Examining Statistical Methods in Forecasting Financial Energy of Households in Poland and Taiwan

Tomasz Korol [ORCID]

Faculty of Management and Economics, Gdansk University of Technology, Narutowicza 11/12, 80-233 Gdansk, Poland; tomasz.korol@zie.pg.gda.pl

Abstract: This paper examines the usefulness of statistical methods in forecasting the financial energy of households. The study's objective is to create the innovative ratios that combine both financial and demographic information of households and implement them in the forecasting models. To conduct this objective, six forecasting models are developed using three different methods—discriminant analysis, logit analysis, and decision trees separately for households in Poland and Taiwan. Such a research approach will answer the question whether the implementation of constructed ratios can increase effectiveness of the forecasting model and its' versatility between different economic regions. The research relies on four samples of households—two learning samples and two testing samples (one for each country) consisting of 2400 households from both countries. This study is one of the first attempts in the literature globally to develop forecasting models based on ratios constructed with a combination of two different types of variables—one financial with one demographic variable. Findings confirm the high usability of the proposed innovative approach for forecasting the financial energy of households by taking into account a wide spectrum of diagnostic variables representing the financial strength of consumers.

1. Introduction

No other crisis gained greater prominence during the past four decades than the Covid-19 pandemic crisis. This is due to its biological nature and wide scope affecting the entire world simultaneously. The household, being the basic economic unit, has been hit hardest by this crisis, often without the government support that enterprises can count on. Although household finance had been attracting substantial academic attention, it has not yet earned its own title and identity. Hence, there is still no commonly accepted definition of household financial instability in the literature. Conceptualizing the term 'financial energy of the household,' the author of this research aims to capture broader aspects of household finance. Households can count on two main types of resources over their lifetime—tangible wealth and human capital. The current pandemic crisis not only reduces the wealth but also decreases the potential of the labor market. Both factors directly influence the financial standing of the household and increase the need for the credits.

There are a number of studies explaining the increased vulnerability of households due to their increased debt and the negative consequences of household debt [1–5]. In particular, studying the vulnerabilities of the household sector is important for at least two reasons. First, the household sector holds the largest share of wealth in developed economies. Since wealth is one of the most important factors in determining household consumption through its lifecycle, household consumption decisions are influenced by its solvency position, thereby, impacting the economic activity. Second, vulnerable households pose a threat to financial stability of the country due to their ties to financial institutions [6].

Debt is one of the most important components of financial vulnerability of the household. However, the financial vulnerability may also be driven by factors other than debt, such as: low income, consumption behaviors that may be induced by irresponsibility or short-sightedness, and demographic factors (e.g., education, marital status, number of children) [7]. Therefore, based on the term "financial energy of the household," the author of this article understands the financial situation of the household and the ability to repay the debt, taking into account a wide range of financial, but also demographic and behavioral factors. Such a broad term of finance has already been applied in terms of refining the medical surge capacity [8] during the COVID-19 pandemic. The second example of interdisciplinary terms and research approaches in today's literature studies can be a theory of the "life cycle" that combines biological, economic, and engineering aspects into one compact study that analyzes the engineering and financial processes at different stages of enterprise development [9]. The other example of broad application of an interdisciplinary term is called "creative destruction" formulated by J. Schumpeter, which is associated with improper, inefficient allocation of assets. From this point-of-view, liquidation of an inefficient enterprise improves the functioning of the whole economy of the country [10]. According to the Cambridge Dictionary, the term "enterprise" is the willingness and energy to do something new that takes plenty of effort. This means that money resources in the direct sense of the word can also be defined as energy. Another example is the theory of energy value. In this theory, the researchers have deduced the "energy value" in terms of the dollar exchange rate of 1.4×10^5 J [11]. The latest research recognizing the term "financial energy" does not exclude the scenario that, in the future, all money will be associated with energy [12].

Majority of studies devoted to forecasting the financial standing of the household is based on traditional credit scoring methods that help banks decide whether or not to grant credit to consumers who apply to them [13]. Most of the scholars use the following determinants in the models, e.g., [14–22]: age, education level, gender, income level, mortgage expenditures, mortgage length, marital status, number of dependents, employment status, credit cards expenditures, number of credits, and value of assets. In other words, existing forecasting models available in the literature contain only individual financial and demographic information. These models lack any constructed ratios unlike in corporate finance where financial ratios are used to evaluate a company's profitability, liquidity, efficiency, and solvency. In the literature, no attempts have been made to verify the possibility and usability of creating ratios based on financial and demographic data that can be implemented in credit scoring models for households. The question arises whether the implementation of constructed ratios can increase not only effectiveness of the forecasting model but also its versatility between different economic regions. Such a model contains the variables based on economic data expressed as a percent (e.g., a relation of debt to the value of assets) and not in absolute numbers.

Therefore, the contribution of paper is three-fold. First, it investigates the usefulness of the proposed ratios that are constructed with the combination of both demographic and financial data of the household. Second, it determines the most effective methodology of the model separately for households in Poland and Taiwan. Third, it verifies the influence of implementing ratios in the evaluation of financial energy of consumers on the versatility of estimated models. To address the mentioned research problems, the author formulated the following objectives of the study: to propose innovative ratios that can increase predictive properties of the model forecasting the financial standing of the household, to estimate six forecasting models (three for Poland and three for Taiwan) by using techniques as discriminant analysis, logit analysis, and decision trees (C&RT), and to compare effectiveness and versatility of the created forecasting models between two different economic regions and forecasting techniques.

The paper consists of five sections. In the introduction, the author justifies the topic, the study objectives, the contributions, and innovations in the literature. Section 2 presents the concepts of financial energy of households and factors affecting it. Section 3 introduces this

study's assumptions regarding constructed data, samples, and implemented forecasting methods. In Section 4, the author presents six models forecasting the financial energy of households and discusses the results of effectiveness tests. Section 5 concludes the paper.

