

Paulina Anioła-Mikołajczak  
Uniwersytet Przyrodniczy w Poznaniu

## An Application of Logistic Regression in Identification of the Factors Influencing the Use of Credit by Polish Households

### Summary

The dynamic socio-economic changes, development of financial markets, increasing importance of media, and growing consumerism undoubtedly contribute to changes in the financial behaviour of Polish households. The constantly changing reality enforces the need for a continuous repeat and update research on financial behaviour of households. The main aim of the study is to identify the factors influencing the use of credit by Polish households. The method of logistic regression was applied in order to achieve the main aim of the study. The source material was individual data of the Household Budget Survey in 2011. The survey was conducted by the Central Statistical Office of Poland. The research results show that the use of credits by Polish households is most influenced by the householder's sex, age, socio-occupational status, family type, and income in the household.

**Key words:** household indebtedness, socio-economic determinants, logistic regression.

**JEL codes:** D10, D14

### Introduction

Financial behaviors, including debt behaviors, are one of the most important areas of the general theory of behavior. Financial practices are a subset of household behaviors (Cynamon, Fazzari 2008). Each person in daily life takes a number of financial decisions, eg. about the payment method – the consumers can spend on resources from current income, he can use for this purpose the accumulated savings, or he take out a loan.

Nowadays, the phenomenon of household indebtedness has become a permanent element of the modern consumer society. Households somewhat accustomed to living on credit and treat it as a common source of financing their needs (Lea, Webley, Walker 1995; Raijas et al. 2010). Inadequate financial management, including credit management, often due to a lack of financial awareness resulting from insufficient financial education, may lead to the phenomenon of household over-indebtedness (Bywalec 2009; Świecka 2008; 2009).

The ongoing socio-economic changes, dynamic development of financial markets, the media, as well as increasing availability of modern technologies undoubtedly contribute to changes in the financial behavior of households. Moreover, constantly changing reality enforces the need for continuous repeat and update research on financial behavior of households.

It should be noted that the household sector, both in Poland and in the world, is very diverse. The scale of this variation is due to a number of economic, socio-demographic and psychological factors of closer and more distant environment in which they operate. That is why it is important to conduct research on the factors determining the financial, including debt behavior.

Świecka (2009) attempted to create a classification of the factors influencing the debt behavior of household in financial institutions, highlighting four categories of factors:

- Economic factors – eg. income, availability of the services of the financial market (loans and credits), promotion policy,
- socio-economic and educational factors, eg. the household life cycle, number of household members, having dependent children, level of education and financial knowledge,
- psychosocial factors, eg. personality traits of household members, approach to risk, optimistic or pessimistic attitude to life, lifestyle,
- institutional and legal factors, eg. state macroeconomic policy, affluence of society, availability of credit and loans.

A number of valuable information about the factors determining the households' financial behavior in terms of debt provide the results of various empirical studies. Based on the literature review, it can be generally say that the most important of them are: age, education, gender, socio-professional status and marital status of the head of household, biological type of family and household income.

According to the life cycle hypothesis, one of the most important factor, which fundamentally determines the financial behavior of households in terms of the debt is the age of head of the household. According to the life cycle theory, the demand for credit in the household decreases with age. Households use the loans in the early stages of life, in the middle stages they save and at retirement age they spend accumulated savings (Yilmazer, DeVaney 2005; Wałęga 2010). The assumptions of life cycle hypothesis in the context of debt, have been positively verified in different kind of research (Beer, Shürz 2007; Yilmazer, DeVaney 2005; Carpentier, Van den Bosch 2008; Zajączkowski, Żochowski 2007; Wałęga 2010; Barnes, Young 2003; Mitrakos et al. 2005).

Another factor significantly influencing households' propensity to borrow, as well as determining the ability of debt repayment is the biological type of family and the number of dependent children. Many studies show that are most often taken by persons who are married (Beer, Shürz 2007; Carpentier, Van den Bosch 2008; Yilmazer, DeVaney 2005), which is directly related to the life cycle hypothesis.

Marriage, starting a family, the appearance of offspring, often causes an increase in fixed costs of living, as well as creation of new needs. Relatively low income of young people just entering the labor market involves finding an alternative source of income to finance all of these needs.

The research of Beer and Shürz (2007) shows that the determinant of taking out loans is not so much being married, but the number of dependent children. Households with children report a higher demand for buying their own or larger homes, because the current

housing conditions do not satisfy their needs. Also, the results of Carpentier and Van den Bosch (2008) suggest that households with dependent children, use credits more often than others.

