

Modelling the distribution of loan repayments of households in Poland

Agnieszka Wałęga¹, Grzegorz Wałęga²

Abstract

A rapid growth of household debt in Poland has been observed since the beginning of the twenty-first century. Households more often buy on credit to cover current expenditure and the purchase of durable goods. This situation makes the household budget appears the item “debt repayment”.

The purpose of the article is the choice of theoretical models which reflect as closely as possible the distribution of debt repayments (credit and loans) by households. The study included theoretical distributions used in the analysis of the distribution of wages and income. The most widely used distributions selected for the purpose of the research were: Burr distribution (type III), log-logistic and log-normal distribution. The distributions were tested for consistence of estimation of measures of position, variability and inequality of the empirical distribution of debt repayments. The distributions were tested for consistence of estimation of measures of position, variability and inequality of the empirical distribution of the debt repayments.

The source data consisted of individual information relating to monthly debt repayments in obtained as part of the household budget survey in Poland in 2015.

The results indicate the usefulness of the application of theoretical distributions modelling income distribution in modelling distributions of debt repayments.

Keywords: debt repayments distribution, household indebtedness, income distribution models

JEL Classification: D12, D14, D31

1 Introduction

The twenty first century sees large increase in consumer credits in most European countries. Indebtedness and bankruptcy in households resulting from growing consumer credit use are rapidly increasing in Western societies (Kamleitner and Kirchler, 2006). In Denmark and the Netherlands, household indebtedness is at a record high (gross debt-to-income ratio of households exceeds 200 percent). The volume of household debt burden in new member states of the European Union has doubled, tripled or even quadrupled between 2004 and 2015 (OECD, 2017). Households are increasingly ready to use credits to meet the living standards in developed countries. The same situation is observed in Poland. Between 2004 and 2014, the amount of outstanding debt rose from PLN 98.7 billion to PLN 587.0 billion. In terms of average annual real growth, it is over 19 percent.

¹ Corresponding author: Cracow University of Economics, Department of Statistics, Rakowicka 27, 31–510 Kraków, e-mail: agnieszka.walega@uek.krakow.pl.

² Cracow University of Economics, Department of Microeconomics, Rakowicka 27, 31–510 Kraków, e-mail: grzegorz.walega@uek.krakow.pl.

Credit use by households has various consequences on economic and political levels. For many households, debt repayments become considerable items in the household budget. At the same time, a growing number of borrowers is reporting financial distress and problems with their debt repayments. Currently, personal indebtedness is enjoying public attention. Monetary policymakers have become concerned about the extent of household indebtedness and impact on aggregate economic indicators. While the outstanding amount of household debts can be easily monitored by using administrative data from the financial system, there is no detailed data on the distribution of household debt at the micro level. This paper is an attempt at a closer analysis of this issue.

2 Debt repayment and indebtedness

The phenomenon of widespread household indebtedness is caused by several factors. Consumers rely on credit arrangements because they can implement consumption plans, which at the current level of income are not available to them (Estelami, 2001). Generally, credit use has become socially acceptable. Moreover, in the era of globalization hedonic values as well as an increase in orientation towards the present are important driving forces of consumer behaviour in Western societies (Wood 1998; Watson 2003). Consequently, credit use no longer means investing in one's personal future, but rather has made it acceptable and accessible to acquire luxury goods and other services (Kamleitner and Kirchler, 2006). Also, easy availability of credits facilitates compulsive consumption, which is a characteristic of modern societies (for instance, see Neuner et al., 2005).

The understanding of household repayment distribution is a central question for researchers and policy makers. The economic slowdown in Europe and subprime mortgage crisis in the USA demonstrated that large segments of households are unable to repay their debts (Demyanyk and Hemert, 2009). Naturally, a revived interest in household default behaviour has appeared. The key to the better understanding of the process of excessive debt is to know the distribution of debt repayments. This will recognize appropriate instruments to ensure macroeconomic stability and allow for supporting effectively the households having difficulty repaying their debts. To understand arrears we must analyse both market sides: the households which take loans and the lender's decision to grant the loan (Grant and Padula, 2016). Political debates have also emphasized an important role which is likely to be played by financial market institutions (Duygan and Grant, 2009).

