Eurasian Journal of Business and Management, 4(1), 2016, 11-22
DOI: 10.15604/ejbm.2016.04.01.002

EURASIAN JOURNAL OF BUSINESS AND MANAGEMENT

http://www.eurasianpublications.com

WHAT TRIGGERS LOAN REPAYMENT FAILURE OF CONSUMER LOANS – EVIDENCE FROM BOSNIA AND HERZEGOVINA

Sanela Pasic
Corresponding Author: Sarajevo School of Science and Technology, Bosnia and Herzegovina.
Email: sanela.pasic@ssst.edu.ba

Adisa Omerbegovic Arapovic
Sarajevo School of Science and Technology, Bosnia and Herzegovina.
Email: adisa.omerbegovic@ssst.edu.ba

Abstract

This research explores most dominant lending product to population of Bosnia and Herzegovina, a consumer loan, with aim to answer the question of what factors trigger loan repayment failure. It explores relation of borrower characteristics such as gender, age, level of indebtness to likeliness of loan repayment by use of probit on banking data sample representing 39% of the market share in the country. It identifies factors which lead to loan repayment failure and also provides exact empirical model for default prediction at loan approval stage. Main audience of this research should be banks, which could use the finding of the study to adjust their credit policies and risk appetite to ensure that lending losses from this strongly present product are minimized, thus leading to stable and financially sound banking sector.

Keywords: Consumer Loans, Default Prediction, Probit Regression, Credit Risk

1. Introduction

Each loan approved by a bank is at risk of not being repaid or risk to default on agreed repayment schedule. Given that default will result in financial loss to the banks, assessing the default risk at the time of approval is crucial. This would involve observing lessons learnt in the past in order to try to predict whether a borrower with particular credit standing and features is likely to default on a loan, as suggested by Mays (2001), Thomas *et al.* (2002), and Anderson (2007). Although default prediction has been researched for more than seventy years, starting with Durand (1941), it remains an area that requires further academic and business research.

2. Research Scope and Purpose

This research covers the area of consumer lending by commercial banks in Bosnia and Herzegovina. The specific research question that we will try to answer through the research is "what are the factors that lead to failure in the payment of consumer loans (triggers of default) within the domestic market?"

The banking market population in Bosnia and Herzegovina has a size of 7.06 billion Convertible Marks (Central Bank of Bosnia and Herzegovina, 2013). The population is indebted to 26.7% of GDP. This is below EU levels of indebtedness, namely EURO ‘A’ zone countries at 67.1%, the surrounding countries (CEE) at 29.7% and SEE at 33.8% (European Central Bank, 2013).

Lending to the population is an important and growing banking business line in Bosnia and Herzegovina as it is a country where consumer loans (simple equal annuity loans without mortgage) dominate amongst the products offered to the population. Of the total population lending business, 58% percent is placed in the form of consumer loans and the remaining portion consists of mortgage loans, car loans and cards (Central Bank of Bosnia and Herzegovina, 2013). Given the fact that these loans represent the main product for banks, the quality of consumer loans is a significant driver of profitability and stability of the banking sector. In the event that the repayment rate for such loans was low then this would result in a high proportion of nonperforming loans and excessive risk impairment costs for the banks. Uncontrolled losses on consumer loans as a result of a low repayment rate would therefore lead to the destabilization of the banking sector. This is why it is important to understand this lending segment and the risks associated with repayment and to ensure that losses are minimized through implementation of prudent risk management practices.

3. Research Hypotheses

Hypothesis for research are derived from expert opinion of author, resulting from decade of practical work in BiH banking sector on approval of loans and from literature review findings. There are three:

First hypothesis (H1) is that there is a relation between borrower characteristics and the loan repayment outcome. It is possible to predict whether a loan will be repaid by analyzing the borrower characteristics. This hypothesis will be confirmed by the actual development of a stable logistic model (the statistical features of the model will demonstrate high stability if a non linear relation exists, as is the assumption).

Second hypothesis (H2) is that higher indebtedness compared to earnings will decrease a borrower’s chances of repaying a loan. This hypothesis will be proven in the model by determining a strong relationship between the independent variable of indebtedness and the dependant variable of loan default.

Finally, third hypothesis (H3) is that borrower characteristics, such as gender, level of education and marital status and location of residence influence the likelihood of loan repayment. This hypothesis will be proven by determining strong marginal effects of the independent variables on the final dependent variable of loan repayment. If the research proves this hypothesis to be untrue then other factors that are more important to loan repayment will be identified and stated in the conclusions of the research.