A higher level of education typically translates into higher levels of earnings, and consequently, the greater creditworthiness – thus determine the ability to take out a loan. Many studies (Beer, Shürz 2007; Mitrakos et al. 2005; Wałęga 2010) point to a positive relationship between household debt and the level of education of head of the household. With the increase of the educational level of the head of the household, increase the percentage of households with debt.

Education as a determinant of households' financial behaviors, is strongly linked to the level of their income. Better educated households generally reach higher income, which improves their creditworthiness and allows to take a loan. The positive correlation between income level and debt is confirmed by many studies. Barnes and Young (2003), Mitrakos et al. (2005), Beer and Schürz (2007), as well as Wałęga (2010), Zajączkowski and Żochowski (2007) noted that with the increase of income, increases the percentage of households with credits. The households with lower income not only use credit less often, but also have different motives of taking a loan. This kind of households most often use credit to finance current needs, whereas households with high income use credits to finance durable goods, holidays, buying a house or flat (Wałęga, Wałęga 2016).

### **Aim of the paper, source materials and methodology**

The aim of this study is to identify factors influencing the use of credit by Polish households. The study was based on data from Household Budget Survey in 2011, which was conducted by Central Statistical Office of Poland. The study involved 37375 households, of which 30% used credit. Analysis was performed with the use of logistic regression. Logistic regression is a method used with the dichotomous dependent variable. Independent variables can be both qualitative and quantitative. This method is often used to examine the probability of occurrence of an event Y, provided the occurrence of events  $x_1, x_2, \dots, x_n$  (Stanisz 2007). The logistic function that the logistic regression model is based on has the following formula (Stanisz 2007):

$$P(Y) = \frac{e^{(\beta_0 + \beta_1 X_1 + \beta_k X_k)}}{1 + e^{(\beta_0 + \beta_1 X_1 + \beta_k X_k)}} \quad (1)$$

where:

Y – dichotomous dependent variable

P(Y) – the probability that variable Y will equal 1 for the independent variable value  $X_k$ ,

X – individual socioeconomic traits of the household and the head of the household,

b – structural parameters of the model.

The following dichotomous variables were used in a model:

- The variable assumes the value of 1 for the households with debt,
- The variable assumes the value of 0 for the households without debt.

In order to interpret the logit model an expression called the odds ratio is usually used, which is the ratio between the probability of occurrence of a phenomenon and the probability that the phenomenon will not take place:

$$OR = \frac{P_i}{1 - P_i} \quad (2)$$

In order to estimate the parameters of logistic regression model a set of independent variables was assumed, which characterises different socioeconomic aspects of households. Then the variables were presented in Table 1.

In order to avoid collinearity in the estimation of logit model parameters selected categories of each qualitative variable were omitted, which in consequence led to the generation of a reference group in comparison with which the results were analysed (Daras, Jerzak 2005). The reference group in logit models consists of the households where the heads of households are:

- man
- aged 25-34
- people with higher education
- living in big cities over 500000 people
- staff of private or public sector,
- childless couples,
- with low income (I quintile).

## Research results

This part of paper attempts to identify the strength and direction of impact of socioeconomic traits on the use of credits in Polish households. The results of logit model estimations are shown in table 2. The variable assumes the value of 1 for the households with credit and the value of 0 for the households without credit. Presented model, due to the unbalanced research sample, takes into account the adjusted cut-off point (0,3). In addition, bold typed traits in table 2 are reference categories in relation to which the interpretation of the results are made.

Of the assumed set of potential statistically independent variables the following factors proved to be significant: the householder's sex and age, socio-occupational group, biological type of family and income. As a result of statistical insignificance the following variables were eliminated from the model: level of education and type of place.

Overall classification accuracy of the model is quite high and amounts 61,3%. The model correctly classified 63,3% of households with debt and 60,5% of households without debt.

One of the factors determining the use of credit by the Polish households is the sex of head of the household. The parameters of the model (table 2) show that in the households run by women the chances of having debt are 22% higher than in households, where a man is a head of the household.