2. Literature Review

Consumer financial health has been a significant problem in many countries all over the world. Typically, studies devoted to evaluating financial energy of households focus on forecasting the risk of consumer insolvency by conducting the analysis of macroeconomic determinants of nonperforming loans (NPLs) in the country or the microeconomic analysis of risk insolvency for individual households using the demographic and behavioral factors of consumers.

In a majority of the studies that investigate the factors affecting the scale of bankruptcies in households in the economy, the macroeconomic determinants are mainly used as explanatory variables. Berge and Boye [23] state that NPLs depend on real interest rates and the unemployment rate, while Rinaldi and Sanchis-Arellano [24] proved that disposable income, the unemployment rate, and monetary conditions have a strong influence on households' financial health. Similarly, Espinoza and Prasad [25] and Kauko [26] found that NPLs are affected by interest rates, along with fiscal and external deficits. Beck et al. [27] estimated that the most significant factors affecting the scale of consumer insolvencies are GDP growth, interest rates, and the exchange rate. Messai [28] also proved that GDP growth, the unemployment rate, and real interest rates have significant influences on NPLs. Ozili [29] revealed the correlation between NPLs and the stages of the business cycle, while Baboucek and Jancar [30] investigated economic developments in the Czech banking sector through unemployment, real GDP growth, exports, imports, the real effective exchange rate, and the inflation rate as indicators of NPLs. Quagliariello [31] presents a regression between the evolution of NPLs as a dependent variable and a set of explanatory variables for the case of Italy. This study revealed that decreasing real GDP growth, increasing unemployment, the real exchange rate, and the consumer price index have a significant effect on loan portfolio quality.

Another branch of the literature focuses on forecasting the risk of consumer bankruptcy with the use of microeconomic factors. Most of the authors, e.g., Aristei and Gallo [14], Diaz-Serrano [15], Ghent and Kudlyak [16], Guiso, Sapienza and Zingales [17], Haughwout, Okah and Tracy [18], Hira [19], Jackson and Kaserman [20], Patel, Balmer and Pleasence [21], Worthington [22], and I-Cheng and Che-hui [32] use the following determinants: age, education level, gender, income level, mortgage expenditures, mortgage length, marital status, number of dependents, employment status, credit card expenditures, number of credit cards, and value of assets.

Despite an exhaustive research on forecasting models for consumer bankruptcy, these studies have mainly focused on either the macro or the micro level. No studies have been devoted to the combination of financial factors with the demographic and behavioral determinants of consumers in the forecasting process of financial strength of households. To fill this gap, the main objective of this research is to develop the innovative ratios describing the financial energy of households and then to estimate six statistical models (three for Central European and three for Far-East Asian households) with constructed ratios.

3. Data, Samples, and Modeling Methods

The author created two learning and two testing samples in the study with one for each research region. Each testing sample consists of 1000 consumers who took consumption credits in Poland and Taiwan, while each learning sample consists of 200 consumers.

The learning samples were used to estimate the forecasting models (the models were taught on a given sample of consumers to recognize solvent and insolvent households). The testing samples were used to check the effectiveness of created models on the data unknown to the models. Such a research approach lets the models' testing in a situation

very close to the real case where the analysts, e.g., in the bank, do not know if a given consumer is a future solvent or insolvent one.

It is important to note that there is a balanced number of households with good (nonbankrupt) and bad (bankrupt) financial health in both samples. Such a research approach ensures the reliability of the conducted study as all the models are tested on an equal number of bankrupt and nonbankrupt cases. The intention of conducting a separate study on forecasting "the financial energy" of households in Poland and Taiwan was to evaluate the usefulness of created ratios that combine economic and demographic data on the example of two very different countries from economic and cultural perspectives' behaviors. Such a research approach lets us conclude the versatility and predictive properties of proposed ratios and models.

In Table 1, there is a presented gender distribution of consumers in both testing samples. There are 58% males and 42% females in the testing sample of Polish consumers, while, in a testing sample of Taiwanese households, there are 63.6% males and 36.4% females. The higher percentage of males in the Far-East Asian sample can be caused by cultural factors. It is also interesting to note that, in both samples, a higher percentage of male consumers (58.62% in Central Europe, i.e., 340 out of 580 male consumers and 56.91% in Far-East Asia, i.e., 362 out of 636 male consumers) are in the state of bad financial standing than female ones (38.09% in Poland, i.e., 160 out of 420 female consumers and 37.91% in Taiwan, i.e., 138 out of 364 female consumers).

Table 1. Distribution of consumers' gender in two testing samples. Source: based on own studies.

	Bankrupt	**Nonbankrupt**
Central Europe (Poland)		
Female	160	260
Male	340	240
Total	500	500
Far-East Asia (Taiwan)		
Female	138	226
Male	362	274
Total	500	500

The status of "nonbankrupt" is given to the households that were repaying the credits with no delays and "bankrupt" to those who had delays in repayment longer than 3 months.

Looking more closely at the demographic characteristic of the testing samples, it can be seen that, in the case of analyzed female consumers in Poland (Figure 1a), women in the group age of 27–50 stated the highest share in the sample (67.5% of bankrupt consumers and 62.3% of nonbankrupt ones). Similarly, in the case of male consumers (Figure 1b), men in the group age of 27–50 were the biggest part of the bankrupt and nonbankrupt households. Here, the noticeable share of male consumers with solvency problems was a group of consumers older than 60 years old—6.2%, while, in the case of female consumers in the state of bankruptcy, it was only 1.9% (Figure 1).

