Bank i Kredyt 53(5), 2022, 475-496

BANK I KREDYT
BANK & CREDIT

A number of capital structure models presented even in prominent papers are estimated with incorrect estimators

Mieczysław Kowerski*

Submitted: 23 February 2022. Accepted: 8 August 2022.

Abstract

The capital structure is most often measured using a debt ratio, which usually takes values in the interval (0; 1). This makes the linear regression model a linear probability model. A basic shortcoming of such a model is that fitted values for some observations can be less than zero, which is inconsistent with the definition of the debt ratio.

In the vast majority of articles, the authors did not pay attention to these defects and if they noticed them, they proposed solutions that cannot be considered fully satisfactory. A simple solution, known for many decades, is the application of the logit transformation of the dependent variable, which ensures that the fitted values of the debt ratio are in the interval (0; 1).

The aim of the study is to draw researchers' attention to the inappropriateness of using linear probability models and to show the advantages of the logit models of debt ratios.

The considerations were illustrated by models of debt ratios estimated for the companies listed on the WSE in 1998–2019.

Keywords: capital structure models, debt ratio, linear probability model, logit model, unbalanced panel

JEL: C24, C51, G32

* Academy of Zamość; e-mail: mieczyslaw.kowerski@akademiazamojska.edu.pl; ORCID: 0000-0002-2147-2037.

1. Introduction

Franco Modigliani and Merton Miller's seminal article (1958) initiated theoretical and empirical research on the factors of the capital structure. Although more than 60 years have passed since then, the "capital structure puzzle" formulated by Myers, who wrote "How do firms choose their capital structures? … the answer is, we don't know" (Myers 1984, p. 585), is still valid. There is still no single theory explaining what factors determine the structure of capital, and there are no unambiguous empirical results, either. On the other hand, a large number of partial studies were carried out, in which the factors determining capital structures in different countries, different periods and for different groups of companies and sectors of the economy were defined.

Most often, the capital structure was measured by the debt ratio, and to determine its factors linear regression models were used (pooled models, panel models with fixed and random effects, and dynamic models), in which the dependent variable was the debt ratio and the explanatory variables were microeconomic, macroeconomic and, over time, behavioural factors.

The debt ratio is a variable limited at the lower limit which takes values equal to or greater than zero. Analyses show that in most of the conducted studies, debt ratio values were in the interval (0; 1). This makes the linear regression model of the debt ratio a linear probability model. Such a model has at least three shortcomings. First of all, the fitted values for some observations can be less than zero, which is inconsistent with the definition of the debt ratio. Secondly, in such a model there is heteroscedasticity of random disturbance. Third, random disturbance in a linear probability model does not have a normal distribution.

Only a few authors pointed to the limitations of the dependent variable and the possibility of negative fitted values. However, their proposed solutions (tobit model, winsorizing or truncation of data) do not fully overcome the disadvantages of the linear probability model. One of the viable solutions based on the logit transformation of the dependent variable was proposed by Magri (2010). However, Magri's work received no response from researchers, who still used methods specific to linear regression models, in which dependent variables take continuous values.

Thus, the vast majority of capital structure models presented even in prominent papers are estimated with incorrect estimators. On the other hand, different factors determine the structure of capital in different periods of time, different countries (different legal systems) or sectors of the economy. Therefore, these types of models are needed and will continue to be created. So, it is important that the models of the capital structure measured by the debt ratio are properly estimated. On the basis of poorly estimated models it is not possible to draw correct conclusions. The described situation is surprising because there are relatively simple methods to solve this problem which have been known for many decades.

Therefore, the aim of the study is to draw researchers' attention to the inappropriateness of using linear probability models of the debt ratio as well as to show that the logit transformation known since the 1940s and the logit linear model known since the 1960s can be a simple way to solve the problem of negative fitted values of debt ratios and estimator inefficiency.

The rest of the article is as follows. In the second chapter, the literature on the subject is reviewed, with special attention being paid to the problems of the modelling of the debt ratio. The third chapter formulates the problem of the incorrectness of the methods used to estimate the debt ratio. In the fourth chapter, a solution to the problem of negative fitted values in linear model is presented. The fifth chapter is an illustration of the problem discussed in the article. The estimated models are intended to show the incorrectness of using linear probability models (the negative fitted values of the debt ratio) and the benefits of using the logit transformation of a dependent variable. They were estimated on the most up-to date data available of the unbalanced panel of 112 companies listed on Warsaw Stock Exchange in 1998–2019.

2. Literature review

In 1958, Franco Modigliani and Merton Miller (MM) formulated the theorem (known as MM Proposition I) that the value of a company in a perfectly competitive capital market does not depend on the structure of capital but on investment decisions. To illustrate Proposition I, they estimated pooled regressions of values of companies measured by total earnings after taxes to the market value of all securities (x) on financial structure measured by the market value of bonds and preferred stock to market value of all the securities (d) of 43 large electric utilities in 1953 and 42 oil companies in 1947–1948 (Modigliani, Miller 1958, pp. 281–284):[1]

$$\begin{array}{llcccl} & x & = & 5.3 & + & 0.006d \qquad r = 0.12 \\ \text{Electric utilities:} & \text{Standard error:} & & & & (0.008) \\ & \text{t-statistic:} & & & & (0.75) \end{array} \tag{1}$$

$$\begin{array}{llcccl} & x & = & 8.5 & + & 0.006d \qquad r = 0.04 \\ \text{Oil companies:} & \text{Standard error:} & & & & (0.024) \\ & \text{t-statistic:} & & & & (0.25) \end{array} \tag{2}$$

In the case of both types of companies, the coefficient on the variable d: the structure of capital was statistically insignificant, from which it can be concluded that the dependence of value on the structure is irrelevant, a conclusion consistent with Proposition I.

MM’s theorems were called the capital structure irrelevance theory.

The publication of MM attracted great interest of both academics and practitioners. There were also critical works. The criticism primarily concerned unrealistic assumptions about a perfect market, to which MM soon responded by removing the assumption of non-taxation (Modigliani, Miller 1963).

At the same time, extensive research aimed at supplementing MM findings began.

The theoretical basis for these studies were the trade-off (substitution) theory, the pecking order theory, the agency theory and the free cash flow theory.

Initially, based on the trade-off theory, the target debt ratio trading off costs and benefits of leverage were sought. Graham and Harvey’s survey evidence (2001) shows that, indeed, 81% of firms consider the target debt ratio when making their debt decisions (Flannery, Rangan 2006). In 1984, Myers proposed to clarify the relationship between the value of a company and its capital structure using an updated version of Donaldson’s (1961) pecking order theory, according to which investments

[1] According to the author, it is worth presenting these models, which are the first empirical example of the verification of the theory. This presentation is also interesting in the context of the progress that financial econometrics has made.

are financed first with internally generated funds and if internal funds are insufficient, the firm issues debt. Equity is issued only as a last source of money. According to the agency theory, the capital structure is determined by agency costs, i.e. costs due to conflicts of interest between the managers and shareholders (Jensen, Meckling 1976). "Holding constant the manager's absolute investment in the firm, increases in the fraction of the firm financed by debt increase the manager's share of the equity and mitigate the loss from the conflict of interest" (Harris, Raviv 1991, p. 300). Moreover, as pointed out by Jensen (1986), since debt forces the firm to pay out cash, it reduces the amount of "free" cash available to managers to transfer firm resources to their own personal benefit. This mitigation of the conflicts between managers and equity holders constitutes the benefit of debt financing.