There are two ways of examining household debt. The use of administrative data that are provided by lenders allows for more reliable estimation of debt repayments since the lenders

keep accurate records of credit arrangements, arrears and default. However, the high level of data aggregation has some limitations. These collections are difficult to obtain and they do not allow for a detailed analysis of the behaviour of households on the credit market. Antczak (2013) writes that in order to understand household consumption, the possibility of consumer reaction to the financial shock and the ability to service the debt by households is necessary to analyse the data at the micro level. The best sources of this type of data are cross-sectional surveys, in which one of the modules (sections) is indebtedness. Similar observations for the USA are formulated by Dynan et al. (2003).

In our paper, we focus on debt repayments which reflect the burden of income and determine the current consumption level. The analysis of debt repayment distribution can also provide guidelines on how to identify a household that has too much debt. There have been a few studies that have examined this issue by constructing various indicators of over-indebtedness. In the literature, there are four common indicators that are used to test over-indebtedness: making high repayments relative to income, being in arrears, making heavy use of credit and subjective perception of debt burden (see: Kempson et al., 2004; D'Alessio and Iezzi, 2013). The first two capture the burden imposed by debt repayments and put arbitrary limits on repayments relative to gross income, beyond which they are thought to represent a significant burden for households. As a complement to the above set of indicators, debt repayment distribution is proposed.

Finding the best possible model of debt repayment distribution allows one to carry out a detailed and reliable evaluation of the debt repayment distribution, examine its characteristics thoroughly and compare it with other repayment distributions.

3 Selected models of income distribution

Research conducted by the authors on the shape of household debt repayments in Poland has confirmed that the repayment distribution is characterized by unimodality and positive skewness. Therefore, it seems reasonable to apply theoretical models used for describing distribution of wages (income) to model the household debt repayment distribution.

Research on income distribution has over a hundred years of history in economic theory, which is why this problem has received many theoretical solutions. Using a mathematical function to describe income distribution was first proposed by Pareto (1895). In literature on the subject, the following distributions can often be met: the log-normal distribution popularized in the area of research on income by Aitchison and Brown (1957), distributions of the Burr type - the Burr Type XII (Singh-Maddala) distribution described by Singh and

Maddala (1976) and the Burr III (Dagum) distribution described by Dagum (1977), as well as the log-logistic distribution (Fisk, 1961). Other models that have been used to describe the income distribution included the Gamma distribution, 2-parameter Weibull distribution, Beta distribution and generalized Beta distribution of the second kind described by McDonald (1984), special cases of which are the Burr Type XII and Burr Type III distributions.

In this paper, to describe the household debt repayment distribution, the log-normal distribution, the log-logistic distribution and the Burr type III distribution were used, a brief description of which is presented below³.

The density function of the log-normal distribution is recorded as follows:

$$f(y) = \frac{1}{y\sigma\sqrt{2\pi}} \exp\left\{-\frac{(\ln(y) - \mu)^2}{2\sigma^2}\right\}, \quad (1)$$

where: μ , σ - distribution parameters (μ - location parameter, σ - shape parameter).

In the case of both empirical and theoretical distributions, their analysis is based primarily on the determination of values of distribution characteristics, which are measures of the distribution position, measures of variability, measures of asymmetry and measures of concentration. The expected value can be determined using the formula:

$$E(Y) = \exp\left(\mu + \frac{\sigma^2}{2}\right), \quad (2)$$

and the standard deviation:

$$D(Y) = \sqrt{\exp(2\mu + \sigma^2)[\exp(\sigma^2) - 1]}. \quad (3)$$

Another theoretical distribution used in the analysis of the shape of the debt repayment distribution was the Fisk log-logistic distribution. The density function in the log-logistic distribution can be recorded as follows:

$$f(y) = \frac{\beta \exp(-(\alpha + \beta \ln y))}{y[1 + \exp(-(\alpha + \beta \ln y))]^2}, \quad (4)$$

where α and β are distribution parameters.