4. Literature Review

Literature review identified number of papers written both in area of methodology assessment used for default prediction studies and in the area of research of factors leading to default.

In terms of methodology, literature review showed techniques used for the determination of probability of default are numerous. There are conventional (parametric) statistical techniques, such as discriminate analysis, linear regression and logistic regression, and there are also more advanced unconventional (nonparametric) methods, such as neural networks, support vector machines, decision trees and expert systems. Several authors, such as Rosenberg and Gleit (1994), Hand and Henley (1997), Oliver and Hand (2005), Sabato (2010), Abdou and Pointon (2011) and Genriha and Voronova (2012), have written papers that are similar in content. All contain a literature review of earlier papers in this area and descriptions of the techniques as well as their comparison. A general conclusion of these papers is that ultimately there is no best technique. The use of a particular technique depends on the problem to be researched: the size and structure of the data sample, the cleanness of the data, the variables used to describe the consumer taking the loan, the extent to which it is possible to classify good and bad samples, the research objective and the availability of time and money for software application. Many authors have found nonparametric methods to have higher accuracy but at the same time that they are expensive and complicated to use. Parametric methods are more conventional and have been used in numerous researches over the past decades. They have a slightly lower level of accuracy but one that is still sufficiently high and they are easier and cheaper to conduct. Amongst them, over the past two decades logistic regression (logit and probit) has been stated as having the highest level of accuracy and reliability. It is more convenient to use parametric methods, via widely available software, for non-institutional researcher with limited resources. More specifically, probit seems to be highly accurate and the most frequently used in recent history.

In terms of specific researches on factors of default, number of authors, such as Mavri *et al.* (2008), Lina *et al.* (2009), Elul *et al.* (2010), Kocenda and Vojtek (2011), Riley (2013), Agarwal *et al.* (2012) have written recently on the topic of the probability of default within the banking portfolio of consumers in Asia and CEE, with the research question similar to the author's research question: What triggers the default of consumer loans in emerging markets?

In terms of conclusions on the factors that influence the likelihood of loan repayment, the review has shown that most authors arrived at similar conclusions. For example, studies show that some of the factors that result in higher default are the male gender of the borrower, the marital status 'single', lower education, higher job positioning, customers using a friend as opposed to a family member as a guarantor, bad credit history and a short relationship with the bank. Furthermore, the review shows that loan contract features that result in higher risk of default are smaller loan values, a high loan to value ratio (if the loan is very high compared to the value of the collateral) and collateral that is not the customer's private residence or collateral that is not located in an urban area.

Literature review has also shown that there are few behavioral studies on private individual portfolios in emerging markets, especially in the area of consumer loans. Researchers seem to focus on lending to legal entities, or in area of private individual lending, they focus on products such as credit cards and mortgage loans, while region wise they concentrate on Europe and the USA. Therefore, it seems that further research into the emerging markets, particularly on the portfolio of consumer loans, will provide added value to academic and business knowledge on this topic.

Lastly, the literature review shows that no similar research on private lending has been conducted in Bosnia and Herzegovina and that consumer loans have not been researched at all for their delinquency. The only related work published in Bosnia and Herzegovina is a PhD paper written by Memic and Rovcanin (2012) in the area of default triggers for legal entities (companies).

In view of the conclusions of the literature review the question remains of how the present research fits into the academic and business framework on this topic. In relation to the array of similar works mentioned in the above sections, the attempt of the authors is to make a contribution by covering the specifics of the market in Bosnia and Herzegovina. This work can make a contribution to science through its exploration of consumer default through a new sample on the unexplored market and hopefully bring its results into line with those of earlier studies. Since earlier researchers focused mainly on mortgage rather than plain consumer loans with no physical collateral, this study will expand on knowledge by focusing on consumer loan default. Moreover, this work will contribute to the business requirements of banks in Bosnia and Herzegovina by identifying those factors that are or are not relevant to loan default. By using these factors in the form of permission or restriction the banks will be able to define their credit policies and therefore minimize their default losses, increase profit and contribute to the stability of the financial sector.

The intention is that this research will be relevant in terms of its focus on the most significant banking product in the Bosnia and Herzegovinian market, through the significant size of the local market sample and through its use of one of the most relevant methods (probit) and observation of the same and extended factors used by other researchers in this field.

5. Methodology and Data

The research area is consumer lending by commercial banks in Bosnia and Herzegovina. The specific question this research will try to answer is “what are the factors that lead to failure in the payment of consumer loans (triggers of default) within the domestic market?”