The results of logit model (Table 2) also indicate that with the increase of household's age the chances of being indebted decreases. The greatest chances of having debts are in households run by people aged 25-34 years, which constitute a reference group. Then, with the increase of age, the chances of having debt decreased. In the households run by people aged between 45-54 years and 55-64 years, the chances of having debts are almost a quarter lower (about 23%), and in the households of the oldest – over 65 years, more than half lower than in the households from the reference group (25-34 years old). These results are consistent with the assumptions of the life cycle hypothesis. In the early stages of household's life cycle, when the demand for cash is highest, credit allows to smooth consumption, which translates into an increased propensity to borrow. The only exception is a group of households run by the youngest people (below the age of 24) in which the chances of being indebted are about 1/3 lower than in the reference group (25-34 years old). The households of very young people, often did not enter a phase of increased demand for cash yet. Nearly 30% of them are single-person households, which did not start a family yet, and over 12% are young couples without children. In addition, nearly 30% of young people (under 24 years of age) remains on unearned sources, which affects their creditworthiness and limits their ability to borrow.

Another factor significantly differentiating households' financial behavior in aspect of debt is belonging to a socioeconomic group. The highest chances of being indebted was characterized by household of staff of private or public sector, which formed a reference group. Households of self-employed had 36% less chances for being indebted, than the reference group. Almost half lower chances than the reference group were characterized for households of farmers (48%) and households living on unearned sources.

The next determinant of being indebted is a biological type of family. The results of logit model show that the chances of having a credit increase in households with dependent children. Moreover, the probability of having debt increases with the number of dependent children. In the households of couples with one child the chances of having debt were about 58% higher than in the households of childless couples. In turn, in the households of couples with two children the odds ratio was as high as  $\exp(\beta) = 1,67$ ; and in households with three or more children the chances of having debt were the biggest, the odds ratio amounted in them  $\exp(\beta) = 1,72$ . This means that taking out credits by households with at least three children is over 70% more likely than households without children. The relatively high chances of taking out a loan in relation to childless couples, have households of single people with dependent children and other persons ( $\exp(\beta) = 1,68$ ), as well as households of a couples with dependent children and other persons ( $\exp(\beta) = 1.51$ ). Whereas households of a single parent with dependent children have almost 20% higher chance of being indebted than households of childless couples.

**Table 1**  
**Independent variables assumed in logistic regression model**

Trait (Independent variable)	Response categories <sup>a</sup>
Biological type of family	<b>Childless couples</b>
	couples with 1 child
	Couples with 2 children
	Couples with 3 or more children
	single-parent families
	Couples with dependent children and others
	a single person with dependent children and others
	other
	single non-familial
Type of place	<b>city with population over 500 thousand inhabitants and more</b>
	City with population 200 - 499 thousand inhabitants
	City with population 100 - 199 thousand inhabitants
	City with population 20 - 99 thousand inhabitants
	City with population less than thousand inhabitants
	village
Age of the head of household	<b>25-34</b>
	do 24
	35-44
	45-54
	55-64
	over 65
Socio-occupational group	<b>staff of private or public sector</b>
	farmers
	self employed
	retirees and pensioners
	living on unearned sources
Income	<b>I quintile</b>
	II quintile
	III quintile
	IV quintile
	V quintile
Level of education of the head of household	primary school and lower
	vocational/middle school
	secondary
	<b>higher education</b>
Sex of the head of household	<b>man</b>
	woman

<sup>a</sup> Reference categories marked in bold type.

Source: The authors' own compilation based on Household Budget Survey in 2011.

**Table 2**  
**The parameters of logit model (1-with debt)**

Variable <sup>a</sup>	$\beta$	Signifi- cance <sup>b</sup>	Signifi- cance level	Exp(B)
<b>Sex (man)</b>				
woman	0.206	***	0.000	1.229
<b>Age (25-34)</b>		***	0.000	
less than 24	-0.413	***	0.000	0.662
35-44	-0.016		0.677	0.984
45-54	-0.256	***	0.000	0.774
55-64	-0.264	***	0.000	0.768
65 and more	-0.728	***	0.000	0.483
<b>Socio-occupational group (staff of private and public sector)</b>		***	0.000	
farmers	-0.647	***	0.000	0.524
self employed	-0.450	***	0.000	0.638
retirees and pensioners	0.065		0.120	1.067
living on unearned sources	-0.616	***	0.000	0.540
<b>Biological type of family (childless couple)</b>		***	0.000	
couples with 1 child	0.459	***	0.000	1.582
couples with 2 children	0.513	***	0.000	1.670
couples with 3 or more children	0.544	***	0.000	1.723
single-parent families	0.182	*	0.047	1.199
couples with dependent children and others	0.412	***	0.000	1.510
a single person with dependent children and others	0.519	***	0.000	1.680
other	0.043		0.260	1.044
single non-familial	-0.428	***	0.000	0.652
<b>Income (I quintile)</b>		***	0.000	
II quintile	0.352	***	0.000	1.423
III quintile	0.541	***	0.000	1.718
IV quintile	0.826	***	0.000	2.285
V quintile	1.178	***	0.000	3.249
Constant	-1.372	***	0.000	0.254
N	37375			
Cox and Snell's pseudo R <sup>2</sup>	0.070			
Nagelkerke's pseudo R <sup>2</sup>	0.999			
Adjusted <i>cut-off point</i>	0.3			
Overall classification accuracy	61.3%			

<sup>a</sup> The bracketed and bold typed traits are reference categories.

