Soft Data Modeling via Type 2 Fuzzy Distributions for Corporate Credit Risk Assessment in Commercial Banking

Brkic, Sabina and Hodzic, Migdat and Dzanic, Enis

American University in Bosnia and Herzegovina, International University Sarajevo, Univerzitet u Bihaću

15 June 2018

Sabina Brkić, MBA

American University in Bosnia and Herzegovina,

MijeKeroševićaGuje 3, 75000 Tuzla

prof. MigdatHodžić, PhD

International University Sarajevo,

Hrasničkacesta 15, 71000 Sarajevo

assoc. prof. EnisDžanić, PhD

Univerzitet u Bihaću

Ivana Pavla II b.b., 77000 Bihać

SOFT DATA MODELING VIA TYPE 2 FUZZY DISTRIBUTIONS FOR CORPORATE CREDIT RISK ASSESSMENT IN COMMERCIAL BANKING

Abstract The work reported in this paper aims to present possibility distribution model of soft data used for corporate client credit risk assessment in commercial banking by applying Type 2 fuzzy membership functions (distributions) for the purpose of developing a new expert decision-making fuzzy model for evaluating credit risk of corporate clients in a bank. The paper is an extension of previous research conducted on the same subject which was based on Type 1 fuzzy distributions. Our aim in this paper is to address inherent limitations of Type 1 fuzzy distributions so that broader range of banking data uncertainties can be handled and combined with the corresponding hard data, which all affect banking credit decision making process. Banking experts were interviewed about the types of soft variables used for credit risk assessment of corporate clients, as well as for providing the inputs for generating Type 2 fuzzy logic membership functions of these soft variables. Similar to our analysis with Type 1 fuzzy distributions, all identified soft variables can be grouped into a number of segments, which may depend on the specific bank case. In this paper we looked into the following segments: (i) stability, (ii) capability and (iii) readiness/willingness of the bank client to repay a loan. The results of this work represent a new approach for soft data modeling and usage with an aim of being incorporated into a new and superior soft-hard data fusion model for client credit risk assessment.

Key words: Soft data, Type 2 fuzzy distributions, credit risk, default risk, commercial banking

JEL: C53, G21, G32

1 INTRODUCTION

Fuzzy set theory and fuzzy logic were introduced by Lotfi A. Zadeh in 1965 who was almost single-handedly responsible for the early development in this field. Fuzzy Set Theory is a mathematical theory for describing impreciseness, vagueness and uncertainty.The first fuzzy logic framework is referred to as Type 1 Fuzzy Logic Sets and Systems. As an extension of his theory of fuzzy sets and fuzzy logic (Zadeh, 1965, pp. 338-353), the "Theory of Possibility" was developed by Zadeh in 1978 in which he explained that possibility distributions were meant to provide a graded semantics to natural language statements by interpretation of membership functions of fuzzy sets as possibility distributions. He introduced the concept of possibility fuzzy distributions, contrary to random and probabilistic distributions, and noticed that what is probable must initially be possible, but not vice versa. Zadeh (1978) wrote that in dealing with soft data, encountered in various fields, the standard practice was to rely almost completely on probability theory and statistics and he stressed out that those techniques could not cope effectively with those problems in which the softness of data is non-statistical in nature. Soft data encounter predominance of fuzziness. Author's rationale for using fuzzy logic for soft data analysis "rests on the premise that the denotations of imprecise terms which occur in soft database are for the most part fuzzy sets rather than probability distributions" (Zadeh, 1981, pp. 515-541). The difference between probability and possibility is that the concept of possibility is an abstraction of our intuitive perception while concept of probability depends on likelihood, frequency, proportion or strength of belief.

Fuzzy logic has been utilized in various industry areas such as, in artificial intelligence, computer science, control engineering, decision theory, expert systems, logic, management science, operations research, pattern recognition and robotics (Zimmermann, 2001, pp. 158-241; 369-404). Considering risk assessment, many studies of fuzzy logic have appeared in different business areas such as information security, software development, ground water nitrate risk management, system failure, civil hazardous materials, natural hazards, bank, etc. (Zirakja&Samizadeh, 2011, pp. 99-100).