In order to answer this question one should analyze the existing consumer loans in Bosnia and Herzegovina (still under repayment), observe their characteristics and try to find a relation between one or more of these characteristics and the quality of loan repayment. Absence of quality of repayment will be defined as the default status and apply to loans where repayment was overdue by more than ninety days at any moment during the observation period. This implies that the research will observe a number of independent variables, all characteristics of individual loans and borrowers, which can be either binary (married or not married), continuous (age for example) or of a fixed value (monthly income). The research only observes one dependent variable and that is whether or not a loan has gone into default (a binary expression of this variable is most suitable). On the basis of the earlier researches and the fact that the statistical analysis will observe a number of independent variables (binary, continuous or fixed) but just one time series and one binary dependent variable (Finney, 1952; Stock and Watson, 2007), it was concluded that the most appropriate method to use for this research was logistic regression in the form of probit. Software used for modeling was Stata (support for Stata use generated from Long and Freese, 2006; Stock and Watson, 2007; Acock, 2008; and Katchova, 2013).

Defined type of data required for this research is information on individual loans, including borrower and loan characteristics and information on whether the loan is being repaid on a timely basis or not. The sample for the research was obtained directly from individual banks, seeking real time data on clients. In order not to miss any important data the decision was made not to specify a set of variables but rather to ask the banks to deliver all of the data contained in their customer loan systems. Of the banks available in the market data was requested from eight larger and more developed banks; these banks hold a 70% share of the retail lending market. A cut-off date of 31 December 2014 was set for the banks to deliver the available data, which was intended to provide a snapshot of all loans approved between 1 January 2010 and 31 December 2012. This means that each loan in the data set sample was observed for possible delay and default for a minimum of two years for those approved at end of 2012 and for up to four years for those approved at the beginning of 2010.

In order to answer the research question, the aim was to obtain as large a sample as possible. The research targeted and obtained 201,653 loans representing a loan portfolio of approximately 2 billion BAM, which accounted for 50% of the total consumer loans portfolio. The final size of the sample was 154,154 loan observations, loan portfolio representing 39% of market with 26 independent variables and one dependent variable, after the data had been cleaned and aligned into a database with the same structure of attributes. Initial variables kept in final model described following factors related to the client: loan maturity, number of guarantors, approved loan amount, monthly annuity amount, whether salary is kept at account with bank, credit history of client, age, gender, education, marital status, number of dependents, whether client owns his own residence, how long is client living at current address, how long has client been generally employed and employed at current job, type of employment, list of incomes, length of relationship with bank, statistics on first and last application for a loan, total loan exposure, history on payment (in days overdue in observation period), relation of debt to income on monthly basis.

6. Model Results

Probit modelling on final research sample showed following results in Table 1:

Table 1. Probit modeling on data pool

	(1)
Variables	**Probit coefficient**
Number of guarantors	-1.087***
	(0.0238)
Salary payment bank account	-0.587***
	(0.0189)
Age	-0.104***
	(0.00550)
Gender female	-0.242***
	(0.0202)
Marital status ‘M’	-0.127***
	(0.0219)
Dependants	0.0670***
	(0.00992)
Residential status	-0.214***
	(0.0465)
Employment type	0.171***
	(0.0318)
First application	0.0797***
	(0.00900)
Max days overdue 1	0.00265***
	(7.55e-05)
DTI	1.695***
	(0.0564)
Constant	-1.398***
	(0.0728)
Observations	154,154

Notes: Standard errors in parentheses. *** signifies p<0.01, ** signifies p<0.05, and * signifies p<0.1.

Table 2. Statistical results of probit model

Statistical results	
Pseudo R^2	0.3573
Prob>chi2	0.0000
Pearson chi2(136,581)	5853785.50
Hosmer - Lemeshowchi2(8)	82.15

As Table 2 demonstrates, the pseudo R^2 of the model is 0.36 (McFadden’s). As the model is higher than 0.25 this indicates that the independent variables have good explanatory power regarding the dependent variable. In other words, as the characteristics of the client and loan provide a good basis to explain the likelihood of loan default the models fits well (as suggested by Veall and Zimmermann, 1996). Both the Chi Squared and Hosmer-Lemeshow tests showed non-significance on the goodness of fit (Hosmer and Lemeshow, 1980; 2000), indicating model prediction that is not significantly different from the observed values. Thus, these tests show that the model fits well. The model showed 98.16% correctly classified observations. Good cases in the sample amounted to 98.3%; therefore, the model appears to be well fitted, since these two figures do not deviate.

