CEFAGE
centro de estudos e formação avançada em gestão e economia
universidade de évora

CEFAGE-UE Working Paper
2011/28

Functional form issues in the regression analysis of financial leverage ratios

***Joaquim J.S. Ramalho*[1], *J. Vidigal da Silva*[2]**

[1] *Department of Economics and CEFAGE-UE, Universidade de Évora*

[2] *Department of Management and CEFAGE-UE, Universidade de Évora*

Functional form issues in the regression analysis of financial leverage ratios

Joaquim J.S. Ramalho*
Department of Economics and CEFAGE-UE, Universidade de Évora

J. Vidigal da Silva
Department of Management and CEFAGE-UE, Universidade de Évora

Abstract

Linear models are typically used in the regression analysis of capital structure choices. However, given the proportional and bounded nature of leverage ratios, models such as the tobit, the fractional regression model and its two-part variant are a better alternative. In this paper, we discuss the main econometric assumptions and features of those models, provide a theoretical foundation for their use in the regression analysis of leverage ratios and review some statistical tests suitable to assess their specification. Using a dataset previously considered in the literature, we carry out a comprehensive comparison of the alternative models, finding that in this framework the most relevant functional form issue is the choice between a single model for all capital structure decisions and a two-part model that explains separately the decisions to issue debt and, conditional on the first decision, on the amount of debt to issue.

Keywords: capital structure, zero leverage, fractional regression model, tobit, two-part model.

JEL Classification: G32, C25.

*Corresponding author. Address for correspondence: Joaquim J.S. Ramalho, Department of Economics, Universidade de Évora, Largo dos Colegiais 2, 7000-803 ÉVORA, Portugal. E-mail: jsr@uevora.pt. Tel.: +351 266740894. Fax: +351 266740807.

1 Introduction

The regression analysis of the financing decisions of firms has been a key theme in applied corporate finance for more than thirty years. Typically, empirical studies on capital structure decisions use linear models to examine how a given set of potential explanatory variables (X) influences some leverage ratio (Y). However, leverage ratios (e.g. debt to capital or total assets) possess two basic characteristics that may render the linear model inadequate for explaining them: (i) by definition, they are bounded on the closed interval [0,1];[1] and (ii) many firms have null leverage ratios.[2] Therefore, regression models that take into account (at least one of) those characteristics of leverage ratios are potentially a better alternative for modelling the conditional mean of leverage ratios, $E(Y|X)$, which is usually the main interest in applied work. Because many firms have no debt, a popular alternative to linear modelling has been the use of the tobit model for data censored at zero (e.g. Rajan and Zingales, 1995; Wald, 1999; Cassar, 2004). Other alternatives include the fractional regression model proposed by Papke and Wooldridge (1996), which was specifically developed for dealing with fractional or proportional response variables such as leverage ratios, and its two-part variant (Ramalho and Silva, 2009), which treats separately the decisions on using debt or not (using a binary choice model) and, conditional on this decision, the decision on the relative amount of debt to issue (using a fractional regression model).

Tobit, fractional and two-part fractional regression models are based on very distinct assumptions about the data generating process of leverage ratios, i.e., how firms make their capital structure decisions. For example, the tobit model assumes that the accu-

[1] Actually, this is strictly valid only for market leverage ratios. Indeed, because some firms may have negative book values of equity, book leverage ratios may display values higher than one. However, given that firms with negative book values of equity are typically excluded from empirical studies on capital structure (e.g. Baker and Wurgler, 2002; Byoun, 2008; Lemmon, Roberts and Zender, 2008) or their leverage ratios are re-coded to one (e.g. Faulkender and Petersen, 2006), book leverage ratios are also, in practical terms, effectively restricted to the unit interval in most cases.