Concerning the age of consumers in the testing sample from Taiwan (Figure 2), 71.7% of bankrupt women were from the age group of 27–50. The highest share of nonbankrupt female consumers was women in the age of 27–50 (54.0%) and 51–60 (20.8%). In the case of male consumers from Taiwan, the share of bankrupt and nonbankrupt consumers was at a similar level between 58.3% and 61.7%. It is worth noticing that quite a high share of bankrupt male consumers was older than 60 years old (11.6%). It is almost twice as big a share as in the case of European bankrupt male consumers (11.6% vs. 6.2%).
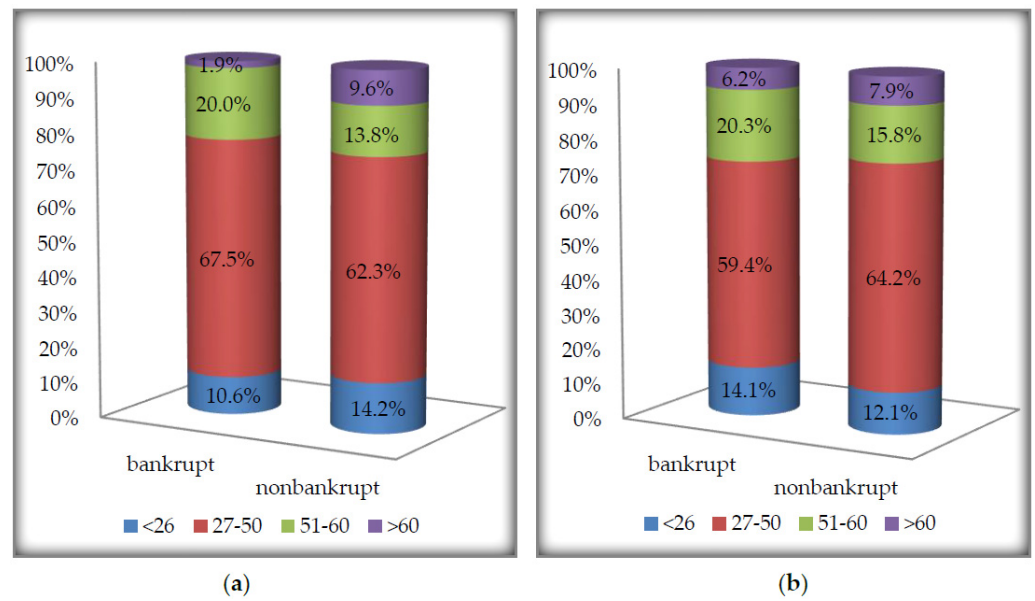
Figure 1. The age structure of bankrupt and nonbankrupt consumers in Poland, in testing samples of: (**a**) female and (**b**) male. Source: Based on own studies.
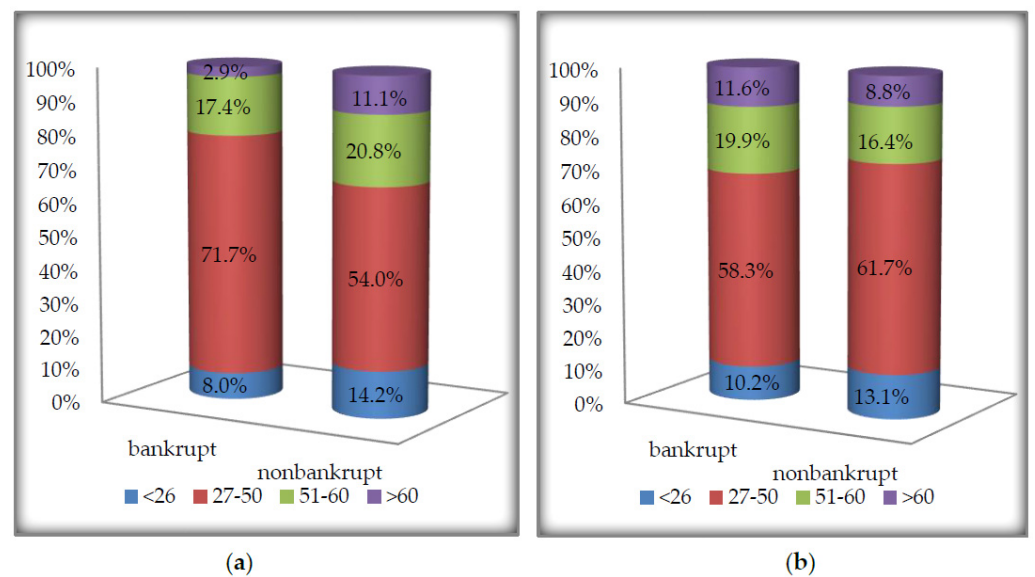

Figure 2. The age structure of bankrupt and nonbankrupt consumers in Taiwan, in the testing sample of: (**a**) female and (**b**) male. Source: Based on own studies.