The free cash flow theory (Jensen 1986) suggests that managers have a tendency to overinvest if the threat of bankruptcy is not serious enough (empire-building). This moral hazard problem can be mitigated if the firm uses debt as a disciplinary device. If a manager spends funds inefficiently, the firm will not be able to generate enough cash to cover existing debt and the probability of bankruptcy will increase. In this case, the probability that managers will lose their jobs increases (Miglo 2020, p. 7).

After the works of Myers (1984) and Myers and Majluf (1984), the emphasis changed toward exploring the factors that influence the capital structure.

The structure of capital can be measured in many ways, but as noted by Rajan and Zingales (1995, p. 1429), the most appropriate and most commonly used ratio is the total debt to the total book value of assets, or the debt ratio. Similar conclusions were later reached, among others, by Baker and Wurgler (2002), Chang and Dasgupta (2009), Fama and French (2002), and Kayhan and Titman (2007).

The factors determining the capital structure were most often identified using linear regression models, in which the dependent variable was most often the debt ratio and the explanatory variables were microeconomic, macroeconomic and, over time, behavioural factors. The first comprehensive review of the factors determining the capital structure that had been taken into account in previous studies was carried out by Harris and Raviv (1991). According to these authors, "there is the association of capital structure with fixed assets, non-debt tax shields, growth opportunities, firm size, volatility, advertising expenditures, research and development expenditures, bankruptcy probability, profitability and uniqueness of the product" (Harris, Raviv 1991, p. 334). Four years later, Rajan and Zingales (1995, p. 1451), using the results of Harris and Raviv's analysis in their research, adopted four variables: firm size (log of assets), tangibility (ratio of fixed assets to the book value of total assets), investment opportunities (market to book value of equity) and profitability (proxied by EBIT to assets), which over time became the starting point in most estimated regression models. Of course, subsequent authors did not stop at these four factors, expanding the set of explanatory variables. Kumar, Colombage and Rao (2017) reviewed 167 papers published between 1972 and 2013 in various peer-reviewed reputed journals and specified 27 factors which were used to explain capital structure changes at the firm, industry and economy level. The authors' meta-analysis showed that at the 99 percent confidence level the relationship between the capital structure and profitability, liquidity, age, non-debt tax shield and risk are negative and the relationship between the capital structure and tangibility, and size and growth are positive (Kumar, Colombage, Rao 2017, p. 124). Nejada and Wasiuzzaman (2013, p. 468) point to the particular importance of dividends in explaining the volatility of the capital structure.

Nehrebecka, Białek-Jaworska and Dzik-Walczak (2016, pp. 94–98) analysed 21 articles published between 1997 and 2012 and identified 75 explanatory variables that the authors of these articles had used to explain the volatility of the capital structure.

Initially, the relationship between the debt ratio and its determinants was treated as pooled regression and estimated using the ordinary least squares (OLS) method. The development of research resulted in the emergence of larger and larger data sets of panel nature; hence linear panel models were used to analyse the relationships. Since the 1970s, the Fama-McBeth method (1973) has enjoyed great popularity among financial market researchers. This method consisted of estimating the coefficients of the pooled models each year on the basis of cross-sectional data and then testing the significance of the mean values of the coefficients from the entire analysed period using the t-Student test. Fama and French (2002, pp. 11–12) suggested that in the inference process the critical value of the t-statistic should be increased 2.5 times due to the autocorrelation in time. Examples of the application of this method for estimating the linear regression of the debt ratio can be found in the works of Welch (2004) and Flannery and Rangan (2006).

Over time, the Fama-McBeth method was replaced with the panel models with fixed (LSDV estimator) or random (GLS estimator) specific effects estimation methods (Maddala 2006, p. 644–648). In 1981 a method of estimating dynamic panel models was proposed by Anderson and Hsiao, and in 1991 Arellano and Bond proposed the generalized moments method (GMM).

One example of dynamic capital structure models are partial adjustment models, in which the debt ratio in year t depends on the debt ratio in year $t - 1$ and selected control variables that make it possible to estimate the speed of adjustment and the target capital structure (Chang, Dasgupta 2009; Drobetz, Wanzenried 2006; Iliev, Welch 2010; Mukherjee, Mahakud 2010; Chen, Huang, Lee 2022).

It should be noted that the debt ratio is a variable limited from the bottom. Rajan and Zingales (1995, p. 1452) were among the first to point out that in the case of limited dependent variables, the use of linear regression models can lead to negative fitted values of dependent variables. Therefore, they estimated coefficients using a censored tobit model. Tobit models were used, among others, by Akhtar (2005), Magri (2010), and Espinosa et al. (2012) in their analyses of the capital structure.

As a solution to the problem of negative fitted values in linear models of the debt ratio, some authors have proposed winsorizing (Nejda, Wasiuzzaman 2013; Haque, Varghese 2021; Khaki, Akin 2020; Stradomski, Schmidt 2020; Welch 2013) or, less frequently, truncation (Rajan, Zingales 1995).

Magri (2010, p. 448) rightly noted that the debt ratio is usually in the interval (0; 1) and suggested that while estimating the factors determining the debt ratio, a model in which the dependent variable is transformed into logits in order to have a variable that is unbounded should be used. She also presented the results of estimating the long-term debt ratio of non-public Italian firms, using fixed effects panel logit models.

Given that a typical problem of dynamic panel models of the capital structure is endogeneity, the GMM is increasingly used to estimate them (Abdullah, Tursoy 2021; de Miguel, Pindado 2001; Espinosa et al. 2012; Liaqat et al. 2021; Mardones, Cuneo 2020). This method yields consistent estimators of the coefficients in situations in which the OLS estimator is inconsistent. It is preferable to use the GMM in two stages with a robust estimator, since it is more efficient and reduces the loss of information (Arellano, Bover 1995).

Dao and Ta (2020) based on a meta-analysis of capital structure models included in 245 articles covers research during 2012–2019, with a data set dated from 2000 to 2017 published in 34 papers, calculated that pooled OLS is a dominant method, used in 40.8% of the selected papers, the fixed-effects model ranks second, with 30.2%, the random effects model ranks third, with 26.1%, and only 2.9% of the studies use GMM as their preferable method.