The area under ROC curve is 92%, it is concave and yields a point in the upper left corner of the ROC space representing high sensitivity (no false negatives) and high specificity (no false positives), as shown in Figure 1. Therefore, the classification of the model is good and the power to predict is fairly high.

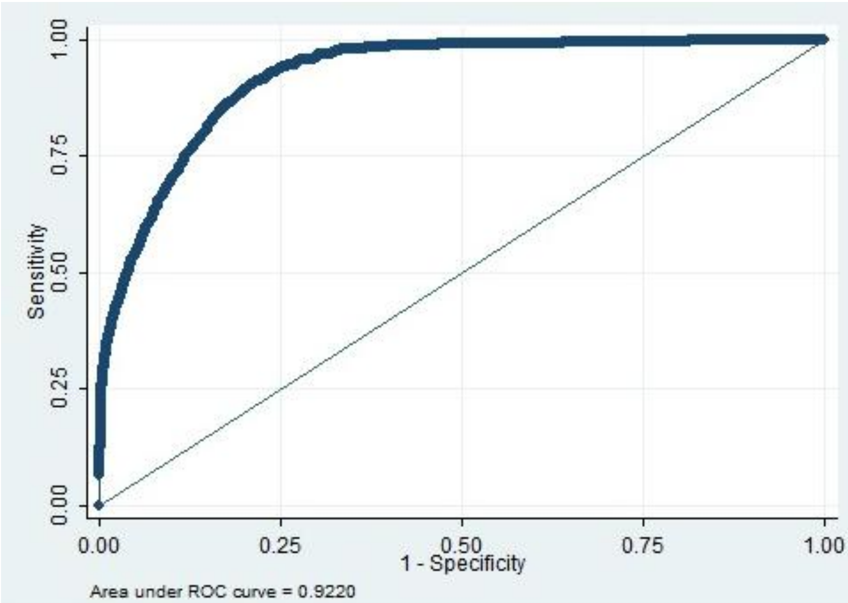

Figure 1. ROC curve

A cross validation was performed in order to estimate the performance of the model building algorithm or the predictive power. Cross validation in STATA was done using the technique of K fold cross validation. This consisted of fitting the model five times and each time leaving out one-fifth of the observations and then comparing the tested sets. Therefore, no new data was introduced to the model but rather cross validation was done internally on the existing observations of the main model.

The Relative Mean Square Error (RMSE) of the main model was 10% (RMSE in individual variables being 10 or smaller). Cross validation showed that the RMSE of each of the subsets ranged from 11.4% to 11.6%, implying that error was rather the small. The average was also the same for all subsets and very close to the error in main model. The Mean Absolute Error (MAE) ranged for all subsets between 2.5% and 2.6%, which implies rather small error. The same error was also for all subsets and smaller than the RMSE, as it should be. Both the RMSE and MAE were less than half of the standard deviation for all of the measured data, which speaks further in favour of model validation. The Pseudo-R^2 of the subsets ranged from 18.9% to 20.8%. This was lower than the Pseudo-R^2 for the total model (34%), but this is understandable due to the subsets being smaller. However, this distribution of Pseudo-R^2 implies that they are similar in all subsets: none of the subsets showed deviation from each other or the main model. To conclude, since all independently created subsets had similar explanatory power and error the cross validation confirmed the model fit.

In terms of the effects of the independent variables on the dependent variable, the model resulted in the coefficients and marginal effects shown below in Table 3. Marginal effects of independent variables seem to be the most suitable way to understand results of probit analysis, as suggested by Hand and Henley (1997), Norton and Wang (2004), and Katchova (2013).

Table 3. Marginal effects results

	(1)
Variables	**Marginal effects**
No guarantors	-0.0365***
	(0.000928)
Salary payment bank account	-0.0197***
	(0.000681)
Age	-0.00350***
	(0.000190)
Gender female	-0.00812***
	(0.000686)
Marital status ‘M’	-0.00428***
	(0.000737)
Dependents	0.00225***
	(0.000335)
Residential status	-0.00719***
	(0.00156)
Employment type	0.00573***
	(0.00107)
First application	0.00268***
	(0.000304)
Max days overdue 1	8.89e-05***
	(2.64e-06)
DTI	0.0569***
	(0.00202)
Observations	154,154

Notes: Standard errors in parentheses. *** signifies p<0.01, ** signifies p<0.05, * signifies p<0.1.