<sup>b</sup> Symbols: \* – significant variables for  $p < 0.05$ ; \*\* – significant variables for  $p < 0.01$ ; \*\*\* – significant variables for  $p < 0.005$ .

Source: as in Table 1.

On the other hand, by far the least likely to be indebted are single non-familial households. The odds ratio for these households amounts  $\exp(\beta)=0,65$ , which means that having a credit in this type of household is more than 1/3 (35%) less likely than in a households of childless couples, which are a reference category. The lower propensity to borrow in households of single persons can be influenced by the following factors:

- almost half of single non-familial households (46%) is represented by the elderly above 65 years old,
- as much as 65% of single non-familial households are the households of retirees and pensioners.
- more than 60% of single non-familial households had very low income (I quintile), and half of them are households of people aged 65 and more. On one hand it is connected with the assumptions of the life cycle hypothesis (a high proportion of older people with low incomes), on the other hand with the lack of creditworthiness of households with the lowest incomes.

The last important and therefore included on the model factor differentiating the financial behavior of Polish households in terms of debt is the level of income. The logit model shows that with the increase in income increases the probability of having debt. The least likely to be indebted were households with the lowest incomes (I quintile), which were a reference group. On the other hand, the use of credits was most likely in households with the highest incomes, in which the chance of being indebted was three times higher than in the reference group.

## Conclusions

The research results show that the use of credits by Polish households are most influenced by the householder's sex, age, socio-occupational status, family type and income in the household. The following conclusions can be made on the basis of the conducted logit analysis:

1. In households run by women the chances of having credit are higher than in households where the head of household is a man.
2. With the increase of the householder's age the chances of being indebted decreases, which is consistent with the assumptions of the life cycle hypothesis.
3. The greatest chances of being indebted, compared to other socio-economic groups were characterized by the staff of private and public sector.
4. The chances of indebtedness increase in households with dependent children. In addition, the probability of having debt increases with the number of dependent children. On the other hand, the least likely to be indebted are single households.
5. With the increase of income increases the probability of having debt. The least likely to be indebted are households with the lowest income.



## Acknowledgement

The project was financed from the resources of The National Science Centre based on the decision number DEC-2012/05/N/HS4/00227

## References

- Barnes S., Young G. (2003), *The rise in US household debt: assessing its causes and sustainability*, "Working Paper", No. 206, Bank of England.
- Beer Ch., Schürz M. (2007), *Characteristic of household debt in Austria. Does household debt pose a threat to financial stability?*, "Monetary Policy and The Economy", No. Q2/07.
- Bywalec Cz. (2009), *Ekonomika i finanse gospodarstw domowych*, Wydawnictwo Naukowe PWN, Warszawa.
- Carpentier S., Van den Bosch K. (2008), *Problematic debts in Belgium: Indicators based on Belgian SILC 2004 and profile of the population with (problematic) indebtedness*, "Working Paper Social Security", No. 7.
- Cynamon B.Z., Fazzari S.M. (2008), *Household Debt in the Consumer Age: Source of Growth-Risk of Collapse*, "Capitalism and Society", Vol. 3, Iss. 2.
- Daras T., Jerzak M. (2005), *Wpływ cech społeczno-demograficznych osób bezrobotnych na możliwość znalezienia pracy. Badanie na podstawie danych BAEL w latach 1993-2003*, „Materiały i Studia”, Zeszyt nr 189, NBP, Warszawa.
- Lea S., Webley P., Walker C. (1995), *Psychological factors in consumer debt: Money management, economic socialization, and credit use*, "Journal of Economic Psychology", No. 16(4).
- Mitrakos T.M., Simigiannis G.T., Tzamourani P.G. (2005), *Intebtedness in Greek households: evidence from a survey*, "Bank of Greece Economic Bulletin", No. 25.
- Raijas A., Lehtinen A., Leskinen J. (2010), *Over-Indebtedness in the Finnish Consumer Society*, "Journal of Consumer Policy", No. 33.
- Stanisz A. (2007), *Przystępny kurs statystyki z zastosowaniem STATISTICA PL na przykładach z medycyny. Tom 3. Analizy wielowymiarowe*, StatSoft, Kraków.
- Świecka B. (2008), *Bankructwa gospodarstw domowych. Perspektywa ekonomiczna i społeczna*, Difin, Warszawa.
- Świecka, B. (2009), *Niewypłacalność gospodarstw domowych. Przyczyny-skutki-przeciwdziałanie*, Difin, Warszawa.
- Wałęga G. (2010), *Determinanty zadłużenia gospodarstw domowych w Polsce w świetle wybranych teorii konsumpcji*, (in:) Dach Z., *Otoczenie ekonomiczne a zachowania podmiotów rynkowych*, PTE, Kraków
- Yilmazer T., DeVaney S.A. (2005), *Household debt over the lifecycle*, "Financial Services Review", No. 14.
- Wałęga A., Wałęga G. (2016), *Bogaci i biedni na rynku kredytowym w Polsce. Analiza porównawcza dla lat 2000–2013*, „Zeszyty Naukowe Uniwersytetu Ekonomicznego w Krakowie”, nr 2(950).
- Zajączkowski S., Źochowski D. (2007), *Obciążenia gospodarstw domowych splatami długu: rozkłady i stress testy – na podstawie badań budżetów gospodarstw domowych GUS*, „Materiały i Studia”, Zeszyt nr 221, NBP, Warszawa.