Comparing the educational level of analyzed consumers (Figure 3), it can be noticed that the highest share of both bankrupt and nonbankrupt households in the Polish sample were consumers at the master or doctorate level (45.6% of bankrupt consumers and 49.8% of nonbankrupt ones). In Taiwanese households, the highest share was devoted to consumers at the bachelor level (35.4% of bankrupt and 35.2% of nonbankrupt households).
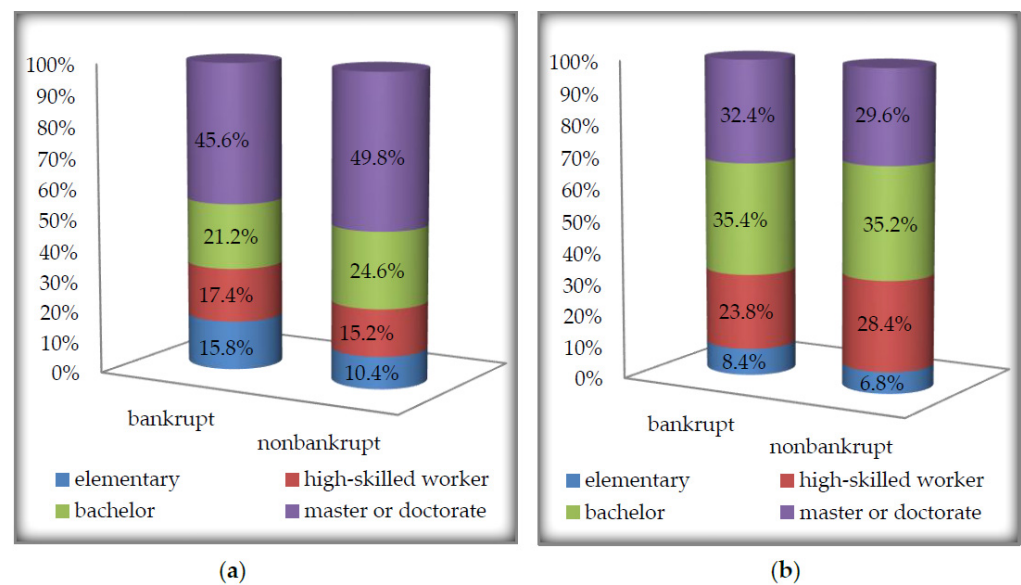
Figure 3. The education level structure of bankrupt and nonbankrupt consumers in the testing sample of: (**a**) Central Europe and (**b**) Far-East Asia. Source: Based on own studies.

All consumers were described by six financial (annual income, monthly income, the value of total assets owned, the value of total credits taken, the value of credit cards, value of monthly interest rates paid) and five demographic (age, education level, number of children, marital status, length of employment) variables. Based on these eleven variables, the author of this study has created twelve ratios that combine the financial and demographic information. It is worth noting that it is one of the first attempts in literature worldwide to use such ratios that combine different types of information in evaluating the financial situation of the households. Table 2 presents the created ratios with the given formulas and the description. Such a research approach uses a wide range of important information about consumers and increases the versatility of the study. These ratios are not denominated in money value nor strictly in demographic units that would be limited to use in only one country, but it can be widely used in other regions of the world.

Table 2. The list of created ratios used in evaluating the financial energy of households. Source: based on own studies.

Symbol	Formula	Description
X1	annual income/value of total assets	It represents the share of annual income in value of total assets owned by the consumer. The higher the value the stronger is the financial energy of analyzed consumer.
X2	annual income/total credits	It represents the relation of annual salary of consumer to the value of total credits. In other words, it can be used to evaluate how many years it will take to pay back all the credits. The higher the value, the stronger is the financial energy of the analyzed consumer.
X3	interest rates paid/monthly income	It represents the information about the burden of a monthly salary with payments of interest rates of credits. The lower the value, the stronger is the financial energy of the analyzed consumer.
X4	credits cards/total credits	It represents the share of credit cards in the value of total credits taken by the consumer. The higher the value, the weaker is the financial energy of the analyzed consumer.
X5	(value of total assets—total credits)/total credits	It shows the relation of net value of total assets to the value of total credits taken by the consumer. The higher the value, the stronger is the financial energy of the analyzed consumer.
X6	value of total assets/total credits	It shows how many times bigger are total assets owned by consumer in relation to the value of all credits taken by him/her. The higher the value, the stronger is the financial energy of the analyzed consumer.
X7	monthly income/credits cards	It shows how many times bigger is the monthly salary compared to the credit card debt. The higher the value, the stronger is the financial energy of the analyzed consumer.

Table 2. *Cont.*

Symbol	Formula	Description
X8	education/age	The variable 'education' is described by the following values: 1—elementary school, 2—high skilled worker, 3—bachelor degree, 4—master or doctorate degree. The variable 'age' is described by the following values: 1—from 27 to 50 years old, 2—from 51 to 60 years old, 3—younger than 26 or older than 60 years old. The higher the value of this relation, the stronger is the financial energy of the analyzed consumer.
X9	education/number of children	The variable 'number of children' is described by the following values: 1—from 0 to 2 children, 2—from 3 to 4 children, 3—more than 4 children. The higher the value of this relation, the stronger is the financial energy of the analyzed consumer.
X10	marital status/length of employment	The variable 'marital status' is described by two values: 1—married, 2—single, widowed. The variable 'length of employment' is described by three values: 1—up to 5 years of job experience, 2—from 6 to 10 years of work experience, 3—more than 11 years of work experience. The higher the value of this relation, the stronger is the financial energy of the analyzed consumer.
X11	education/(total credits/annual income)	It represents the ability of the consumer to pay back the credits taking into account the annual salary and the level of education of the consumer. The higher the value of this relation, the stronger is the financial energy of the analyzed consumer.
X12	age/(total credits/annual income)	It represents the ability of the consumer to pay back the credits by taking into account the annual salary and the age of the consumer. The higher the value of this relation, the stronger is the financial energy of the analyzed consumer.

The author used three methods to estimate the forecasting models—multivariate discriminant analysis (MDA), logit (LOG), and decision tree models (C&RT).

The multivariate discriminant analysis method is classified as a pattern (teacher) classification technique because the discriminant function's value, determined for the analyzed households, is compared to a pattern, thereby defining the consumer belonging to a class. First the linear discriminant function is estimated, which is a weighted sum of the analyzed diagnostic variables in the following form.

$$Z = d_0 + d_{1 \times 1} + d_2 x_2 + ... + d_n x_n \tag{1}$$

where Z is the dependent variable, x_i is the independent variables ($i = 1, 2, ..., n$), and d_i is the discriminant weights. Then, the discriminant function's cutoff value is determined. When the value of the function Z for the analyzed household is smaller than the cutoff value. The consumer is classified as a consumer with high risk of insolvency (low financial energy and capabilities). However, if the value is higher than the threshold, the consumer is classified as not at risk of bankruptcy (high financial energy and capabilities).