7. Conclusion and Result Discussion

7.1. H1 Conclusion

We consider first hypothesis confirmed. It is possible to predict whether loan will be repaid by analysing characteristics of borrower. Relation of characteristics to loan repayment is nonlinear. Confirmation comes from fact nonlinear relation between characteristics of the customer and default of the loan was determined, as Figure 2 shows. Also confirmation comes from fact statistical model to explain relationship was developed and it is stable and fit. This is manifested by relatively strong R square of 35.7%, fact all variables are individually and jointly significant, showing no correlation, both CHI Squared and Hosmer-Lemershow tests showed non-significance on goodness of fit, area under ROC curve was large – 92% and 98.16% loans were correctly classified/predicted, which corresponds with presence of non-defaults in main sample.

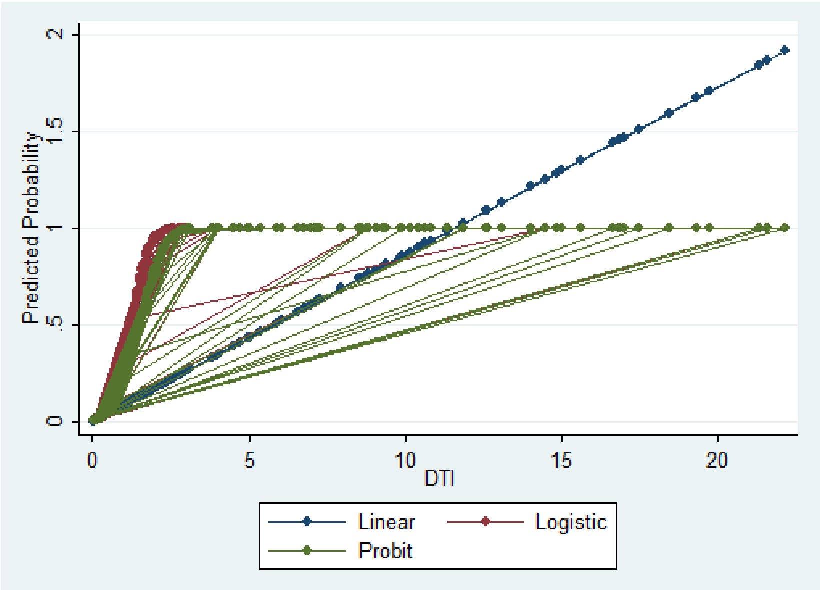

Figure 2. Linear versus logistic spread of relation of model variables

7.2. H2 Conclusion

We consider this hypothesis confirmed. Thus higher indebtness of borrower compared to earnings will decrease chances of repayment of loan. This is confirmed by fact customer indebtness to income, measured by variable DTI (debt to income, monthly annuities divided by monthly income), is significant variable in the final model, showing positive marginal effect to default 0.0569, implying that with each percentage increase od DTI, likeliness of default increases by 5.6%. Further to this, a test was conducted to calculate how increase of DTI from 10% to 100% affects likeliness of default. It showed Customer whose full salary is used to repay loan is 8.7% likely to default, compared to customers whose 10% of salary are used for loan repayment with likeliness of default of 0.19%.

7.3. H3 Conclusion

We consider this hypothesis partially confirmed and partially rejected. Some borrower characteristics such as gender and marital status are confirmed as hypothesized, to be important factor influencing loan repayment. However, level of education and location of living seem not to have any impact on loan repayment.

Figure 3 bellow shows that from total of twenty six characteristics analyzed, only eleven affect repayment capacity. Biggest drivers are level of indebtness (DTI), primary relation with the bank (holding salary payment at the bank) and number of guarantors.

Among them, highest impact comes from DTI, whether customer has guarantors and whether he keeps salary at bank. DTI has positive marginal effect to default of 0.0569. This means client with higher DTI (higher indebtness to income) has higher chances to default. With each unit increase in DTI, chances of default increase by 5.7%.

Number of guarantors has negative marginal effect of 0.0365. This means client with higher number of guarantors is less likely to default on loan. Increase of guarantors by one will decrease likeliness of default by 3.6%. This probably comes from fact that guarantors support payment when client has difficulties, or for moral reasons, client does not want guarantors to be contacted by bank and makes sure repayment is timely.

Salary payment with bank has negative marginal effect of 0.0197. This means client who has salary with bank is 1.97% less likely to have default on the loan.