The expected value in this distribution is obtained using the formula:

$$E(Y) = \frac{\pi}{\beta \sin\left(\frac{\pi}{\beta}\right)} \exp\left(-\frac{\alpha}{\beta}\right), \quad (5)$$

³ The model description was prepared based on: (Kot, 1999), (Ulman, 2011) and (Salamaga, 2016).

and the standard deviation:

$$D(Y) = \sqrt{(E(Y))^2 \left[\frac{\beta}{\pi} \operatorname{tg} \left(\frac{\pi}{\beta} \right) - 1 \right]}. \quad (6)$$

Another theoretical distribution frequently used to describe the shape of the income distribution is the Dagum (Burr III) distribution. The density function of this distribution can be recorded as follows:

$$f(y) = \frac{cb \exp(-a) y^{-(b+1)}}{[1 + \exp(-a) y^{-b}]^{c+1}}, \quad (7)$$

where a, b, c are distribution parameters.

The ordinary moment of order r in this distribution is expressed as follows:

$$m_r = B \left(1 - \frac{r}{b}, c + \frac{r}{b} \right) \exp \left(-\frac{a}{b} r \right) c, \quad (8)$$

where: $B(p, q)$ is the Euler's beta function. Using the formula (10), the average value of income can be determined as the ordinary moment of order $r = 1$, and the standard deviation - using the formula:

$$D(Y) = \sqrt{m_2 - (m_1)^2}. \quad (9)$$

The fundamental issue for the practical application of theoretical functions as models of wage distributions is the knowledge of distribution parameters. Among many estimation methods, the most frequently used is the maximum likelihood estimation (MLE), which provides consistent, asymptotically unbiased, asymptotically efficient and asymptotically normal estimators of parameters.

To assess the matching rate of the theoretical distribution and the empirical debt repayment distribution, there were applied values of the Akaike information criterion.

It should be emphasized that the fact that theoretical distribution fits specific empirical data better does not mean that this distribution will also be the best model in case of other data (for a different community, in a different period, for a different phenomenon). For this reason, it is worth considering various theoretical distributions to describe empirical debt repayment distributions.

4 Data and research results

The research used data from a household budget survey conducted by the Central Statistical Office in 2015. Members of households participating in the study declared revenues that had involved taking out a credit or a loan and expenses (expenditure) on the repayment of debts.

Since 2005, debt service expenses have been provided broken down by mortgage loans, loans relating to the use of credit cards, as well as other bank borrowings, repayments of debts contracted at institutions other than banks and from private individuals.

Amounts of household income and expenditure that relate to loans and credits are given in the total amount, i.e. without breakdown by the principal part and interest part.

For the purpose of the study, it was assumed that an indebted household was a household that met at least one of the following conditions:

- during the examined period, it took out a mortgage loan, a loan in a bank using a credit card, another loan in a bank, a loan from another institution or a cash loan from private individuals;
- during the examined period, it repaid a principal instalment of and/or interest on a mortgage loan, a loan in a bank taken out using a credit card, another loan taken out in a bank, a loan taken out from another institution, a cash loan taken out from private individuals;

In the study, there was used a repayment amount of the principal instalment of and/or interest on the loan. The analysis included only those households whose amount of debt repayments did not exceed income obtained in the month of the study. In 2015, such households constituted 25.3% of the surveyed households. On the other hand, households whose debt repayments exceeded the amount of the obtained income amounted to 1.2% of the indebted households.

In 2015, the surveyed households spent PLN 748.24 per month, on average, on debt repayment, while half of the households spent not more than PLN 500 per month on debt repayment. However, it should be emphasized that the amount of debt repayments is characterized by a great diversity and significant right-tailed asymmetry (Table 2). There can also be observed considerable disparities of the distribution, measured with the use the Gini coefficient.

In order to estimate parameters, the GRETl econometric package was used. All the estimated parameters of distributions were found to be statistically significantly different from zero – in each case, the test probability is significantly lower than any reasonably assumed level of significance (Table 1).

Based on the value of the likelihood function and the Akaike criterion, it can be stated that for the examined set of data, the distribution of loan repayments was best reflected by the Dagum (Burr III) theoretical distribution, followed by the Fisk (log-logistics) distribution.