[2] For example, Strebulaev and Yang (2007), Byoun, Moore and Xu (2008), Bessler, Drobetz, Haller and Meier (2011) and Dang (2011) report that an average of 8.9% of U.S. firms (sample period: 1962-2003), 12.2% of U.S. firms (1971-2006), 11.0% of G7 firms (1988-2008) and 12.2% of U.K. firms (1980-2007), respectively, had zero outstanding debt. In the last year of the sample period, those figures rise to 18.2%, 22.6%, 14.2% and 23.7%, respectively, which shows that the zero-leverage phenomenon has been increasing over time.

mulation of observations at zero is the result of a censoring problem (e.g. the firms with zero debt would really like to have negative debt) and should be modelled as such, the fractional model ignores the causes of that accumulation and treats the zero observations as any other value (as the linear model also does), and the two-part fractional model assumes that the zero and the positive leverage ratios are generated from different, independent mechanisms. Thus, while in the fractional (and also in the linear) regression model it is only relevant to calculate $E(Y|X)$, in the other cases choosing a functional form for $E(Y|X)$ automatically defines expressions for the probability of a firm using debt, $\Pr(Y>0|X)$, and the conditional mean of leverage ratios for firms that do use debt, $E(Y|X,Y>0)$, which may be also of interest for researchers. Moreover, while in the tobit model each explanatory variable is restricted to influence in the same direction $E(Y|X)$, $\Pr(Y>0|X)$ and $E(Y|X,Y>0)$, the two-part fractional model allows the covariates to affect in independent ways each one of those quantities. Finally, while the tobit model requires distributional assumptions, the two-part fractional model typically only requires such assumptions for its binary component, and the fractional model does not require them at all.

Because they are based on very different assumptions, we may suspect that the results produced by each model may be also very distinct. If that is the case, then using an incorrect functional form for $E(Y|X)$ may generate misleading conclusions about how financial leverage decisions are made. To choose the most appropriate specification in empirical work, practitioners may resort to theoretical arguments (e.g. if the zeros are interpreted as the result from a unique firm value-maximizing decision, then using a two-part fractional regression model makes no sense because this model assumes that the zeros result from two independent decisions) and/or use econometric specification tests. Nevertheless, most empirical studies on capital structure decisions assume *a priori* a given specification for $E(Y|X)$ and do not test the assumptions underlying the model chosen or justify theoretically their option.

The main aim of this paper is the analysis of the main functional forms issues that may arise when studying the determinants of capital structure choices. In particular, we discuss the econometric specification, estimation and evaluation of linear, tobit, fractional and two-part fractional regression models and provide a theoretical foundation for their use in this context. As little is known about the consequences of using an incorrect model in the analysis of capital structure decisions, we use a data set of Portuguese firms previously considered in the literature (Ramalho and Silva, 2009) to compare the results yielded by each model at various levels: (i) the significance, direction and magnitude of the marginal effects of covariates; and (ii) the prediction of leverage ratios.

The most closely related paper to ours is Ramalho and Silva (2009). In fact, in addition to using the same data set, they have considered the same regression models. However, since Ramalho and Silva (2009) were mainly interested in justifying the use of a two-part fractional regression model to study the financial leverage decisions of Portuguese firms, the other specifications were only briefly addressed and the focus of the empirical analysis was the assessment of several hypotheses about capital structure choices. In contrast, in this paper we deal with all models in a comprehensive and balanced way, establish clear links between all of them and capital structure theories, propose statistical tests suitable to assess the model assumptions and focus the empirical illustration on comparisons across models. The ultimate aim of this paper is to provide a sound econometric basis for analyzing leverage ratios bounded in the unit interval.

The paper is organized as follows. Section 2 reviews the four alternative regression models that we consider in this paper for analyzing financial leverage decisions. Section 3 shows why some capital structure theories imply the use of particular regression models and propose some econometric tests for assessing the specification of each model. Section 4 compares the alternative regression models using Ramalho and Silva's (2009) dataset. Finally, section 5 contains some concluding remarks.

2 Regression models for capital structure choices

In this section, we discuss the main characteristics of linear, tobit, fractional and two-part fractional regression models, stressing their advantages and drawbacks when applied to the regression analysis of fractional response variables, in general, and leverage ratios, in particular.