Remaining marginal effects for other variables are explained in the same manner as for main three and listed in the table with marginal effect results.

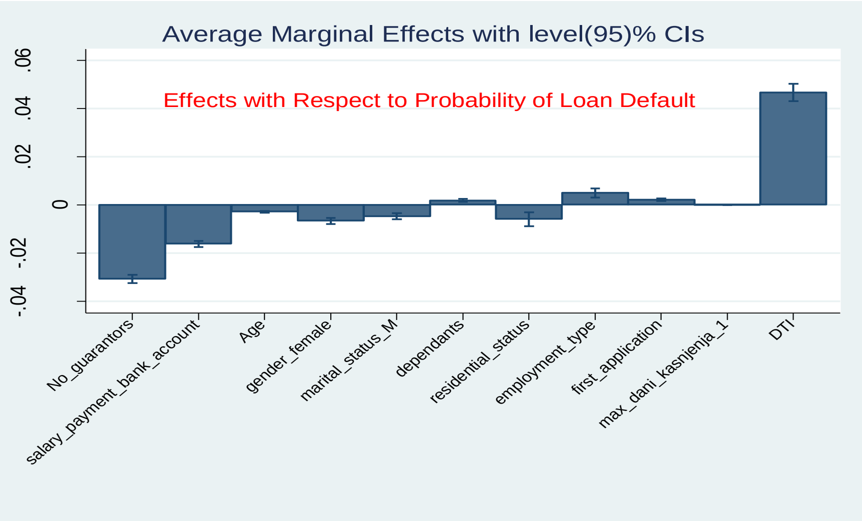

Figure 3. Marginal effects of variables in final model to default

8. Implications of the Research
8.1. Implications for Business

Results of the research are the most significant to banks in BiH, in other words, banks are the main audience of this study. Direct implications/results of this research for banks are:

8.1.1. The Actual List of Factors Which Strongly Lead to Loan Repayment Failures and Those Which Are Not Relevant or Have No Impact

Factors which increase risk of failure in repayment of loan are higher indebtness (DTI), not having salary with bank, not having guarantors, not having permanent employment, frequent usage of loans, having dependents in household, and overdues in installment repayment. Positive influence, which reduce likeliness of default, will have client with higher age, female gender, being married, owning residence. Factors which are not relevant for repayment of loans are level of education, region of residence, industry of employment, loan maturity and amount, monthly annuity and sum of incomes, history of repayment in other banks, time spent living at current location or working on current job, overall time of employment, length of time in relation with this bank, time since last application for loan to a bank, current overdue days.

Implication of this finding is that bank can alter its credit policy accordingly. This means, that credit policy should only introduce certain restrictions to the factors identified as strongly leading to default, either completely ask for rejection of those clients or discourage it at some level. Process of managing default is about optimizing risk and reward. This implies that banks do not need to reject clients for which there is some likeliness of default shown by these factors. On contrary, banks should determine acceptable level of risk, perhaps it is a current one, perhaps a lower or higher one, and compare likeliness of default shown by this research to a desired level of default. By doing this, they would manage their risk and reward optimum better. This implies credit policies should be structured in the manner to observe all these factors in relation to desired level of risk appetite. Furthermore, banks can apply risk based pricing based on this model and create appropriate reward from more or less risky clients. Therefore, this implies banks should use knowledge of factors to create their own strategy and risk appetite optimum.

8.1.2. Logistic Prediction Model

Banks can use actual logistic model defined in this research for default prediction, meaning to calculate whether some applicant will or will not fail in repayment at the time of request for a loan. Final model is as follows:

Default 90 dpd = -1.087 *no of guarantors –‘0.5587 *salary payment with bank -0.104*Age – 0.242*gender female – 0.127 *marital status + 0.067*dependants -0.214*residential status +0.171 *employment type + 0.0797 *first application + 0.0026 *max days overdue + 1.695*DTI

Use of this model would be at approval stage, when client applies for a loan. Bank would need to ask for these eleven factors and insert them into this model formula. Formula would deliver binary result of “1”, if that particular client is expected to default on the loan or “0” if that particular client is not expected to default. Results do not have to be rounded, but decimal and bank can set up a cut off for accepting default expectation (for example, for all clients having result higher than 0.60, we will conclude default is highly expected). For those clients where default is highly expected, bank should reject them a loan and move to a different customer.