3. The problem of the incorrectness of the methods used to estimate the debt ratio

The debt ratio takes values equal to or greater than zero. Usually, these are also values less than unity.[2] This makes the linear regression model of the debt ratio a linear probability model. Such a model has at least three shortcomings. First, as Goldberger wrote as early as in 1964,[3] the adoption of such a model does not preclude fitted values for some observations from being outside the interval (0; 1), which is contradictory to the definition of the dependent variable and to the interpretation of its expected value as probability. It is not possible for the debt ratio to be negative. Also, introducing to analysis observations with a debt ratio greater than unity is at least debatable – more on that later. Secondly, in such a model there is heteroscedasticity of random disturbance, which means that estimators of OLS coefficients will not be effective (Maddala 2006, p. 369). Thirdly, random disturbance in a linear probability model does not have a normal distribution, which also makes OLS not fully effective (Gruszczyński 2001, p. 56) and may negatively affect the quality of the statistical inference process. Therefore, the assumption of the unlimited values of the dependent variable in linear models of the debt ratio is incorrect and inference based on such models will always be inaccurate.

4. Proposal to solve the problem of negative fitted values in linear model

As shown in literature review, only a few authors noted the incorrectness of using linear models to estimate the factors determining the debt ratio, proposing three solutions:

1. Replacement of the linear model with the censored regression model developed by James Tobin (1958), which Goldberger (1975, p. 325) gave the commonly used name "tobit model," derived from the name of its author.

2. Replacement of the outlier observations (top and bottom) with more reasonable ones (winsorizing) or their removal (truncation).

3. Entering the logit transformation of the value of the dependent variable.

Unfortunately, the first two proposals do not remove all the errors resulting from the use of linear probability models. The tobit model, although it takes into account the fact of the limitation of the dependent variable, does not protect against negative fitted values. Similarly, winsorizing and truncation, although they probably reduce the likelihood of negative fitted values, do not completely exclude them. It is interesting that none of the authors who applied one of the two proposed solutions checked whether they actually received positive fitted values as a result.

On the other hand, the logit transformation of the dependent variable (debt ratio) proposed by Magri (2010) seems to be appropriate. It makes it possible to convert the value of the debt ratio from interval (0; 1) to interval $(-\infty; \infty)$ and thus use the linear models (pooled models, panel models with fixed or random specific effect or dynamic panel models) to determine the factors of the capital structure. However, Magri's work received no response from researchers who still did not pay attention to the fact of the limited values of the debt ratio and the estimative consequences associated with it,

---

2 The author did not conduct a comprehensive study, but in more than 40 articles (about half of the analyzed ones) in which the authors gave the minimum and maximum debt ratio in three-quarters, the maximum debt ratio was less than 1.

3 Goldberger A.S. (1964), *Econometric Theory*, John Wiley and Sons. The second Polish edition was used in Goldberger (1975, p. 321).

and used methods specific to linear models, in which dependent variables take continuous values. Interestingly, even the authors who had read Magri's work, as evidenced in their citations, did not use logits to estimate debt ratio models (Abad, Sánchez-Ballesta, Yagü 2017; Domenichelli 2015). Perhaps the reason for the lack of any application of the solution proposed by Magri was that in her article the author did not present more extensively the negative consequences of using linear models in the case of the limited values of the dependent variable but only stated that the debt ratio should be transformed into logits (Magri 2010, p. 448).

Therefore, it is worth discussing the issue of the use of logit transformation in the modelling of the debt ratio in more detail.

The logit transformation requires that the debt ratio takes values in the open interval from 0 to 1. Thus, observations with zero values of the debt ratio as well as observations with ratios equal to 1 or higher are not taken into account. From the point of view of analyses that aim to find significant factors that (in the long term) have a relatively stable relationship with the debt ratio, which is, in the author's opinion, the basic goal of this research, the exclusion of observations with a debt ratio equal to or greater than 1 will improve the results. In the case of debt exceeding the value of assets, the owner's equity is negative, which makes it impossible for the company to pay off all its liabilities, and this can even lead to bankruptcy (Jaworski, Czerwonka 2022, p. 200). On the basis of data from entities threatened with bankruptcy, it is difficult to draw correct conclusions regarding the functioning of all enterprises in the long term.

The situation with zero debt ratios is slightly more complicated. Analysing a sample of industrial companies from 20 developed countries in 1988–2011 (31,820 companies with a total of 315,464 firm-year observations), Bessler et al. (2013) observed an increasing share of zero-debt companies from only 8.47% of all companies in 1988 to 25.70% in 2011 (12.71% in civil law countries and 34.38% in common law countries). Detailed research led the authors to conclude that the increasing share of zero-debt firms involved financially constrained companies, which do not have the chance to obtain debt financing. These firms tend to be smaller, younger, more risk prone, and less profitable, but sometimes have big investment opportunities. Constrained firms do not possess sufficient debt capacity and thus (have to) maintain a zero-debt policy for longer periods of time. On the basis of information from such companies, it is difficult to draw correct conclusions about the functioning of all companies in the long term, so they can be removed from the set of analysed companies.[4] However, the zero-debt firms group is not homogeneous and there is a rather small subsample of firms that deliberately choose to adopt a zero-leverage policy. These financially unconstrained firms are more profitable, distribute higher dividends, and are older as well as larger than their constrained zero-debt peers (Bessler et al. 2013, p. 197). For example, Apple had no debt in 2012–2014. The company's earnings were steadily growing between 2005–2012 and many analysts and managers spoke about its excessive liquidity problems (Miglo 2020, p. 2).

It is difficult to determine to which group a given company belongs only on the basis of zero values of the debt ratio. Therefore, in the authors' opinion, observations with a zero debt ratio could be winsorized,[5] and variables explaining such a situation (an inverted U-shaped variable or a product of two variables) should be introduced to the models.

[4] Some of them are often not accepted for research due to their failure to meet the assumed criteria (too small, too young, with negative equity).

[5] This does not preclude a separate, detailed analysis of such companies; see e.g. Bessler et al. (2013), Miglo (2020).

The logit transformation converts the value of the debt ratio *DR* belonging to the interval (0; 1) into an interval (-∞; ∞) using the formula:

$$Y = logitDR = ln\frac{DR}{1-DR} \tag{3}$$

where *DR* – debt ratio $\in (0;1)$.

Having values of the dependent variable in the form of logits, we can estimate linear regression models. The fitted value of debt ratio for *i*-th company in year *t* ($\widehat{DR_{it}}$) is calculated on the basis of the fitted value of Y_{it} (which we denote $\widehat{Y}_{it}$) using the definition of logarithm:

$$\widehat{DR_{it}} = \frac{\exp\left(\widehat{Y}_{it}\right)}{1+\exp\left(\widehat{Y}_{it}\right)} \tag{4}$$

Estimated fitted values are always in the interval (0; 1), which is consistent with the assumptions made about the value of the debt ratio.

5. Illustration of the proposed method

The primary purpose of this chapter is to show empirically the problem identified in the previous chapters, which is that in linear probability models, dependent variables can take fitted negative values. The results of the estimated linear probabilities models were compared with linear models in which the dependent variable was logit transformed. This part of the work is only intended to illustrate the proposed method and does not pretend to fully explain the factors determining the debt ratio on the Polish capital market. Anyway, such a noteworthy attempt has appeared recently (Jaworski, Czerwonka 2022), although the authors also did not notice the issue of the possibility of negative fitted values of debt ratios. According to the author, the illustrations were made on the basis of properly specified models, estimated on the basis of a well-chosen panel of companies and appropriate explanatory variables.