Considering the above matching rates of the model and the actual data, the log-normal distribution proved to be worst fitting.

parameter	estimator	st. error	z	p-value
Log-normal distribution				
(Log-likelihood=-79960.96; Akaike=159926; Schwarz=159940)				
μ	6.1650	0.0104	594.22	<0.0001
σ	0.9915	0.0073	136.63	<0.0001
Log-logistic distribution				
(Log-likelihood=-71448.38; Akaike=142901; Schwarz=142915)				
α	-10.8106	0.1090	-99.19	<0.0001
β	1.7467	0.0177	98.60	<0.0001
Burr III distribution				
(Log-likelihood=-71432.22; Akaike=142870; Schwarz=142892)				
a	-17.1700	0.0185	-927.46	<0.0001
b	2.5180	0.0024	1031.79	<0.0001
c	0.4301	0.0007	638.04	<0.0001

Table 1. Estimation of parameters of the log-normal, log-logistic and Burr Type III distributions.

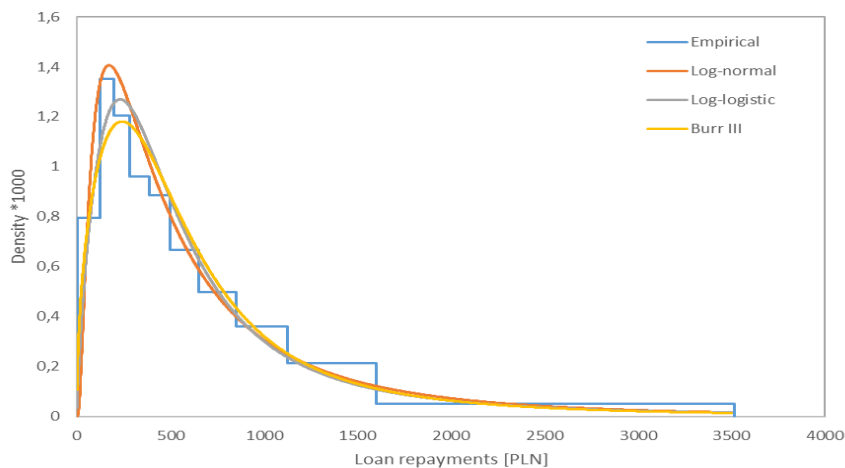


Fig. 1. Empirical distribution and theoretical the log-normal, log-logistic and Burr type III distributions of loan repayments.

Figure 1 shows the course of the density function of the individual theoretical distributions in relation to the empirical decile distribution. In the empirical distribution, the first class of

loan repayments was closed on the left side with the zero value, while the final class with the value of 3515.20 determined based on the knowledge of the average value.

specification	distribution			
	empirical	log-normal	log-logistic	Burr III
Mean	748.24	777.85	900.06	747.44
Median	500.00	475.80	487.49	527.05
Mode	184.61	178.02	231.26	206.58
Coef. of variation	1.10	1.29	0.76	1.40
Gini Index	0.488	0.517	0.573	0.489
Skewnees	0.683	0.596	0.983	0.517

Table 2. Numerical characteristics of the empirical distribution and theoretical distributions of debt repayments.

In Table 2, there are results of calculations of relevant characteristics of the empirical debt repayment distribution. In this case, the mode was determined based on the grouped data (decile distribution), while other characteristics were determined on the basis of individual data. The comparison of characteristics of the empirical distribution and individual theoretical distributions enabled the authors to evaluate how these distributions fit the actual data.

When analysing the numerical characteristics calculated on the basis of the estimated theoretical models, it is hard to state unequivocally which model gives the best results. Undoubtedly, as regards the mean and the Gini coefficient, the Burr III distribution has values most similar to the empirical distribution. In the case of the evaluation of the coefficient of variation level, the log-normal distribution is closer to the empirical value. General principle states that the relative differences between descriptive parameters should not exceed 5%. In our analysis mean, median and mode of log-normal distribution and mean, median and Gini index of decomposition Burr type III comply only with the above principle. The largest differences in characteristics can be found in the log-logistic distribution.

Conclusions

The rising level of household debt and the burden on their budgets repayment of loans affects the long-term processes of consumption. In the context of emerging, not only in the period of crisis, insolvency households particularly interesting is the analysis of the wider distribution

of repayment of loans. Calculated numerical characteristics indicate on positive skewness of repayments of loans in 2015, with an average value repayment in the amount of 750 PLN.