The main criticism of Type 1 Fuzzy Logic Systems is in its limited capability to directly handle data uncertainties (Mendel, 2007, pp. 20-29), considering that the membership grade in the fuzzy set is expressed exactly i.e. making it a crisp value. Therefore Type-2 fuzzy sets and systems are introduced by the inventor of fuzzy sets (Zadeh, 1975, pp. 199-249) which generalize Type-1 fuzzy sets and systems so that more uncertainty can be tackled. The main strength of type-2 fuzzy logic is its ability to deal with the second-order uncertainties that arise from several sources (Karnik et.al, 1999, pp. 643–658) and are preferred over Type 1 fuzzy systems in highly uncertain environment to better handle uncertainty (Wu, 2012, pp. 832-848). Considering Type 2 Fuzzy Systems advantages (Wu, 2012, pp. 832-848; Mendel, 2003, pp.10-13) they have generated lot of interest in the research community.

The purpose of this study is to design and develop possibility distribution modeling of soft data used for corporate client credit risk assessment in commercial banking by applying Type 2 fuzzy membership functions (distributions) for the purpose of developing a new expert decision-making fuzzy model for evaluating credit risk of corporate clients in a bank. Terms fuzzy and possibility distributions are used interchangeably. The paper is an extension of previous research conducted on the same subject which was based on Type 1 fuzzy distributions. Our aim in this paper is to address inherent limitations of Type 1 fuzzy distributions so that broader range of banking data uncertainties can be handled and combined with the corresponding hard data, which all affect banking loan decision making process.

Expert sample is created ad hoc with a commercial bank in Bosnia and Herzegovina that was willing to take part in this project at this initial phase. We are now in a process of adding data from other local banks for the purpose of expanding the relevant soft database. Top senior credit risk assessment experts from this bank were interviewed and they have provided all information about the process, data processing and inputs used for credit risk assessment. Experts have provided inputs for generating universe of discourse, as well as the number and description of membership functions related to each soft variable. Data processing is done by listing all identified soft variables and by mapping their membership values into membership functions based on inputs from interviewed experts. Results of Type 1 fuzzy distributions from previous research (Brkic&Hodzic&Dzanic, 2017) are incorporated in this study.

The results of this work represent a new approach for soft data usage/assessment with an aim of being incorporated into a new and superior soft-hard data fusion model by applying the method of Uncertainty Balance Principle (Hodzic, 2016a, pp. 58-66, Hodzic, 2016b, pp. 17-32) for the purpose of creating a new decision-making fuzzy model of credit risk assessment that will assist bank managers in identifying credit risk factors and improve evaluation of the corresponding default risks of their loan applicants. Design and development of Type 1 and Type 2 possibility distributions of soft data/variables used for corporate client credit risk assessment serve as critical steps in this process.

In this paper, we first present a general overview of credit risk assessment in commercial banking. Following section provides an overview of the results of this study based on which Type 2 fuzzy distribution model of identified soft variables is developed, used by the bank for assessing the credit risk of a corporate loan applicant. Finally, we make conclusions and give directions for future research.

2 CREDIT RISK ASSESSMENT IN COMMERCIAL BANKING

As it was described in our previous research, credit risk is one of the largest risks faced by commercial banks and it is assuming increased importance in a changing regulatory regime and quite volatile market conditions. Risk analysis techniques are powerful tools that help professionals manage uncertainty and can provide valuable support for decision making. These techniques can be either qualitative or quantitative depending on the information available and the level of detail that is required (Bennett and Bohoris, 1996, pp. 467-475). Quantitative techniques rely heavily on statistical approaches while qualitative techniques rely more on judgment than on statistical calculations.

The complex and uncertain nature of loan processing has enforced banks to make loan decisions by utilizing experienced lending officers to perform the essential tasks and evaluations. A loan officer has to fully understand the level of risk a loan would entail and thus has to understand and assess the following: the financial position, repayment ability and strength of the company, whether the company has a sound record of credit worthiness, work history, what is applicants experience and management skills, does the company have a sound business plan which demonstrates his/her understanding of the business and his/her commitment to the success of the business, is company's cash-flow solid and stable, willingness to repay debt and many more. Such analysis incorporates not only the economic data but also the qualitative information concerning the borrower. Data which is subject to this analysis can be classified as hard and soft data. Hard data is usually objective, they express a measure and thus are measurable, quantitative and crisp, while soft data is linguistic, qualitative, subjective and non-measurable.