8.1.3. Using the Profile of Best Customer

Best customer (least risky for a bank) is mid aged client, female, with family and residence, who keeps salary at bank and has low DTI (best bellow 35%). This information can be used to target these customers with more loans, even to pre-approve them, since it is concluded in the study that these customers bear low risk. Variations can be made in pre-selection from customer database, by altering one or more of these factors, for example, banks can draw product offer for customers with age of 20 for housing, customer of age of 30 for renovation, etc.

This would enable banks to do acquiring cheaper – by data analyses, rather than waiting for walk in customers. Also risk cost of such loans would be lower, so at the end implication for bank is cheaper lending, higher profit and bigger stability.

8.1.4. Changed Data Requirement to Clients

Banks tend to request from clients a lot of data. Various data sheets are filled and repetitive requests for additional information sent to clients in order to approve loan. Consequence of this practice on bank side is timely and more expensive loan processing. On client side, this process is frustrating and often leads to decision not to take loan or to take loan with competing bank. Implication of this research with respect to data requirement is that only eleven type of information matters (our independent variables) and thus bank should reduce data requirement only to this information. By doing this, bank will reduce paperwork for client, reduce time of loan processing, have better service, yet same level of risk, since all other data currently sought is redundant from perspective of risk of default. This will also lead to cheaper operations, therefore higher profit and bigger bank stability.

9. Implications to Science

The findings of this research were in line with findings of similar researches in other countries. It confirmed that relation between customer characteristics and loan default can be identified and that is nonlinear. It confirmed that lower indebtness, married status, higher age, gender female and owning residence lead to lower default. However, some findings were different from other studies. For example this study implied information on historical repayment of loans, length of relationship with banks and level of education are not relevant for loan default whereas some other researchers identified these factors as relevant.

In the array of similar works, this research has a biggest contribution to the science by providing analysis of the consumer default on a new sample from unexplored market. Therefore, it may observe same factors and use same methods, but it is unique due to the fact it observes only market in Bosnia and Herzegovina.

Research is also unique in having a bigger sample than any of the other researchers conducted in this area, which should make its results credible. Typically, similar researches in other countries or regions have contained sample of 3 to 10percent of the market, where as this research has covered with its sample 39 percent of the local market. It is conducted with real time exact banking data and with use of probit, which seems to be most relevant method for this type of research, which should further compliment it s credibility.

This research has indicated several areas of interest for further research, which would take this research to a new level of quality and removing its limitation. Following areas are recommended for further research: a) effect of increased indebtness in the country to loan default (to expand data pool to 2015 where indebtness of population to GDP is higher than in previous years), effect of job positioning (manager or junior associate) in relation to default, effect of type of guarantor to default (whether a friend or a family), effect of time since first application, effect of industry of employment to default.

10. Limitations

To our knowledge, this research has several limitations. The first one is the fact it is orientated specifically to the market of Bosnia and Herzegovina and can be used by scholars and banks on this market best.

The second limitation is that *reject interference* is not used in research. This means that data on already rejected clients was not part of main research sample and main probit model. This reduces the learning on the factors which really influence the repayment failure.

Last, but possibly biggest limitation of research may be the fact that sample will contain too small number of bad loans or defaults, namely 5%. Small number of bad loans may lead to lack of model stability and affect conclusions of probit. This potentially affects level of certainty with which results and conclusions can be claimed, thus the reader is warn of possible problems in this area. However, real level of defaulted portfolio in the country is this low, according to statistics of Banking Agencies for Federation and RS, so this is as far as bad rate could be represented in the research.

It is our belief none of these limitations reduces the importance and quality of findings significantly and that conclusions fairly represent the market behaviour.

References

Abdou, H. and Pointon, J., 2011. Intelligent systems in accounting, finance and management. *Wiley Online Library*, [online] Available at: <http://onlinelibrary.wiley.com/doi/10.1002/isaf.325/full> [Accessed 3 December 2015].

Acock, A., 2008. *A gentle introduction to stata*. Second edition. Texas: Stata Press.

Agarwal, S., Ambrose, B., Chomsisengphet, S., and Sanders, A., 2012. *Thy neighbor's mortgage: Does living in a subprime neighborhood affect one's probability of default. Real Estate Economics,* 40, pp.1-22. http://dx.doi.org/10.1111/j.1540-6229.2011.00311.x

Anderson, R., 2007. *The credit scoring toolkit: Theory and practice for retail credit risk management and decision automation*. New York: Oxford University Press.

Central Bank of Bosnia and Herzegovina, 2013. *Publications/annual report 2013*. [online] Available at: <www.cbbh.ba> [Accessed 14 December 2014].