5.1. Procedure for verifying the validity of the proposed method

To verify the validity of the proposed method, panel linear models of the capital structure were applied. Using GRETL (Cottrell, Lucchetti 2021; Kufel 2018), fixed effects panel models (with the de-meaned data method, which is numerically equivalent to the LSDV method), random effects panel models (with the GLS method), a tobit panel model and dynamic panel models (with the two-step GMM method) of two variables DR and logitDR were estimated.

Panel models of capital structure with fixed specific effects:

$$DR_{it} \text{ or } Y_{it} = \alpha_0 + \alpha_i + \beta_1 X_{1it(t-1)} + \beta_2 X_{2it(t-1)} + \ldots + \beta_k X_{kit(t-1)} + \varepsilon_{it} \tag{5}$$

Panel models of capital structure with random specific effects:

$$DR_{it} \text{ or } Y_{it} = \alpha_0 + \beta_1 X_{1it(t-1)} + \beta_2 X_{2it(t-1)} + \ldots + \beta_k X_{kit(t-1)} + \gamma_{it} \tag{6}$$

Dynamic panel (partial adjustments) capital structure models with fixed specific effects:

$$DR_{it} \text{ or } Y_{it} = \alpha_0 + \alpha_i + \beta_0 Y_{it-1} + \beta_1 X_{1it(t-1)} + \beta_2 X_{2it(t-1)} + \ldots + \beta_k X_{kit(t-1)} + \varepsilon_{it} \tag{7}$$

Dynamic panel (partial adjustments) capital structure models with random specific effects:

$$DR_{it} \text{ or } Y_{it} = \alpha_0 + \beta_0 Y_{it-1} + \beta_1 X_{1it(t-1)} + \beta_2 X_{2it(t-1)} + \ldots + \beta_k X_{kit(t-1)} + \gamma_{it} \tag{8}$$

where:

DR_{it} – capital structure in *i*-th company at the end of the year *t*,
Y_{it} – logit of capital structure $(logitDR_{it})$ in *i*-th company at the end of the year *t*,
$X_{jit(t-1)}$ – the value of *j*-th factor (explanatory variable) in *i*-th company in year *t* or in year *t* – 1 depending on the definition of the variable,
ε_{it} – random disturbance,
$\gamma_{it} = \alpha_i + \varepsilon_{it}$ – random disturbance which is the sum of fixed effects and random effects (Maddala 2006, p. 645),
$t = 1, 2, \ldots, n$ – number of years,
$i = 1, 2, \ldots, m$ – number of companies,
$j = 1, 2, \ldots, k$ – number of explanatory variables.

Using the Welch's F test for the group intercept difference and the Hausman test, it was checked whether there were fixed or random effects and which estimator was more effective.

In the dynamic panel models, the Sargan test was used to assess the correctness of the selection of instrumental variables and the Arellano-Bond test to verify hypotheses on the autocorrelation of the random disturbance of the first and second order.

To assess the goodness-of-fit of models to empirical data, the linear correlation coefficients between fitted and actual values of *DR* were used.

The significance of the estimated coefficients was assessed using robust standard errors (Cottrell, Lucchetti 2021, pp. 181, 199–202) and their signs for model pairs (*DR* and *logitDR*) were compared.

A logit model is a linear model of logit in regard to explanatory variables. Hence, individual coefficients are derivatives of logits relative to the corresponding explanatory variables and their estimates are interpreted similarly to those in a linear regression model. The sign of the estimated coefficient on the variable X_j in the logit model determines the direction of the association of this variable with the variable *Y*.

The significance of the estimated coefficients was assessed and their signs for model pairs (*DR* and *logitDR*) were compared.

In all the estimated models, the fitted values of dependent variables were analysed. In the case of logit models, the fitted values were calculated on the basis of formula 4.

5.2. Data

To illustrate the proposed method, data were used from companies listed on the Warsaw Stock Exchange. For this purpose, an unbalanced panel was constructed, which included domestic companies, excluding banks, forming part of the WIG20, mWIG40 and sWIG80 indices at the end of 2019 and having been listed for at least 3 years prior, i.e. companies that had entered the WSE no later than 31 December 2016. It was decided that the data would go back to 1998. Thus, for companies that entered the WSE before that year, data were collected starting from 1998. For companies that entered the WSE in 1998 or later data were collected since their entry. The developed database is therefore an unbalanced panel of 1509 observations and does not include observations with negative owner's capital (Kowerski, Charkiewicz 2021).

The panel consists of 112 companies included in the three basic WSE indices, i.e. 80% of such companies. The companies covered by the panel accounted for only 27.93% of all (401) domestic companies listed on the WSE at the end of 2019, but their share in the capitalization of the domestic market amounted to 60.5%, and the share in the book value of equity was 64.2%. Therefore, these companies were significantly larger than other companies listed on the WSE (t-Student test).

Consequently, the panel included liquid companies, mostly with a better economic and financial situation and a longer stock exchange history than other companies listed on the main market. Although this may cause the phenomenon of survival bias, according to the author, if we want to formulate and verify hypotheses regarding factors permanently determining the debt ratio, companies with the adopted characteristics should be used. Information coming from young companies in poor condition can only disrupt the results.

5.3. Variables

The following were adopted as dependent variables:

1. *DR* – debt ratio (total assets – owner's capital)/total assets,
2. $Y = logitDR$.

The author did not conduct a separate numerical procedure for the selection of variables, but as explanatory variables he adopted variables that most often occurred in the previously discussed studies (profitability, risk, growth opportunities and propensity to pay dividends). Compare to the most frequently identified significant determinats of the debt ratio used in studies of Polish enterprises (Cwynar, Cwynar, Dankiewicz 2015; Hajduk 2018; Jaworski, Czerwonka 2022) selected set do not include the size of the enterprise, which turned out to be insignificant in any model, and liquidity due to the lack of data. However, as is often the case with other financial categories, the relationship between the debt ratio and growth opportunities was proposed in the shape of an inverted U (Graham, Leary 2011, p. 6; Kowerski 2013; Kowerski, Bielak 2018). This should capture the phenomenon of very low debt ratios for both financially constrained companies and mature and healthy companies.

The following explanatory variables were measured:

1. ROA – profitability – the value of the net profit in year t to the average of total assets at the end of year $t - 1$ and t (%);

2. AssetGrowth – growth opportunities – asset growth rate; assets at the end of year t to assets at the end of year t – 1) / assets at the end of year t – 1 (%);

3. AssetGrowth2 – asset growth square;

4. Risk – the coefficient of the volatility of share prices calculated as the quotient of the difference between the maximum and minimum share price to the maximum price in year t;

5. DividProp – previous propensity to pay dividends – the number of payments made to the maximum possible number of payments from the year following the debut on the WSE till year t – 1(%) (Kowerski 2011, pp. 111–112).