Besides loan officers, banks usually use various types of scoring models to assess credit risk of a borrower before disbursing a loan. Scoring models were initially introduced to standardize the decision making process and to increase the transparency of a bank's business. They are usually estimated with historical data and statistical methods. Scoring models generally do not follow the Basel II regulatory capital framework definitions since their primary aim is not to fulfill the supervisory requirements but to provide internal decision support. Credit scoring could also be considered as a data mining technique, introduced in 1950s, and since then many methods for applying this technique for credit scoring have been proposed. They can, in general be classified as hard and soft models of data mining (Thomas & Edelman & Crook, 2002; Lando 2004). Soft techniques use fuzzy logic compared to crisp in case of hard techniques. In order to overcome shortcomings of credit scoring models many researches have suggested the use of hybrid methods, which use the advantages of various models. Thus, a single model may not be sufficient in order to identify all the characteristics of the data (Khashei&Bijari&Hejazi, 2012). An example of such model is a soft version of traditional multi-layer perceptrons which is proposed as an alternative classification model, using the unique soft computing advantages of fuzzy logic in which instead of crisp weights and biases, fuzzy numbers are used in multi-layer perceptrons for better modeling of the uncertainties in financial markets (Khashei&Mirahmadi, 2015).

Statistical theory offers a variety of methods for statistical risk assessment. In general, such statistical models use the borrower's characteristic indicators (usually data from financial statements) and (if possible) macroeconomic variables which were collected historically and are available for defaulting (or troubled) and non-defaulting borrowers. Depending on the statistical application of this data, various methods can be used to predict the performance of a borrower. These methods have a common feature in that they estimate the correlation between the borrowers' characteristics and the state of default in the past and use this information to build a forecasting model. The Internal Rating Based Approach (IRBA) of the New Basel Capital Accord allows banks to use their own rating models for the estimation of probabilities of default (PD) as long as the systems meet specified minimum requirements. Most common statistical methods for building and estimation of such models are Regression Analysis, Discriminant Analysis, Logit and Probit Models, Panel Models, Hazard Models, Neural Networks, Decision Trees (Hayden &Porath, 2011, pp.1-12).

Traditional risk models are based on probability and classical set theory which are widely used for assessing market, credit, insurance and trading risk. However, many risks still cannot be analyzed sufficiently by applying classical probability models because of lack of sufficient experience data, lack of knowledge and vagueness, as well as complex cause-and-effect relationships that are inherent in certain risk types. Many authors believe that the best way to solve obstacles in facing with any type of uncertainties is by utilizing fuzzy logic and theory of possibility. It provides a mathematical advantage to capture the uncertainties associated with human cognitive processes, such as thinking and reasoning. "Fuzzy logic is a superset of conventional (Boolean) logic that has been extended to handle the concept of partial truth, truth values between completely true and completely false" (Gupta &Celtek, 2001, p.20). By applying fuzzy logic most variables of a model are described in linguistic terms which makes fuzzy logic models more intuitively similar to the human reasoning. For risks that do not have a proper quantitative probability model, a fuzzy logic system can help model the cause-and-effect relationships, assess the degree of risk exposure and rank the key risks in a consistent way, considering both the available data and experts opinions (Shang &Hossen, 2013, p.3).

3 SOFT DATA MODELING VIA TYPE 2 FUZZY DISTRIBUTIONS

With an aim to eliminate the uncertainty of Type 1 fuzzy distribution results we have extended our previous research (Brkic&Hodzic&Dzanic, 2017) to Type 2 fuzzy distributions and have created a comprehensive database of soft data fuzzy Type 1 and 2 based on inputs provided by interviewed experts. In this section we show the results of Type 2 fuzzy distributions which are used by the targeted bank for the purpose of credit risk assessment of corporate clients.

We conducted a series of interviews with credit risk specialists from a local bank which resulted in a database of soft data used for credit risk assessment of corporate clients. It contains 12 main soft variables that have been analyzed from the perspective of five possible outcome/states per variable, generating a total of 60 fuzzy distributions per expert, as well as per type of fuzzy logic sets and systems (i.e. 60 possibility distributions from the perspective of Type 1 and 60 per Type 2 fuzzy distributions). All identified soft variables can be grouped in following segments: stability, capability and readiness/willingness of the client to repay a loan. Each of these segments have a variety of impact on the assessments going from low impact to medium and high. Considering the resulted number of possibility distributions we are not able to show all results, thus here we consider only some examples of the results with a focus on comparison of different perception of the same variable by different expert, provided in Figures 1-12. Moreover, due to confidentiality we do not disclose estimation results that have been given by the bank experts but we are instead showing graphical illustration of the possibility distribution results. The examples of results shown here are chosen so that they can illustrate various types of distributions e.g., triangles, trapezoids, S-membership function, Z-membership function etc.