Comparison of characteristics of particular theoretical distributions with the empirical distribution shows that for, each of them has certain drawbacks. In the case of log-normal and log-logistic distributions was overstated the value of the mean, while in the distribution of log-logistic and Burr type III inflated value is mode. In turn, the distribution log-normal and log-logistic undercut the value of the median. The evaluation of model fitting based on descriptive parameters indicates that the best model from the proposed ones is Burr III distribution. However, it should be noted, that methods of assessing goodness of fit yielded inconclusive results. It all makes the research on modelling the distribution of debt repayments need to continue. It would be appropriate to compare methods for parameter estimation and the inclusion of analysis of other models used for example in the analysis of distribution of income.

Acknowledgements

This paper was supported by funds from the National Science Centre, Poland through grant No. 2015/19/D/HS4/02569.

References

- Aitchison, J., & Brown, J. A. (1957). *The lognormal distribution with special reference to its uses in economics*. Cambridge: University Press.
- Antczak, R. (2013). Koncepcja i pomiar zasobności gospodarstw domowych. *Wiadomości Statystyczne*, GUS, 11(630).
- Dagum, C. (2008). A New Model of Personal Income Distribution: Specification and Estimation. *Modeling Income Distributions and Lorenz Curves*, 3-25.
- D'Alessio, G., & Iezzi, S. (2013). Household Over-Indebtedness: Definition and Measurement with Italian Data, Bank of Italy Occasional Paper No. 149.
- Demyanyk, Y., & Hemert, O. V. (2009). Understanding the Subprime Mortgage Crisis. *Review of Financial Studies*, 24(6), 1848-1880.
- Duygan-Bump, B., & Grant, C. (2009). Household debt repayment behaviour: what role do institutions play? *Economic Policy*, 24(57), 107-140.
- Dynan, K., Johnson, K., & Pence K. (2003). Recent changes to a measure of US household debt service, *Federal Reserve Bulletin*, 89(10).

- Estelami, H. (2001). Determinants of discount rates in consumer credit decisions. *Journal of Marketing Theory and Practice*, 9(1), 63-73.
- Fisk, P. R. (1961). The Graduation of Income Distributions. *Econometrica*, 29(2), 171.
- Grant, C., & Padula, M. (2016). The repayment of unsecured debt by European households. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*.
- Kamleitner, B., & Kirchler, E. (2007). Consumer credit use: a process model and literature review. *Revue Européenne de Psychologie Appliquée/European Review of Applied Psychology*, 57(4), 267-283.
- Kempson, E., McKay, S., & Willitts, M. (2004). Characteristics of Families in Debt and the Nature of Indebtedness, Corporate Document Services, DWP Research Report.
- Kot, S. (1999). Teorie genezy rozkładu płac. In *Analiza ekonometryczna kształtowania się płac w Polsce w okresie transformacji* (Kot S.M. ed.). Warszawa-Kraków: Wydawnictwo Naukowe PWN.
- Mcdonald, J. B. (1984). Some Generalized Functions for the Size Distribution of Income. *Econometrica*, 52(3), 647.
- Neuner, M., Raab, G., & Reisch, L.A. (2005). Compulsive buying in maturing consumer societies: an empirical re-inquiry. *Journal of Economic Psychology*, 26(4), 509-522.
- OECD (2017), Household debt (indicator). (Accessed on 08 February 2017)
- Salamaga M. (2016). *Badanie wpływu metody estymacji teoretycznych modeli rozkładu dochodów na jakość aproksymacji rozkładu dochodów mieszkańców Krakowa*, Zeszyty Naukowe Uniwersytet Ekonomiczny w Krakowie 3(951), 63-79.
- Sierminska, E. (2014). Indebtedness of households and the cost of debt by household type and income group, Research note 10/2014, European Commission.
- Singh, S. K., & Maddala, G. S. (2008). A Function for Size Distribution of Incomes. *Modeling Income Distributions and Lorenz Curves*, 27-35.
- Ulman, P. (2011). *Sytuacja ekonomiczna osób niepełnosprawnych i ich gospodarstw domowych w Polsce*. Kraków: Wydawnictwo Uniwersytetu Ekonomicznego w Krakowie.
- Watson, J. J. (2003). The relationship of materialism to spending tendencies, saving, and debt. *Journal of Economic Psychology*, 24(6), 723-739.
- Wood, M. (1998). Socio-economic status, delay of gratification, and impulse buying. *Journal of Economic Psychology*, 19(3), 295-320.