

Variables	Coefficient	Std. Error	Wald	P-value	Odds Ratio
Age	0.005	0.001	11.786	0.001	0.995
Car owner	1.322	0.374	12.491	0.000	0.267
Person	0.169	0.095	3.188	0.074	0.845
Credit card	-0.210	0.041	26.017	0.000	0.811
Edu	-0.196	0.040	23.725	0.000	0.822
House size	-0.004	0.001	44.975	0.000	0.996
log income per person	0.221	0.062	12.788	0.000	0.801
log rent2	-1.027	0.027	1473.507	0.000	0.358
Sechouse	-0.176	0.066	7.023	0.008	0.839
Savings	-0.486	0.040	144.974	0.000	0.615
Poscode	0.351	0.044	63.746	0.000	0.704
Position	0.187	0.043	19.253	0.000	1.205
House owner	0.346	0.105	10.784	0.001	1.414
Goodness of fit tests	**Value**	**p-value**			
Cox & Snell - R^2	39.2 %	n/a			
Nagelkerke- R^2	42.13 %	n/a			
McFadden - R^2	45. 81%	n/a			
-2 Log Likelihood	2742	0.000			

Notes: Definitions of the variables are presented in Table 1