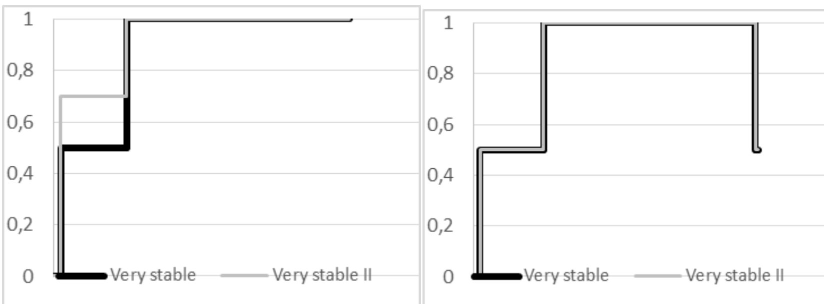

Figure 1 – Very stable Type 1 and Type 2 fuzzy distributions with comparison of perception between two experts of Stability/Capability of the loan applicant based on company size considering its total assets

Fuzzy membership functions examples of Type 2 Stability and capability of a company assessed considering company's size based on its total assets, total income, as well as total number of employees, are shown in figures 1-3. The results for example in Figure 1 show a shift in Type 2 membership function in case of one expert, while no shift in case of the perception of another expert for the same variable. In examples shown in Figure 2 and 3 there is a clear shift of membership function assessed by both experts in case of Type 2 evaluation.

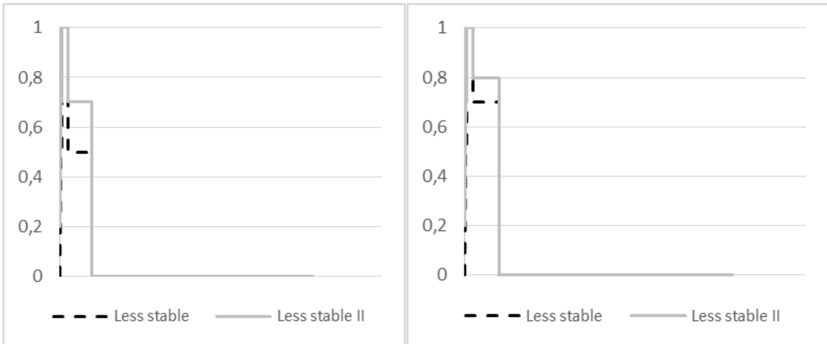

Figure 2 – Less stable Type 1 and Type 2 fuzzy distributions with comparison of perception between two experts of Stability/Capability of the loan applicant based on company size considering its total income

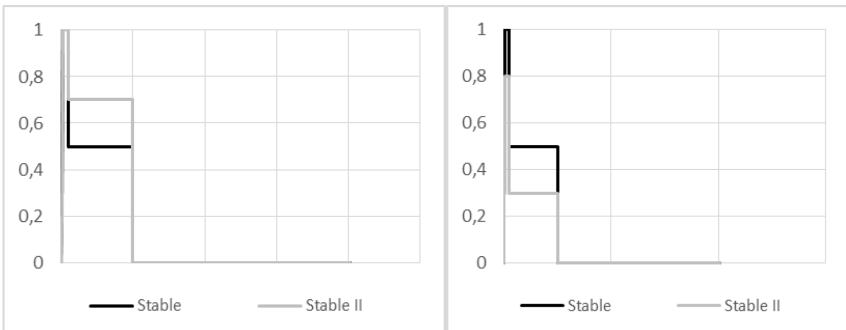

Figure 3 – Stable Type 1 and Type 2 fuzzy distributions with comparison of perception between two experts of Stability/Capability of the loan applicant based on company size considering its total number of employees

Figure 4 demonstrates comparison between less stable Type 1 and Type 2 fuzzy distribution of Stability of the loan applicant considering number of years the company is doing business, as well as comparison in the perception of two experts from the same bank. In case of the first expert Type 2 membership function is shifted to the right while there is an increase in the membership grade in the case of the second expert.