5.4. Results of the estimation of debt ratio models

In the analysed panel, the minimum observed *DR* value was 0.0046, and the maximum value was 0.9996, which means that no data truncation or winsorization had to be carried out for the logit transformation.

Of course, the distribution of the *DR* variable was not normal, which was confirmed by the Doornik-Hansen and Shapiro-Wilk tests. The Breusch-Pagan test showed that the pooled OLS model is not adequate and that the random effects alternative is more reliable. The test for the group intercept difference showed that intercepts differ significantly. The Hausman test indicated that the random effects model is consistent. Therefore, a model with random effects is more appropriate. Nonetheless, the author decided to present the results of the estimation of both models.

In the dynamic model of *DR* and *logitDR*, Sargan's test showed that the whole set of instruments is exogenous, which means that there is no endogeneity problem. In turn, the Arellano-Bond test showed that autocorrelations of the disturbance of the first order were statistically significant, but autocorrelations of the second order were already insignificant, which is consistent with the assumptions of the GMM method.

The linear correlation coefficients between fitted and actual values of *DR* in all models were statistically significant at the level of 0.05 (t-Student test on the significance of the correlation coefficient), with the highest value obtained for the dynamic panel logit model (0.8941).

In all the estimated models, the signs of coefficients on the same explanatory variables were identical. Only in the dynamic model of the *logitDR* were all the coefficients significant ($p < 0.0001$). In the other models, one of the coefficients was insignificant at the level of 0.05.

As predicted, all the models of the *DR* variable had observations with negative fitted values of *DR*: 3 in the panel fixed model, 6 in the panel random model, 18 in the dynamic model, and 1 in the tobit model. Particularly large was the spread between the minimum and maximum fitted values in the dynamic panel model of *DR* (from -0.6931 to 1.2696). Of course, in logit models all fitted values are in the interval (0; 1). This result confirms the thesis put forward in the work that linear probability models are not able to accurately establish the factors determining the debt ratio, since some of its estimated fitted values are negative, which is contrary to the definition of this variable. For example, how to interpret the debt ratio of -9.4%? Only models with a logit-transformed debt ratio do not have this disadvantage. On the other hand, the lack of differences in the parameter signs of both pairs of models (*DR* and *logitDR*) does not allow us to completely dismiss the results obtained in so many studies published even in well-known world journals. It can be said that these results are less accurate, because the used estimators are ineffective.

In the conducted study, the dynamic panel model of *logitDR* turned out to be the best and it will be used to interpret the results. And although the aim of the work was not to find the best set of determinants of the capital structure in Poland, the estimated model can be considered the right direction of inquiry.

The signs of the estimated coefficients of the dynamic panel model seem to properly reflect the relationship between the selected explanatory variables and the debt ratio.

The negative value of the parameter at ROA is consistent with the pecking order theory – more profitable companies first benefit from the profits generated, so their debt ratios are low. Companies with low profitability must "support" themselves with debt by increasing the debt ratio.

A higher debt ratio is a characteristic of riskier companies, which results from higher costs of the debts incurred. Also, a higher debt ratio is a characteristic of companies with a greater propensity to pay dividends. According to the agency's theory, dividends "force" companies to replace "lost" equity with external capital. In addition, companies paying dividends are usually companies with high creditworthiness, which allows them to raise external capital more cheaply.

The negative value of the coefficient on the AssetGrowth2 variable confirms the relationship between growth opportunities and the debt ratio in the shape of an inverted U. In dynamic terms, this means that initially, as the company's growth opportunities increase, so does its debt ratio – the company develops using external capital, but after reaching the maximum value of the debt ratio the company is already mature enough to afford less and cheaper debt. In cross-section terms, this means that companies with the fewest investment opportunities, i.e. mature companies that invest much less and pay large dividends, and financially constrained companies are characterized by the lowest debt ratio. Also companies financially constrained with large investment opportunities, i.e. young and developing, which, however, may have problems with obtaining external funds, are characterized by very low debt ratios. Maximum debt ratios are recorded for companies with average investment opportunities.

6. Conclusions

The literature review clearly confirms that the vast majority of the models of the capital structure measured by the debt ratio published in Polish and foreign journals was estimated using incorrect estimators. Typically, these are linear pooled models or panel models with fixed or random effects, which, when the debt ratio takes values from the interval (0; 1), are actually linear probability models. When estimating such models with linear regression, assuming the continuous value of debt ratio, estimators will not be effective and some observations of fitted values of dependent variable are negative, which is inconsistent with the definition of the debt ratio. And from an interpretative point of view, it does not matter how many negative fitted values there are. There are more cases of negative fitted values when there are more observations of the actual debt ratio being close to zero.

One solution to the problem may be the use of the logit transformation of the debt ratio. In the case of panel data, the logit dynamic model seems to be the most appropriate, especially since it allows verification of one of the basic theories trying to explain the "capital structure puzzle" – the trade-off (substitution) theory. This is well illustrated by the example of modelling the debt ratio of companies listed on the WSE.

Of course, the discussed method has some drawbacks, which relate primarily to the possibility of companies having zero values of debt ratios (especially in the case of constrained firms but also for a small group of companies in a very good condition). Zero values are not logit transformed and such observations must be replaced with others (winsorizing) or removed from the panel. In both cases, this distorts the final results.

On the other hand, as shown in the presented example, the differences in estimations of linear probability models and logit models do not have to be large (the lack of differences in the coefficient signs of both pairs of models – *DR* and *logitDR*), which does not allow us to completely dismiss the results obtained in so many studies published even in well-known world journals. However, according to the author, it is more appropriate to use models with the logit transformation of the debt ratio.

References

Abad D., Sánchez-Ballesta J.P., Yagü J. (2017), The short-term debt choice under asymmetric information, *SERIEs Journal of Spanish Economic Association*, 8(3), 261–285.

Abdullah H., Tursoy T. (2021), Capital structure and firm performance: evidence of Germany under IFRS adoption, *Review of Managerial Science,* 15(2), 379–398.

Akhtar S. (2005), The determinants of capital structure for Australian multinational and domestic corporations, *Australian Journal of Management*, 30(2), 321–341.

Anderson T.W., Hsiao Ch. (1981), Estimation of dynamic models with error components, *Journal of the American Statistical Association*, 76(375), 598–606.

Arellano M., Bond S. (1991), Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations, *Review of Economic Studies*, 58(2), 277–297.

Arellano M., Bover O. (1995), Another look at the instrumental variable estimation of error components models, *Journal of Econometrics*, 68(1), 29–51.

Baker P., Wurgler J. (2002), Market timing and capital structure, *The Journal of Finance,* 57(1), 1–32.

Bessler W., Drobetz W., Haller R., Meier I. (2013), The international zero-leverage phenomenon, *Journal of Corporate Finance*, 23, 196–221.