In the case of the logit model, the result of the logistic regression function is the likelihood of an event p_i. Estimating a consumer financial health is the probability of an

analyzed consumer belonging to one of two sets: "bankrupt" or "nonbankrupt." The p_i function takes the following form:

$$P(Y = 1) = 1/(1 + \exp^{-z}) = \exp^{z}/(1 + \exp^{z}) \tag{2}$$

where $P(Y = 1)$ is the dependent variable, the probability of adoption by variable Y is the value of 1. Z is the value of the linear function Z, where $Z = d_0 + d_1 x_1 + d_2 x_2 + ... + d_n x_n$ [x_i — explanatory variables (i = 1, 2, ..., n), and d_i—weights (i = 1, 2, ..., n)]. The value of indicator $P(Y = 1)$ occurs in the range of 0 to 1. To use the estimated logit model, a certain threshold (P_{cutoff}) of function $P(Y = 1)$ must also be adopted, as in the case of discriminant analysis.

$$P(Y = 1) \leq P_{cutoff} \text{ then } Y = 0 \tag{3}$$

$$P(Y = 1) > P_{cutoff} \text{ then } Y = 1 \tag{4}$$

Decision trees allow the presentation of the rules arising from the collected data using hierarchical, sequential structures. These rules in a recursive way enable division of a multi-dimensional feature space into subsets, which are homogeneous or almost homogeneous in one characteristic. One of the most popular and most effective methods of building a decision tree is the C&RT algorithm. A characteristic feature of this algorithm is the search of the domain of all possible divisions and split into two disjoint and complementary subsets. This method involves making decisions about each division of a given set into subsets, resulting in the creation of a node, or a decision not to divide and classify data to one class, thus, creating a leaf. This procedure is repeated recursively until the moment when one cannot make more divisions, that is, create the next node, and, thus, every branch ends with a leaf. The decision to create the leaf, which is the recognition that the set is sufficiently homogeneous due to the explained variable (end of divisions) or node, which is a continuation of divisions determines the size of the decision tree.

To evaluate the effectiveness of the created models, the following formula was used (the symbols are explained in Table 3):

$$S = \{1 - [(D1 + D2)/(BR + NBR)]\} \times 100\% \tag{5}$$

Table 3. Classification matrix of statistical forecasting models. Source: based on own studies.

Predicted Status Actual Status	Bankrupt	Nonbankrupt
Bankrupt	TRUE	Type I error E1 = D1/BR × 100%, where D1 is the number of bankrupted households classified by the model as nonbankrupt consumers, and BR is the number of bankrupt households in the sample.
Nonbankrupt	Type II error E2 = D2/NBR × 100%, where D2 is the number nonbankrupt households classified by the model as a consumer at risk of bankruptcy, and NBR is the number of nonbankrupt consumers in the sample.	TRUE

It should be stressed that, from the point of view of the banks, in addition to analyzing the overall effectiveness (S) of the model, it is important to distinguish two types of errors generated by the models. The author has used the classification matrix presented in Table 3

with the identification of such errors. Type I error indicates the classification of a consumer that, in the future, will experience problems with repaying the credits as a consumer with no risk of insolvency. Granting a credit to a future insolvent household (Type I error) will result in losses of bank that will have difficulties in collecting back the previously given credit. Conversely, Type II error means the loss of "only" potential profits by deciding not to grant the credit to a healthy consumer, believing that this is a future bankrupt household. For this reason, Type I error is considered to be more costly and dangerous from the point of view of the solvency of the bank than Type II.

4. Results and Discussion

To conduct the study, the author of research first calculated the set of 12 ratios (Table 2) for all consumers in learning and testing samples. In the second stage of study, three forecasting models for Central European and three predicting models for the Far-East Asian household were developed. The first estimated model was the multivariate discriminant analysis model. Using a forward stepwise regression method, the following forms of MDA models were estimated.

- For European consumers:

$$2.1 \times X1 - 3.08 \times X3 + 0.54 \times X5 - 0.275 \times X12 \tag{6}$$

- For Asian consumers:

$$1.017 \times X1 + 1.15 \times X2 - 2.158 \times X3 + 0.984 \times X5 - 0.358 \times X12 \tag{7}$$

Both models are well fitted to empirical data (adjusted coefficient of determination equal to 76.36% and 83.21%), and all selected independent variables were highly statistically significant ($p \ll 0.01$). All the variance inflation factors (VIF) were smaller than five, so the multicollinearity problem does not exist. It is worth noting that the MDA model for Polish consumers comprise four ratios, while, for Taiwanese households, the model contains five ratios. It is also interesting that models for both regions contain the same type ratios (X1, X3, X5, and X12). It is the evident that proposition of constructed ratios that combine financial and demographic information increased the versatility of the models.

The second developed model was the logit model, which was estimated using forward stepwise regression. For individual samples of households, the following forms of this model have been estimated.

- For European consumers:

$$1.954 \times X2 - 2.158 \times X3 + 0.945 \times X5 - 0.469 \times X12 \tag{8}$$

- For Asian consumers:

$$1.216 \times X1 + 0.875 \times X2 + 1.907 \times X5 - 0.715 \times X12 \tag{9}$$

Above both models are also well fitted to empirical data (adjusted coefficient of determination equal to 89.55% and 84.75%), and all selected independent variables were highly statistically significant ($p < 0.01$). All VIF were smaller than five so the multicollinearity problem does not exist either. Both logit models contain four ratios. Three out of these four ratios are the same for both economic regions (X2, X5, and X12). Comparing to MDA models, it can be seen that very similar information was selected in developing logit models. The most common ratios for all four models are four financial ratios (X1, X2, X3, X5) and one ratio that combine demographic variable (age) and financial variables (total credits/annual income). This ratio combines the relation of age to number of years needed for repayment of all the credits by the consumer.