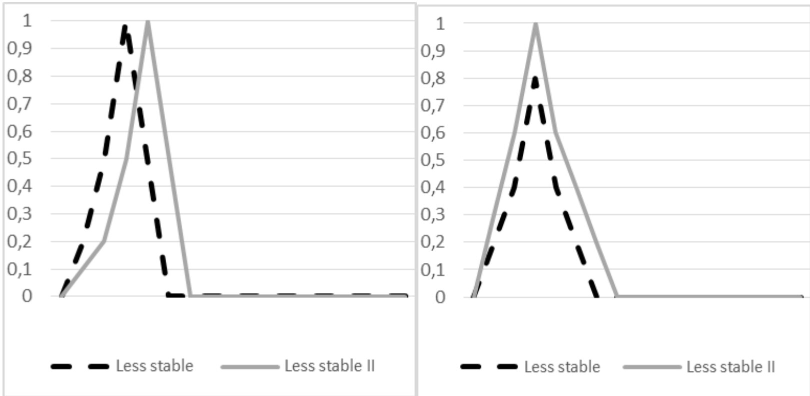

Figure 4 – Less stable Type 1 and Type 2 fuzzy distributions with comparison of perception between two experts of Stability of the loan applicant considering number of years the company is doing business

Extremely stable Type 2 membership function, shown in Figure 5, is shifted to the left in case of one expert's perception, while in the case of another expert the membership function is shifted to the right.

Figures 6-8 demonstrate similar shapes of trapezoid possibility distributions in case of different expert evaluation but are shifted in case of Type 2 evaluation in all shown examples.

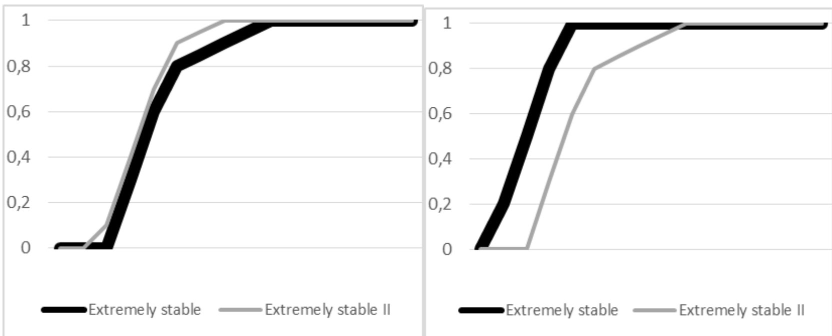

Figure 5 – Extremely stable Type 1 and Type 2 fuzzy distributions with comparison of perception between two experts of Stability of the loan applicant considering number of years it operates profitably (considering operating income)

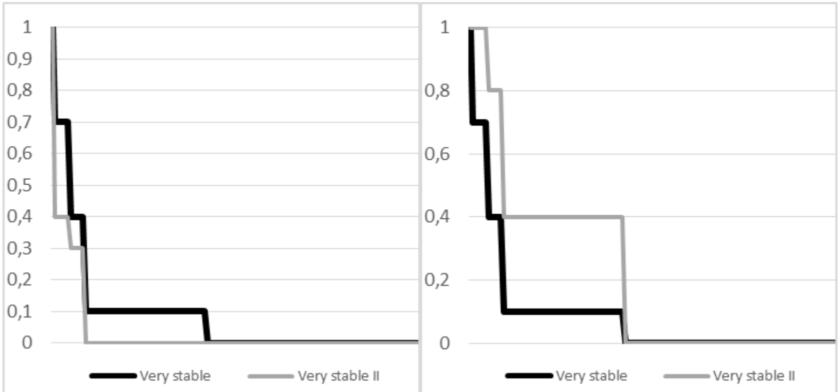

Figure 6 – Very stable Type 1 and Type 2 fuzzy distributions with comparison of perception between two experts of Stability of the loan applicant considering number of days of blocked bank accounts in the last year

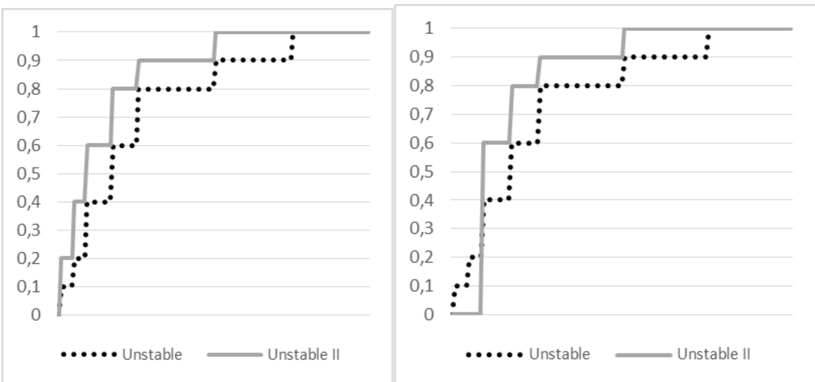

Figure 7 – Unstable Type 1 and Type 2 fuzzy distributions with comparison of perception between two experts of Stability of the loan applicant considering company's repayment history in the bank (if already a client)