Chang X., Dasgupta S. (2009), Target behavior and financing: How conclusive is the evidence, *The Journal of Finance,* 64(4), 1767–1796.

Chen Ch.-C., Huang H.H., Lee Ch.I. (2022), Supply chain, product pricing, and dynamic capital structure, *International Review of Economics & Finance*, 80, 938–952, https://doi.org/10.1016/j.iref.2022.03.001.

Cottrell A., Lucchetti R. (2021), *Gretl User's Guide. Gnu Regression*, *Econometrics and Time-series Library*, Wake Forest University, Università Politecnica delle Marche.

Cwynar A., Cwynar W., Dankiewicz R. (2015), Studies on firm capital structure determinants in Poland: an integrative review, *eFinanse. Financial Internet Quarterly*, 11(4), 1–22.

Dao B.T.T., Ta T.D.N. (2020), A meta-analysis: capital structure and firm performance, *Journal of Economics and Development*, 22(1), 111–129.

Domenichelli O. (2015), An empirical investigation of the debt maturity of Italian family firms, *International Journal of Finance and Accounting*, 4(5), 281–292.

Donaldson G. (1961), *Corporate Debt Capacity: A Study of Corporate Debt Policy and the Determinants of Corporate Debt Capacity*, Harvard Graduate School of Business Administration.

Drobetz W., Wanzenried G. (2006), What determines the speed of adjustment to the target capital structure?, *Applied Financial Economics*, 16(13), 941–958.

Espinosa C., Maquieira V.C., Vieito J.P., Gonzalez M. (2012), Capital structures in developing countries: the Latin American case, *Investigación Económica*, 71(282), 35–54.

Fama E., MacBeth J. (1973), Risk, return, and equilibrium: empirical tests, *Journal of Political Economy*, 81(3), 607–636.

Fama E., French K. (2002), Testing trade-off and pecking order predictions about dividends and debt, *Review of Financial Studies,* 15(1), 1–33.

Flannery M., Rangan K. (2006), Partial adjustment toward target capital structures, *Journal of Financial Economics*, 79(3), 469–500.

Goldberger A.S. (1975), *Teoria ekonometrii*, Państwowe Wydawnictwo Ekonomiczne.

Gruszczyński M. (2001), *Modele i prognozy zmiennych jakościowych w finansach i bankowości*, Oficyna Wydawnicza Szkoły Głównej Handlowej w Warszawie.

Haque S., Varghese R. (2021), *The Covid-19 impact on leverage and financial fragility*, Internationally Found Working Paper, 21(265), https://ssrn.com/abstract=4026495.

Hajduk A. (2018), Determinanty struktury kapitałowej przedsiębiorstwa w świetle statystycznej analizy parametrycznej, *Zarządzanie i Finanse (Journal of Management and Finance),* 16(2), 103–115.

Harris M., Raviv A. (1991), The theory of capital structure, *The Journal of Finance*, 56(1), 297–355.

Iliev P., Welch I. (2010), *Reconciling estimates of the speed of adjustment of leverage technical report*, http://ssrn.com/abstract=1542691.

Jaworski J., Czerwonka L. (2022), *Struktura kapitału przedsiębiorstw w Polsce*, Wydawnictwo CeDeWu.

Jensen M.C. (1986), Agency costs of free cash flow, corporate finance and takeovers, *American Economic Review*, 76(2), 323–329.

Jensen M.C., Meckling W. (1976), Theory of the firm: managerial behaviour. Agency costs and ownership structure, *Journal of Financial Economics*, 3(4), 305–360.

Kayhan A., Titman S. (2007), Firms' histories and their capital structures, *Journal of Financial Economics*, 83(1), 1–32.

Khaki R., Akin A. (2020), Factors affecting the capital structure: new evidence from GCC countries, *Journal of International Studies*, 13(1), 9–27.

Kufel T. (2018), *Ekonometria. Rozwiązywanie problemów z wykorzystaniem programu GRETL*, Wydawnictwo Naukowe PWN.

Kowerski M. (2011), *Ekonomiczne uwarunkowania decyzji o wypłatach dywidend przez spółki publiczne*, Konsorcjum Akademickie.

Kowerski M. (2013), Możliwości inwestycyjne a skłonność do płacenia dywidend, *Bank i Kredyt*, 44(6), 623–646.

Kowerski M., Bielak J. (2018), Self-selection models in determination of target dividend payout ratio of real estate domestic companies quoted on Warsaw Stock Exchange, *Barometr Regionalny. Analizy i Prognozy*, 16(1), 99–112.

Kowerski M., Charkiewicz A. (2021), Panelowa baza danych o spółkach notowanych na Giełdzie Papierów Wartościowych w Warszawie dla potrzeb analiz finansowych, *Facta Simonidis*, 14(1), 161–179.

Kumar S., Colombage S., Rao P. (2017), Research on capital structure determinants: a review and future directions, *International Journal of Managerial Finance*, 13(2), 106–132.

Liaqat I., Khan M.A., Popp J., Oláh J. (2021), Industry, firm, and country level dynamics of capital structure: a case of Pakistani firms, *Journal of Risk and Financial Management*, 14(9), 1–17.

Maddala G.S. (2006), *Ekonometria*, Wydawnictwo Naukowe PWN.

Magri S. (2010), Debt maturity choice of nonpublic Italian firms, *Journal of Money, Credit and Banking*, 42(2/3), 443–463.

Mardones J.G., Cuneo G.R. (2020), Capital structure and performance in Latin American companies, *Economic Research – Ekonomska Istraživanja*, 33(1), 2171–2188.

Miglo A. (2020), Zero-debt policy under asymmetric information, flexibility and free cash flow considerations, *Journal of Risk and Financial Management*, 13(12), 296, 1–25.

de Miguel A., Pindado J. (2001), Determinants of capital structure: new evidence from Spanish panel data, *Journal of Corporate Finance*, 7(1), 77–99.

Modigliani F., Miller M. (1958), The cost of capital, corporation finance, and the theory of investment, *American Economic Review*, 48(3), 261–297.

Modigliani F., Miller M. (1963), Corporate income taxes and the cost of capital: a correction, *American Economic Review*, 53(3), 433–443.

Mukherjee S., Mahakud J. (2010), Dynamic adjustment towards target capital structure: evidence from Indian companies, *Journal of Advances in Management Research*, 7(2), 250–266.

Myers S.C. (1984), The capital structure puzzle, *The Journal of Finance*, 39(3), 575–591.

Myers S.C., Majluf N.S. (1984), Corporate financing and investment decisions when firms have information the investors do not have, *Journal of Financial Economics*, 13(2), 187–221.

Nehrebecka N., Białek-Jaworska A., Dzik-Walczak A. (2016), *Źródła finansowania przedsiębiorstw. Stan badań i ich metaanaliza*, Difin.

Nejada N.R., Wasiuzzaman S. (2013), The empirical analysis of capital structure determinants – evidence from Malaysia, *World Academy of Science, Engineering and Technology*, 74, 466–474.