The third type of a model developed to predict the financial energy of households was a decision tree model (C&RT). Figure 4 shows the structure of the decision tree model for

the European sample of consumers. Figure 5 shows the structure of such decision trees for the analysis of Asian households. Once again, it is interesting to observe that, in the case of the C&RT model for European and Asian consumers, the models share only one common financial ratio (X5), and one demographic ratio (X11). The model for Taiwanese consumers comprises very distinct structure (Figure 5) of five ratios compared to the model for a Polish household that is relatively simpler and consists of only three ratios. Two out of these five ratios are again ratios that include a combination of financial and demographic variables (X11 and X12).
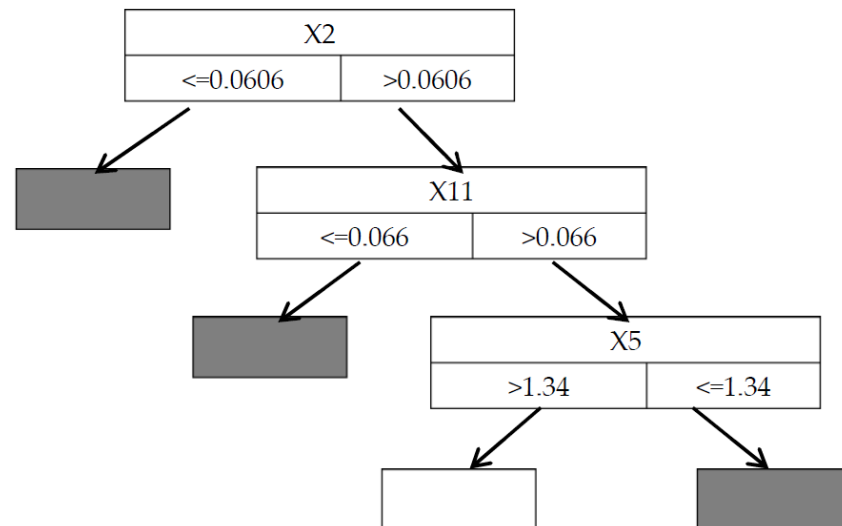

Figure 4. The structure of the Classification and Regression Tree (C&RT) model for Central European households. Gray box indicates consumers at risk of insolvency. White box, consumers at high financial health. Source: based on own studies.
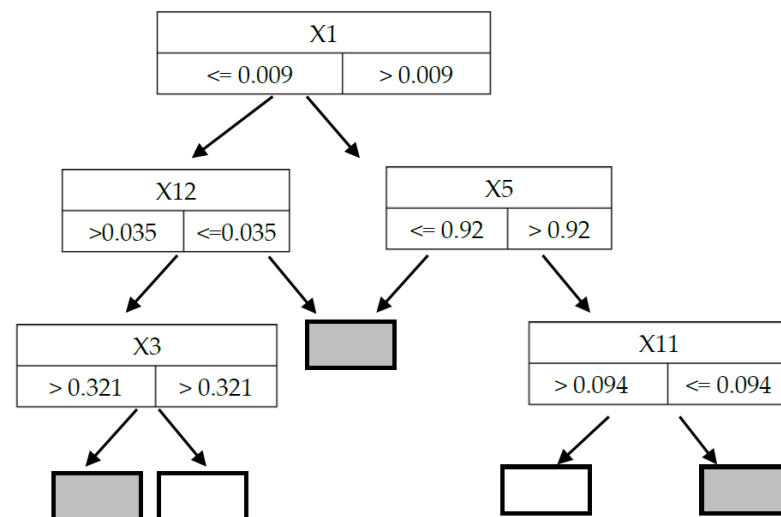

Figure 5. The structure of the Classification and Regression Tree (C&RT) model for Far-East Asian households. Gray box indicates consumers at risk of insolvency. White box indicates consumers at high financial health. Source: based on own studies.

Table 4 presents the results of all created models for European and Asian consumers obtained by the testing samples consisting of 500 bankrupt and 500 nonbankrupt households. Analyzing the results for European consumers, it can be seen that the highest effectiveness was achieved using the logit model, with 92.70% correct classifications. The second highest

effectiveness was obtained by a multivariate discriminant analysis model (89.60% of correct classifications). The decision tree model is characterized by 85.60% effectiveness.

Table 4. The results of the effectiveness of models for Central European and Far-East Asian consumers (testing sample). Source: based on own calculations.

Testing Sample		Effectiveness		
		DA	LOG	C&RT
Central European households	E1	8.20% (41)	6.20% (31)	15.80% (79)
	E2	12.60% (63)	8.40% (42)	13.00% (65)
	S	89.60%	92.70%	85.60%
Far-East Asian households	E1	13.80% (69)	10.60% (53)	17.40% (87)
	E2	10.80% (54)	9.20% (46)	15.20% (76)
	S	87.70%	90.10%	83.70%

Although the C&RT model, for both European and Asian households, achieved the worst results, it is necessary to note that the effectiveness above 80% is considered a very good level of forecasting quality. Looking at the results of models for Asian households, it is evident that the logit model has the best predictive properties (90.10% effectiveness). The second-best model for the Asian sample is the MDA model with 87.70% correct predictions.

It is also interesting to note that all three forecasting models for Central European households generated fewer Type I errors (E1) than did models developed for evaluating Far-East Asian consumers. Such errors indicate how many bankrupt consumers are classified as nonbankrupt ones. Type I errors, for clear financial reasons, are far more dangerous than Type II errors (E2).