## Zastosowanie regresji logistycznej w określaniu czynników wpływających na korzystanie z kredytu przez polskie gospodarstwa domowe

### Streszczenie

Dynamiczne zmiany społeczno-gospodarcze, rozwój rynków finansowych, zwiększające się znaczenie mediów i rosnący konsumpcjonizm niewątpliwie przyczyniają się do zmian w zachowaniach finansowych polskich gospodarstw domowych. Stale zmieniająca się rzeczywistość narzuca potrzebę ciągłego powtarzania i aktualizowania badań nad zachowaniami finansowymi gospodarstw domowych. Głównym celem opracowania jest identyfikacja czynników wpływających na korzystanie z kredytów przez polskie gospodarstwa domowe. Do osiągnięcia głównego celu opracowania zastosowano metodę regresji logistycznej. Materiałem źródłowym były poszczególne dane z badania budżetów gospodarstw domowych w roku 2011. Badanie przeprowadził Główny Urząd Statystyczny. Wyniki badań pokazują, że na korzystanie z kredytów przez polskie gospodarstwa domowe najbardziej wpływają: płeć, wiek, status społeczno-ekonomiczny głowy rodziny, typ rodziny i dochody w gospodarstwie domowym.

**Słowa kluczowe:** zadłużenie gospodarstw domowych, determinanty społeczno-ekonomiczne, regresja logistyczna.

**Kody JEL:** D10, D14

## Применение логистической регрессии в выявлении факторов, влияющих на пользование кредитом польскими домохозяйствами

### Резюме

Динамичные социально-экономические изменения, развитие финансовых рынков, растущее значение медиа и развивающееся потребительство несомненно вносят свою лепту в изменения в финансовом поведении польских домохозяйств. Постоянно изменяющаяся действительность вызывает необходимость в постоянном повторении и актуализации изучения финансового поведения домохозяйств. Основная цель изучения – выявить факторы, влияющие на пользование кредитом польскими домохозяйствами. Для достижения основной цели изучения применили метод логистической регрессии. Исходный материал представляли отдельные данные обследования бюджетов домохозяйств ЦСУ. Результаты изучения показывают, что на пользование кредитом польскими домохозяйствами наиболее влияют пол, возраст и социально-профессиональный статус главы домохозяйства, тип семьи и доходы в домохозяйстве.

**Ключевые слова:** задолженность домохозяйств, социально-экономические детерминанты, логистическая регрессия.

**Коды JEL:** D10, D14

Artykuł nadesłany do redakcji w kwietniu 2016 roku

© All rights reserved

Afiliacja:

dr Paulina Anioła-Mikołajczak

Uniwersytet Przyrodniczy w Poznaniu

Wydział Ekonomiczno-Społeczny

Katedra Ekonomiki Przedsiębiorstw Agrobiznesu

ul. Wojska Polskiego 28

60-637 Poznań

tel.: 61 848 73 05

e-mail: [aniola@up.poznan.pl](mailto:aniola@up.poznan.pl)