Rajan R., Zingales L. (1995), What do we know about capital structure? Some evidence from international data, *The Journal of Finance*, 50(5), 1221–1260.

Stradomski M., Schmidt K. (2020), Firm specific determinants of capital structure in European advanced developing countries, *Bank i Kredyt*, 51(3), 263–292.

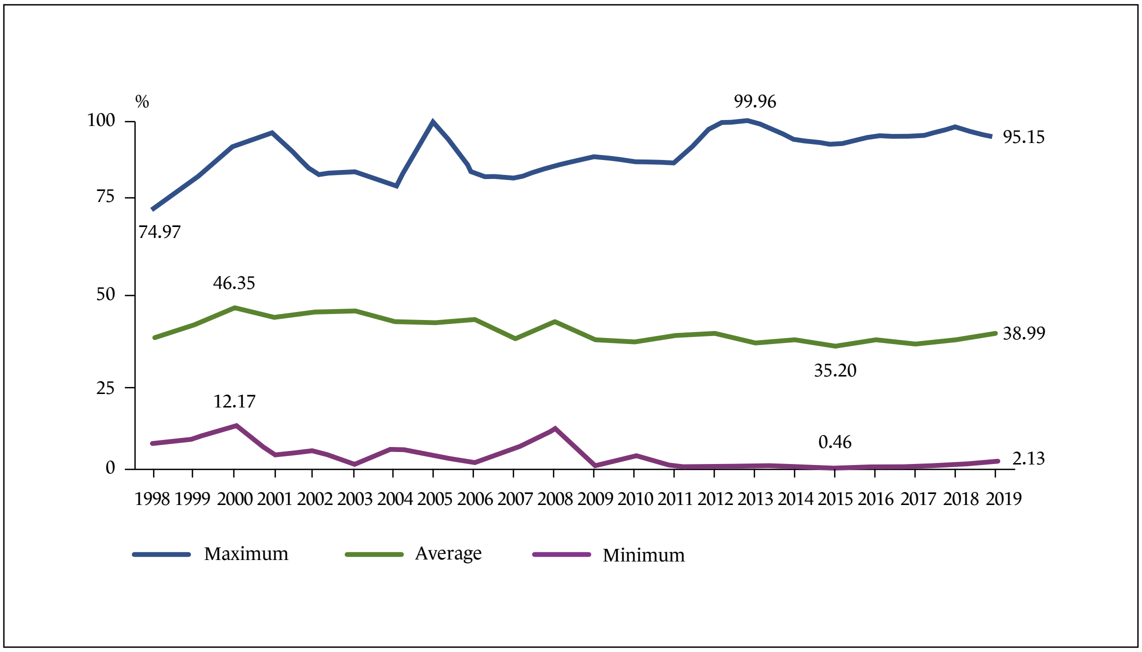
Tobin J. (1958), Estimation of relationships for limited dependent variables, *Econometrica*, 26(1), 24–36.

Welch I. (2004), Capital structure and stock returns, *Journal of Political Economy*, 112(1), 106–132.

Welch I. (2013), A critique of recent quantitative and deep-structure modelling in capital structure research and beyond, *Critical Finance Review*, 2(1), 131–172.

Appendix

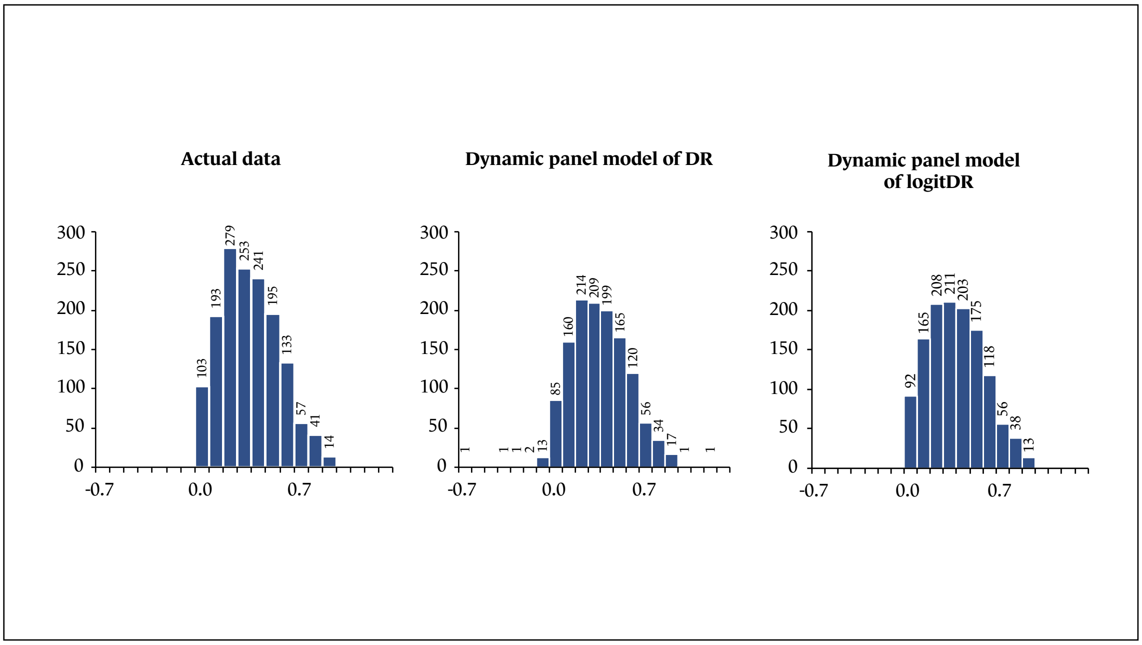
Figure 1
Changes in the average, minimum and maximum values of the debt ratio of panel companies in 1998–2019



Source: own elaboration.

Figure 2

Distribution of actual data and fitted values in dynamic models of *DR* and *logitDR*



Source: own elaboration.