What it is more important to note is that the logit model has the fewest Type I errors for both European and Asian consumers, and, is the only model that achieved effectiveness above 90%. Based on overall effectiveness and Type I errors among all six models, both logit models are shown to be the best bankruptcy predictive techniques for households from analyzed economic regions. The multivariate discriminant analysis models are the second-best models if we take into account overall effectiveness and Type I errors.

To enrich the study, the author conducted comparative analyses of overall effectiveness of created models (MDA, LOG, and C&RT) with effectiveness of models available in literature—MDA and LOG model of I-Cheng and Che-Hui [32] for Taiwanese consumers, and C&RT model of Boyle et al. [33] for British households. In both studies, the authors used traditional demographic and financial ratios without any attempt at combining both variables in one ratio as in the proposed research.

Comparing the overall effectiveness of the models found in literature for the effectiveness of models created by the author, it can be seen (Figure 6) that the author's models are characterized by:

- 10.6 percentage points better effectiveness in case of the C&RT model for Polish consumers and 8.7 percentage points in case of the C&RT model for Taiwanese consumers,
- 10.7 percentage points higher effectiveness in case of the LOG model for Polish households and 8.1 percentage points in case of the LOG model for Taiwanese consumers,
- 15.6 percentage points for better effectiveness in case of the MDA model for European households and 13.7 percentage points in case of the MDA model for Asian consumers.

Such results support the idea of creating the ratios that consists a combination of financial and demographic information. It is evident that proposed research generated much better effectiveness of the models and the created models are characterized by high versatility of usage in completely different economic regions.
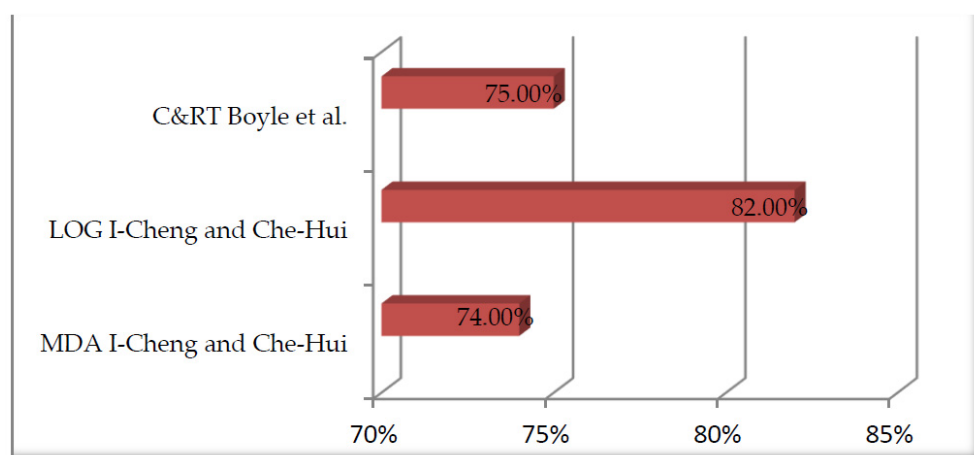
Figure 6. The comparison of effectiveness of the forecasting models. Source: based on own studies.

5. Conclusions

Although many studies have aimed at forecasting the financial standing of the households, the vast majority of them used the individual financial or demographic variables without any attempts of combining these two very different types of information into one relation in the form of the ratios.

The presented empirical study shows that such a research approach aiming at using the set of demographic-financial ratios gives us a much wider spectrum of possibilities at evaluating the financial energy of consumers. We have to remember that a single demographic variable (e.g., age) or financial variable (e.g., annual income) itself does not support the forecasting model with a large load of information important from the point of view of financial strength of the household.

This paper presents one of the first attempts at implementing such ratios in the forecasting process in the literature worldwide. A conducted research approach provided clear evidence that developing the forecasting models with such ratios increases the versatility of the models. First, the models use relative values that allow using them in different countries and regions with different currencies and the level of wealth (e.g., with different values of the average income and credits taken). Second, the models consist of information that contains the relation of financial burden to demographic capabilities (e.g., high level of taken credits compared to young vs. an old age of the consumer, or a high level of taken credits in relation to a well-educated consumer vs. a poorly educated one). All these factors directly and indirectly influence the financial power of the household.

Moreover, this empirical study revealed that, among used statistical techniques, the logit models are the most effective for households in both economic regions. These models showed fewer Type I errors than Type II errors, and obtained the highest overall effectiveness of the forecast. Another conclusion is that the forecasting models for European and Asian consumers very often use the same type of proposed ratios, which means that similar information is necessary to evaluate the process of forecasting the financial energy of consumers, but the models for European households are characterized by a lower cost of the errors (smaller level of type I errors than type II errors).

The main limitation of the research is limited data access. The author met many obstacles and difficulties to collect the reliable data with such a wide variety of information. For each household, the data was collected and calculated individually. It was a long computational process. The author is going to continue the research toward the use of macroeconomic variables of selected countries in predicting the financial standing of consumers.