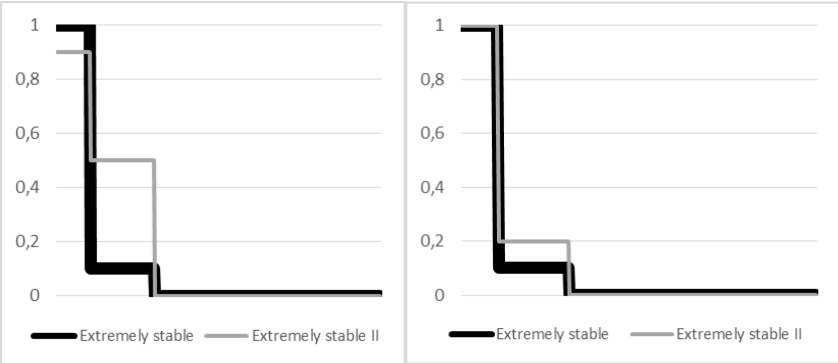

Figure 8 – Extremely stable Type 1 and Type 2 fuzzy distributions with comparison of perception between two experts of Stability of the loan applicant considering company's worst regulatory classification found in the Central Credit Registry

Example of less stable Type 1 and Type 2 fuzzy distributions of Stability of the loan applicant considering future development of the company, shown in Figure 9, express different shapes of membership function between two experts from the same bank. Results of expert evaluation shown in left graph of Figure 9 indicate shift to the right along the universe of discourse, while results from another expert shown in the right part of the Figure 9 indicate a shift upwards along the membership grade.

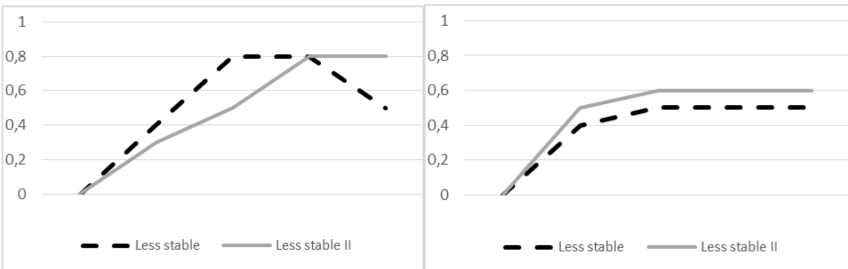

Figure 9 – Less stable Type 1 and Type 2 fuzzy distributions with comparison of perception between two experts of Stability of the loan applicant considering future development of the company

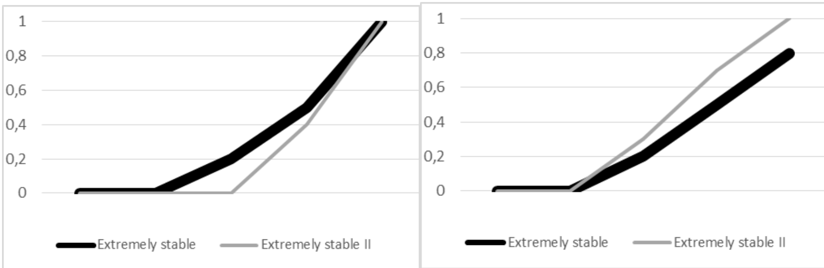

Figure 10 – Extremely stable Type 1 and Type 2 fuzzy distributions with comparison of perception between two experts of Stability of the loan applicant considering company's competition

In case of the Stability of the loan applicant considering company's competition extremely stable Type 1 and Type 2 fuzzy distributions demonstrate in both cases a Z-membership function. However, in the perception of one expert Type 2 membership function is expressed mostly through a slight shift to the right, while an increase in the membership grade is expressed by another expert for the same variable in case of Type 2 estimation.

Result of one expert estimation of Type 2 stable company based on Readiness/Willingness/Character of the management of the company to repay the loan (Figure 11) show a right trapezoid membership function compared to Type 1 where the result indicate a Z-membership function. Type 2 estimation of the same variable of a different expert from the same bank show an upwards shift of the same type of possibility distribution.