Table 1
Results of the estimation of debt ratios models

	Panel fixed models				Panel random models				Dynamic panel models				Tobit model	
	DR		Y = LogitDR		DR		Y = LogitDR		DR		Y = LogitDR		DR	
	coefficient	p	coefficient	p	coefficient	p	Coefficient	p	Coefficient	p	Coefficient	p	Coefficient	p
Const	0.3573	< 0.0001	-0.8026	< 0.0001	0.3455	< 0.0001	-0.8871	< 0.0001	0.0001	0.7019	0.0086	< 0.0001	0.3300	< 0.0001
DR(-1), LogitDR(-1)									0.6472	< 0.0001	0.6330	< 0.0001		
ROA	-0.0043	< 0.0001	-0.0237	< 0.0001	-0.0042	< 0.0001	-0.0235	< 0.0001	-0.0052	< 0.0001	-0.0259	< 0.0001	-0.0039	< 0.0001
AssetGrowth	0.0005	0.0011	0.0025	0.0025	0.0005	0.0008	0.0025	0.0022	0.0020	< 0.0001	0.0111	< 0.0001	0.0005	0.0130
AssetGrowth2	-0.0000007	0.0011	-0.0000035	0.0021	-0.0000007	0.0009	-0.0000035	0.0018	-0.000008	< 0.0001	-0.00005	< 0.0001	-0.0000009	0.0068
Risk	0.0822	0.0042	0.4963	0.0066	0.0856	0.0020	0.5221	0.0037	0.0041	0.6492	0.2818	< 0.0001	0.1513	< 0.0001
DividProp	0.0004	0.2790	0.0029	0.0903	0.0004	0.2392	0.0027	0.0879	0.0008	< 0.0001	0.0069	< 0.0001	0.0003	0.0844
Sargan test Chi2(209)									82.3550	0.9999	93.9449	0.9999		
Arellano--Bond test z statistic AR(1)									-4.6744	< 0.0001	-4,3832	< 0.0001		
AR(2)									-0.1231	0.9020	-0.8501	0.3953		
Correlation between fitted values and DR	0.8310		0.8288		0.2750		0.2799		0.8900		0.8941		0.2827	
Number of observations	1509		1509		1509		1509		1279		1279		1509	

Source: own calculations.

Table 2
Characteristics of the fitted values of DR

	Panel fixed models		Panel random models		Dynamic panel models		Tobit model
	DR	Y = logitDR	DR	Y = logitDR	DR	Y = logitDR	DR
Average	0.3871	0.3815	0.3766	0.3437	0.3903	0.3960	0.3871
Median	0.3645	0.3545	0.3782	0.3408	0.3791	0.3865	0.3866
Max	0.9217	0.9188	0.7267	0.7716	1.2696	0.9999	0.7144
Min	-0.0942	0.0107	-0.0307	0.0531	-0.6931	0.0010	-0.0793
Number of negative values	3	0	6	0	18	0	1
Number of values > 1	0	0	0	0	2	0	0
Variability index (%)	44.26	48.10	15.18	18.89	56.10	53.42	15.06

Source: own calculations.

Wiele modeli struktury kapitału opublikowanych nawet w znaczących czasopismach zostało nieprawidłowo oszacowanych

Streszczenie

Strukturę kapitału najczęściej mierzy się za pomocą stopy zadłużenia, która zazwyczaj przyjmuje wartości z przedziału (0; 1). Sprawia to, że liniowy model stopy zadłużenia jest liniowym modelem prawdopodobieństwa. Podstawową wadą takiego modelu jest to, iż obliczone na jego podstawie wartości teoretyczne zmiennej objaśnianej mogą być ujemne, co jest niezgodne z definicją stopy zadłużenia. Tylko nieliczni autorzy zwracali uwagę na tę wadę. Proponowane przez nich rozwiązania (model tobitowy, obcięcie lub zastąpienie danych ekstremalnych) nie usuwają jednak w pełni wad liniowego modelu prawdopodobieństwa. Właściwe rozwiązanie problemu, polegające na transformacji logitowej zmiennej objaśnianej, zaproponowała Magri (2010). Jej praca pozostała jednak niezauważona przez badaczy, którzy nadal stosowali metody właściwe dla modeli liniowych, ze zmienną objaśnianą o dowolnych wartościach. A przecież na podstawie źle oszacowanych modeli trudno o prawidłowe wnioski.

Z tego względu celem pracy jest zwrócenie uwagi prowadzących badania na niewłaściwość stosowania liniowych modeli prawdopodobieństwa stopy zadłużenia, a także szersze zaprezentowanie możliwości zastosowania transformacji logitowej (znanej od lat 40. ubiegłego wieku) oraz modelowania logitowego (znanego od lat 60.) do rozwiązania problemu ujemnych wartości teoretycznych zmiennej objaśnianej.

W pracy do określenia czynników determinujących strukturę kapitału mierzoną stopą zadłużenia zaproponowano logitowe modele panelowe ze stałymi i losowymi efektami specyficznymi oraz logitowy dynamiczny model panelowy. Wyniki tak oszacowanych modeli porównano z oszacowaniami panelowych liniowych modeli prawdopodobieństwa ze stałymi i losowymi efektami specyficznymi oraz panelowego modelu dynamicznego stopy zadłużenia.

Do zilustrowania zaproponowanej metody wykorzystano dane spółek notowanych na Giełdzie Papierów Wartościowych w Warszawie. W tym celu skonstruowano panel niezbilansowany. Weszły do niego spółki krajowe, z wyłączeniem banków, które na koniec 2019 r. wchodziły w skład indeksów WIG20, mWIG40 i sWIG80 i były notowane przynajmniej przez 3 lata. Zdecydowano, że dane będą sięgać 1998 r. Dla spółek, które weszły na GPW przed tym rokiem, zgromadzono więc dane począwszy od 1998 r. Dla spółek, które weszły na GPW w 1998 r. lub później, wykorzystano dane od roku ich debiutu giełdowego. Bazę tworzy 112 spółek z 1509 obserwacjami.

Jako zmienne objaśniające przyjęto rentowność mierzoną stopą zwrotu z aktywów, możliwości wzrostu spółki mierzone stopą wzrostu aktywów, ryzyko mierzone ilorazem różnicy między ceną maksymalną i minimalną do ceny maksymalnej w ciągu roku oraz skłonność do wypłat dywidend w poprzednich latach.

We wszystkich liniowych modelach prawdopodobieństwa stopy zadłużenia wystąpiły jej ujemne wartości teoretyczne – 3 w modelu ze stałymi efektami, 6 w modelu z losowymi efektami, 18 w modelu dynamicznym i jedna w modelu tobitowym. Dobrze ilustruje to postawioną w pracy

tezę, iż powszechnie stosowane liniowe modele prawdopodobieństwa mogą dawać wyniki sprzeczne z istotą kategorii finansowych. Oczywiście w oszacowanych modelach logitowych nie było ujemnych teoretycznych wartości stopy zadłużenia. Znaki oszacowanych parametrów przy tych samych zmiennych w liniowych modelach prawdopodobieństwa i modelach logitowych były jednakowe. Nie pozwala to na całkowite dezawuowanie wyników uzyskanych w tak wielu badaniach, opublikowanych nawet w znanych światowych czasopismach. Jednak zdaniem autora właściwsze jest stosowanie modeli ze stopami zadłużenia przekształconymi za pomocą transformacji logitowej.

Oczywiście omówiona metoda nie jest pozbawiona pewnych wad, które dotyczą przede wszystkim możliwości przyjmowania zerowych wartości stóp zadłużenia, zwłaszcza przez spółki o słabej kondycji, ale również przez niewielką grupę spółek o bardzo dobrej kondycji. Zerowe wartości nie podlegają transformacji logitowej i muszą być zastąpione innymi lub usunięte z panelu. W obu przypadkach obciąża to ostateczne wyniki.

Słowa kluczowe: modele struktury kapitału, stopa zadłużenia, liniowy model prawdopodobieństwa, model logitowy, niezbilansowany panel