References

1. O'Connor, G.E.; Newmeyer, C.E.; Ching Wong, N.Y.; Bayuk, J.B.; Cook, L.A.; Komarova, Y.; Loibl, C.; Ong, L.L.; Warmath, D. Conceptualizing the multiple dimensions of consumer financial vulnerability. *J. Bus. Res.* **2019**, *100*, 421–430. [CrossRef]
2. French, D.; Vigne, S. The causes and consequences of household financial strain: A systematic review. *Int. Rev. Financ. Anal.* **2019**, *62*, 150–156. [CrossRef]
3. Gross, M.; Poblacion, H. Assessing the efficacy of borrower-based macroprudential policy using an integrated micro-macro model for European households. *Econ. Model.* **2017**, *61*, 510–528. [CrossRef]
4. Bakkeli, N.Z. Health and economic scarcity: Measuring scarcity through consumption, income and home ownership indicators in Norway. *SSM Popul. Health* **2020**, *11*, 1–13. [CrossRef]
5. Barba, A.; Pivetti, M. Rising household debt: Its causes and macroeconomic implications—A long-period analysis. *Camb. J. Econ.* **2009**, *33*, 113–137. [CrossRef]
6. Ampudia, M.; Vlokhoven, H.; Zochowski, D. Financial fragility of euro area households. *J. Financ. Stab.* **2016**, *27*, 250–262. [CrossRef]
7. Anderloni, L.; Bacchiocchi, E.; Vandone, D. Household financial vulnerability: An empirical analysis. *Res. Econ.* **2012**, *66*, 284–296. [CrossRef]
8. Melnychenko, O. The Energy of Finance in Refining of Medical Surge Capacity. *Energies* **2021**, *14*, 210. [CrossRef]
9. Kosanke, K.; Vernadat, F.; Zelm, M. CIMOSA: Enterprise engineering and integration. *Comput. Ind.* **1999**, *40*, 83–97. [CrossRef]
10. Schumpeter, J. *Capitalism, Socialism and Democracy*; Routledge: London, UK, 1994; pp. 82–83.
11. Beaudreau, B.C.; Pokrovskii, V.N. On the energy content of a money unit. *Phys. A Stat. Mech. Appl.* **2010**, *389*, 2597–2606. [CrossRef]
12. Espen, G.H. The smallest possible money unit! When money crashes into the laws of physics. *Phys. A Stat. Mech. Appl.* **2020**, *560*, 125–143.
13. Lyn, C.T. A survey of credit and behavioural scoring: Forecasting financial risk of lending to consumers. *Int. J. Forecast.* **2000**, *16*, 149–172.
14. Aristei, D.; Gallo, M. The determinants of households' repayment difficulties on mortgage loans: Evidence from Italian microdata. *Int. J. Consum. Stud.* **2016**, *40*, 453–465. [CrossRef]
15. Diaz-Serrano, L. Income volatility and residential mortgage delinquency across the EU. *J. Hous. Econ.* **2005**, *14*, 153–177. [CrossRef]
16. Ghent, A.C.; Kudlyak, M. Recourse and residential mortgage default: Evidence from US states. *Rev. Financ. Stud.* **2011**, *24*, 3139–3186. [CrossRef]
17. Guiso, L.; Sapienza, P.; Zingales, L. The determinants of attitudes towards strategic default on mortgages. *J. Financ.* **2013**, *68*, 1473–1515. [CrossRef]
18. Haughwout, A.; Okah, E.; Tracy, J. *Second Chances: Subprime Mortgage Modification and Re-Default*; Staff Reports Federal Reserve Bank of New York: New York, NY, USA, 2009; p. 417.
19. Hira, T.K. Promoting sustainable financial behaviour: Implications for education and research. *Int. J. Consum. Stud.* **2012**, *36*, 502–507. [CrossRef]
20. Jackson, J.R.; Kaserman, D.L. Default risk on home mortgage loans: A test of competing hypotheses. *J. Risk Insur.* **1980**, *47*, 678–690. [CrossRef]
21. Patel, A.; Balmer, N.J.; Pleasence, P. Debt and disadvantage: The experience of unmanageable debt and financial difficulty in England and Wales. *Int. J. Consum. Stud.* **2012**, *36*, 556–565. [CrossRef]
22. Worthington, A.C. Debt as a source of financial stress in Australian households. *Int. J. Consum. Stud.* **2006**, *30*, 2–15. [CrossRef]
23. Berge, T.O.; Boye, K.G. An analysis of bank's problem loans. *Nor. Bank Econ. Bull.* **2007**, *78*, 65–76.
24. Rinaldi, L.; Sanchis-Arellano, A. *Household Debt Sustainability: What Explains Household Non-performing Loans? An Empirical Analysis*; ECB Working Paper; European Central Bank: Frankfurt am Main, Germany, 2006.
25. Espinoza, R.; Prasad, A. *Non-Performing Loans in the GCC Banking System and Their Macroeconomic Effects*; IMF Working Paper No. 10/224; International Monetary Fund: Washington, DC, USA, 2010.
26. Kauko, K. External deficits and non-performng loans in the recent financial crisis. *Econmic Lett.* **2012**, *115*, 196–199. [CrossRef]
27. Beck, R.; Jakubik, P.; Piloiu, A. Key determinants of non-performing loans: New evidence from a global sample. *Open Econ. Rev.* **2015**, *26*, 525–550. [CrossRef]
28. Messai, A. Micro and macro determinants of non-performing loans. *Int. J. Econ. Financ. Issues* **2013**, *3*, 852–860.
29. Ozili, P.K. How bank managers anticipate non-performing loans: Evidence from Europe, US, Asia and Africa. *J. Appl. Financ. Account.* **2015**, *1*, 73–80. [CrossRef]

30. Baboucek, I.; Jancar, M. *A VAR Analysis of the Effects to Macroeconomic Shocks to the Quality of the Aggregate Loan Portfolio of the Czech Banking Sector*; Working Paper Series; No. 1/2005; The Czech National Bank: Prague, Czech Republic, 2005.

31. Quagliariello, M. *Are Macroeconomic Indicators Useful in Predicting Bank Loan Quality: Evidence from Italy*; Bank of Italy: Rome, Italy, 2003.

32. Cheng, Y.; Che-Hui, L. The comparisons of data mining techniques for the predictive accuracy of probability of default of credit card clients. *Expert Syst. Appl.* **2009**, *36*, 2473–2480.

33. Boyle, M.; Crook, J.; Hamilton, R.; Thomas, L. *Methods for Credit Scoring Applied to Slow Payers*; Oxford University Press: Oxford, UK, 1992; pp. 75–90.