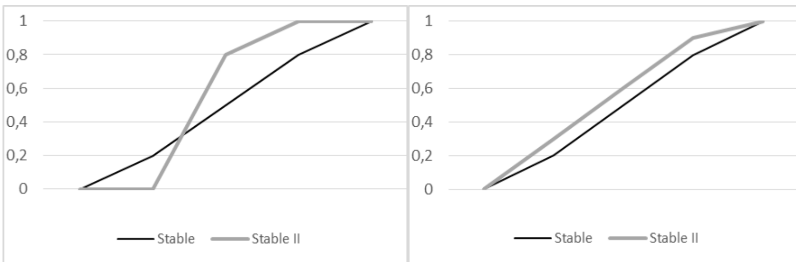

Figure 11 – Stable Type 1 and Type 2 fuzzy distributions with comparison of perception between two experts of Readiness/Willingness/Character of the management of the company to repay the loan

Finally, example provided in Figure 12 with the results of less stable Type 1 and Type 2 fuzzy distributions of Capability/Quality of the company's management, with comparison of perception between two experts, display different shapes of membership functions. However, in case of Type 2 evaluation they demonstrate a shift upwards along the membership grade in the perception of both experts.

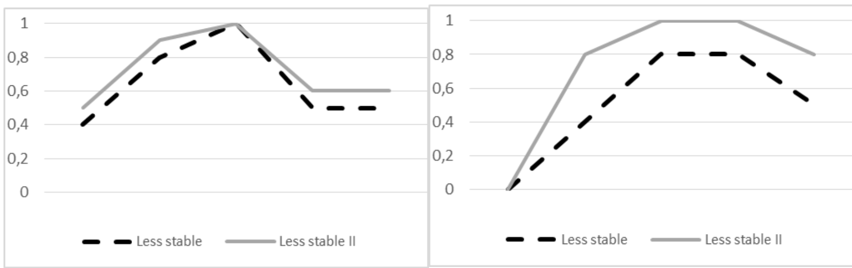

Figure 12 – Less stable Type 1 and Type 2 fuzzy distributions with comparison of perception between two experts of Capability/Quality of the company's management

4 CONCLUSION

In this paper, Type 2 fuzzy logic possibility distribution modeling of main soft variables, used for corporate client credit risk assessment in commercial banking. Banking experts were interviewed about the types of soft variables used for credit risk assessment of corporate clients, as well as for providing the inputs for generating Type 2 fuzzy logic membership functions of these soft variables. Similar to our analysis with Type 1 fuzzy distributions, all identified soft variables can be grouped into a number of segments, which may depend on the specific bank case. In this paper we looked into the following segments: (i) stability, (ii) capability and (iii) readiness/willingness of the bank client to repay a loan.

Our aim in this paper is to address inherent limitations of Type 1 fuzzy distributions. As demonstrated in our results, the Type 2 fuzzy distributions incorporate second-order uncertainties and thus include a broader range of banking data uncertainties which can be handled and combined with the corresponding hard data, in order to finally improve the loan decision making process.

With this we have built a thorough soft database with Fuzzy type 1 and 2 fuzzy logic possibility distributions based on expert interviews. In order to further improve the analysis we are currently in the process of expanding the soft database with inputs from other banks and experts. This database presents a new methodology for transforming linguistic and intuitive (soft) information about bank credit risk data into a series of mathematical fuzzy (possibility) distributions which can be handled quantitatively and combined (fused) with related probabilistic data. The results of this work represent a new approach for corporate client soft data usage/assessment in commercial banking with an aim of finally being incorporated into a new and superior soft-hard data fusion model via the Uncertainty Balance Principle (Hodzic, 2016a, pp. 58-66, Hodzic, 2016b, pp. 17-32) for the purpose of client credit risk assessment and other similar assessments. The Uncertainty Balance Principle (UBP) was defined to express uncertain data vagueness as represented by a fuzzy data models, with a non-uniqueness of related random data distributions (Hodzic, 2017, pp. 785-809).

Type 2 fuzzy logic gives a better basis for applying the method of UBP. This method transfers the fuzzy distribution in equivalent random (hard) distribution which is then combined with the original hard distribution of probability of default via the process of soft hard data fusion. This will be reported in our future papers. Finally, soft hard data fusion offers improved data for credit risk assessment.

Our final aim is to be able to improve bank credit risk assessments by using exact and more precise mathematical methodology.

REFERENCE

Bank for International Settlements. (June 2006). International Convergence of Capital Measurement of Capital Measurement. Basel Committee on Banking Supervision, A Revised Framework Comprehensive Version.