Durand, D., 1941. *Risk elements in consumer instalment financing, studies in consumer instalment financing*. New York: National Bureau of Economic Research.

Elul, R., Souleles, N.S., Chomsisengphet, S., Glennon, D., and Hunt, R., 2010. Mortgage market and the financial crisis: What "triggers" mortgage default?. *American Economic*

Review: Papers & Proceedings, 100, pp.490-494. http://dx.doi.org/10.1257/aer.100.2.490

European Central Bank, 2013. *Statistics/consolidated banking data 2013*. [online] Available at: <https://www.ecb.europa.eu> [Accessed 14 December 2014].

Finney, P.J., 1952. *Probit analysis*. Cambridge: Cambridge University Press.

Genriha, I. and Voronova, I., 2012. *Methods for evaluating the creditworthiness of borrowers*. *Economics and Business*, 22, pp.42-49.

Hand, D.J. and Henley, W.E., 1997. Statistical classification methods in consumer credit scoring: A review. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 160(3), pp.523-541. http://dx.doi.org/10.1111/j.1467-985X.1997.00078.x

Hosmer, D.W. and Lemeshow, S., 1980. A goodness-of-fit test for the multiple logistic regression model. *Communications in Statistic*, 9(10), pp.1043-1069. http://dx.doi.org/10.1080/03610928008827941

Hosmer, D.W. and Lemeshow, S., 2000. *Applied logistic regression.* New York: Wiley. http://dx.doi.org/10.1002/0471722146

Katchova, A., 2013. *Series of video lectures on stata and logistic regression*. [video online] Available at: <www.youtube.com/user/econometricsacademy> [Accessed 15 March 2015].

Kocenda, E. and Vojtek, M., 2011. Default predictors in retail credit scoring: Evidence from Czech banking data. *Emerging Markets Finance & Trade*, 47(6), pp.80-98. http://dx.doi.org/10.2753/REE1540-496X470605

Lina, T., Leeb, C., and Chena, C., 2009. *Impacts of the borrower's attributes, loan contract contents, and collateral characteristics on mortgage loan default*. Taiwan: Department of International Business Publication, National Dong Hwa University.

Long, J.S. and Freese, J., 2006. *Regression models for categorical dependent variables using stata*. Second edition. Texas: Stata Press

Mavri, M., Angelis, V., Ioannou, G., and Gaki, E., 2008. A two-stage dynamic credit scoring model, based on customers profile and time horizon. *Journal of Financial Services Marketing*, 13, pp.17-27. http://dx.doi.org/10.1057/fsm.2008.2

Mays, E., 2001. *Handbook of credit scoring*. Chicago:Glenlake Publishing Company.

Memic, D. and Rovcanin, A., 2012. On the main financial predictors of credit default: Evidence from Federation of Bosnia and Herzegovina. *Our Economy*, 58(1-2), pp.3-15.

Norton, E.C. and Wang, H., 2004. Computing interaction effects and standards errors in logit and probit models. *Stata Journal*, 4(2), pp.154-167.

Oliver, R.W. and Hand, D.J., 2005. Survey of the issues in consumer credit modelling research. *Journal of the Operational Research Society*, 56, pp.1006-1015. http://dx.doi.org/10.1057/palgrave.jors.2602018

Riley, S.F., 2013. Strategic default behavior and attitudes among low-income homeowners. *Real Estate Finance*, 29(5), pp.22-36.

Rosenberg, E. and Gleit, A., 1994. Quantitative methods in credit management: A survey. *Operations Research*, 42(4), pp.589-613. http://dx.doi.org/10.1287/opre.42.4.589

Sabato, G., 2010. Credit risk scoring models. *SSRN Electronic Journal*, [online] Available at: <http://ssrn.com/abstract=1546347> [Accessed 3 December, 2015].

Stock, J.H. and Watson, M.W., 2007. *Introduction to econometrics*. London: Pearson Education, Inc.

Thomas, L.C., Edelman, D.B., and Crook, L.N., 2002. *Credit scoring and its applications.* Philadelphia: Society for Industrial and Applied Mathematics. http://dx.doi.org/10.1137/1.9780898718317

Veall, M.R. and Zimmermann, K.F., 1996. Pseudo-R2 measures for some common limited dependent variable models. *Journal of Economic Surveys*, 10(3), pp.241-259. http://dx.doi.org/10.1111/j.1467-6419.1996.tb00013.x