Bennett, J.C., Bohoris, G.A., Aspinwall, E.M., & Hall, R.C., (1996). Risk analysis techniques and their application to software development. European Journal of Operational Research, vol. 95, (3), pp. 467-475.

Brkic, S., Hodzic, M., &Dzanic, E.,. Fuzzy Logic Model of Soft Data Analysis for Corporate Client Credit Risk Assessment in Commercial Banking, (Nov 2017). Fifth Scientific Conference with International Participation “Economy of Integration” ICEI 2017, Available at SSRN: https://ssrn.com/abstract=3079471

Gupta, Vipul K., &Celtek, S., (2001). A fuzzy expert system for small business loan processing, Journal of International Information Management: Vol. 10, Article 2. Available at: http://scholarworks.lib.csusb.edu/jiim/vol10/iss1/2

Hayden E., Porath D., (2011). Statistical Methods to Develop Rating Models, The Basel II Risk Paramenters. London: Springer, pp.1-12.

Hodzic, M., (2016a). Fuzzy to Random Uncertainty Alignment, Southeast Europe Journal of Soft Computing, 5, pp. 58-66.

Hodzic, M., (2016b). Uncertainty Balance Principle, IUS Periodicals of engineering and natural sciences, 4, No.2, pp. 17-32.

Hodzic, M., (2017). Soft to Hard Data Transformation Using Uncertainty Balance Principle, International Symposium on Innovative and Interdisciplinary Applications of Advanced Technologies, Springer IAT 2017: Advanced Technologies, Systems, and Applications II, pp. 785-809.

Karnik, N. N., Mendel, J. M. & Liang, Q., (1999). Type-2 fuzzy logic systems, IEEE Transactions on Fuzzy Systems 7 (6), pp. 643–658.

Khashei, M., &Bijari, M., &Hejazi, S.R., (2012) Combining seasonal ARIMA models with computational intelligence techniques for time series forecasting. Soft Computing, 16, 6, pp-1091–1105.

Khashei, M., &Mirahmadi, A., (2015). A Soft Intelligent Risk Evaluation Model for Credit Scoring Classification, International Journal of Financial Studies, 3, pp. 411-422.

Kickert, W. J. M., (1979). An example of linguistic modeling, the case of Mulder's theory of power, in: M.M. Gupta, R.K. Ragade& R.R. Yager (eds.): Advances in Fuzzy Set Theory and Applications, North-Holland, Amsterdam.

Kosko, B., &Isaka, S., (1993). Fuzzy logic, Scientific American 269, pp.76-81.

Lando, D., (2004). Credit risk modeling: Theory and applications, Princeton Series in Finance; Princeton University Press: Princeton, NJ, USA.

Mendel, J. M., (2003). "Type-2 Fuzzy Sets: Some Questions and Answers, "IEEE Connections, Newsletter of the IEEE Neural Networks Society 1, pp. 10–13

Mendel, J. M., (2007). "Type-2 fuzzy Sets and Systems: An Overview," IEEE Computational Intelligence Magazine, 2, pp. 20–29.

Shang, K., Hossen, Z., (2013). Applying Fuzzy Logic to Risk Assessment and Decision-Making, Casualty Actuarial Society, Canadian Institute of Actuaries, Society of Actuaries.

Thomas, L.C. & Edelman, D.B. & Crook, J.N. (2002) Credit Scoring and its Applications, SIAM Monographs on Mathematical Modeling and Computation; SIAM: Philadelphia, PA, USA.

Viot, G., (1993). Fuzzy logic: Concepts to constructs, AI Expert, 8, pp. 26-33.

Zadeh, L. A., (1965). Fuzzy sets, Information and Control, vol. 8, pp. 338-353.

Zadeh, L. A. , (1975). "The Concept of a Linguistic Variable and Its Application to Approximate Reasoning–1,"Information Sciences, vol. 8, pp. 199–249.

Zadeh, L. A., (1978). Fuzzy sets as a basis for a theory of possibility, Fuzzy Sets and Systems 1, pp. 3-28.

Zadeh, L. A., (1981). Possibility theory and soft data analysis, Selected papers by LotfiZadeh, pp. 515-541.

Zirakja, M.H., Samizadeh, R., (2011). Risk Analysis in E-commerce via Fuzzy Logic. Int. J. Manag. Bus. Res., 1 (3), pp. 99-112.

Wu, D., (2012). "On the Fundamental Differences Between Interval Type-2and Type-1 Fuzzy Logic Controllers", IEEE Transactions on Fuzzy Systems, vol. 20, no. 5, pp 832-848.