2.1 Linear model

Most empirical studies of capital structure have used linear regression models to explain observed leverage ratios; see *inter alia* Prasade, Green and Murinde (2005) and Frank and Goyal (2008), which summarize the main methodologies used in capital structure empirical research. However, the linearity assumption

$$E(Y|X) = X\beta, \tag{1}$$

where β denotes the vector of parameters of interest, is unlikely to hold in our framework. Indeed, in linear models, the effect on $E(Y|X)$ of a unitary change in the explanatory variable X_j is constant throughout its entire range,

$$\frac{\partial E(Y|X)}{\partial X_j} = \beta_j, \tag{2}$$

which is not compatible with both the bounded nature of leverage ratios and the existence of a mass-point at zero in their distribution. Moreover, the conceptual requirement that the predicted values of Y lie in the interval $[0,1]$ is not satisfied by the linear model. Note also that the use of the linear model in this framework requires the computation of heteroskedasticity-robust standard errors, since the conditional variance of Y is in general a function of its conditional mean: the former must change as the latter approaches either boundary.

While it is straightforward to compute heteroskedasticity-robust standard errors and, to some extent, the problem of assuming constant marginal effects may be overcome by augmenting the model with nonlinear functions of X (which, however, do not correspond to the standard practice in empirical capital structure studies), the predicted values from a linear regression model can never be guaranteed to lie in the unit interval.[3] Therefore, any sensible description of the true data generating process of leverage ratios cannot be based on the use of linear models. Nevertheless, in this paper, as the linear regression model has been used in most previous empirical capital structure studies, we investigate in which conditions, if any, may the linear model constitute a reasonable approximation for that data generating process.

2.2 Tobit model

As a typical random sample of firms contains many firms that do not use debt, some authors have opted for using a tobit approach for data censored at zero for modeling leverage

[3] Basically, the drawbacks of using linear specifications for modeling fractional data are similar to the drawbacks of using the linear probability model for describing binary data.

decisions. The tobit model was originally proposed for cases where the explanatory variables are fully observed for all sampling units but the variable of interest is incompletely observed (only its positive values are observed, while its non-positive values are, just by convenience, represented by zeros). Thus, instead of observing Y^*, the latent variable of interest, we observe Y, which is defined as follows: $Y = Y^*$ for $Y^* > 0$ and $Y = 0$ otherwise. It is also assumed that Y^* has a normal distribution, that there exists a linear relationship between Y^* and the covariates, $E\left(Y^*|X\right) = X\beta$, and that the error term of the latent model, $u = Y^* - E\left(Y^*|X\right)$, is homoskedastic.

While in early applications of tobit models the main interest was inference on Y^*, currently the tobit model is also often used for explaining the influence of X on Y (see e.g. Wooldridge, 2002, pp. 517-521). In the regression analysis of leverage ratios, the main goal of any empirical capital structure study is effectively to explain *observed* leverage ratios, not the *latent* ones. Thus, the specification of the tobit model that is relevant for our purposes, which is implied by the assumptions made above for Y^*, is given by

$$E\left(Y|X\right) = \Phi\left(\frac{X\beta}{\sigma}\right)X\beta + \sigma\phi\left(\frac{X\beta}{\sigma}\right), \tag{3}$$

where $\Phi\left(\cdot\right)$ and $\phi\left(\cdot\right)$ denote the standard normal distribution and density functions, respectively, and σ is the standard deviation of u. The tobit model also implies that

$$\Pr\left(Y > 0|X\right) = \Phi\left(\frac{X\beta}{\sigma}\right) \tag{4}$$

and that

$$E\left(Y|X, Y > 0\right) = X\beta + \sigma\frac{\phi\left(\frac{X\beta}{\sigma}\right)}{\Phi\left(\frac{X\beta}{\sigma}\right)}; \tag{5}$$

see Wooldridge (2002) for details. Given the distributional assumption made for Y^*, the parameters β and σ are estimated by the maximum likelihood method.

The overall partial effects of unitary changes in X_j on Y are given by:

$$\frac{\partial E\left(Y|X\right)}{\partial X_j} = \beta_j\Phi\left(\frac{X\beta}{\sigma}\right). \tag{6}$$

We may also compute the marginal effects of a covariate X_j over the probability of using debt and over the conditional mean leverage ratios of firms that do use debt, which are given by, respectively,

$$\frac{\partial \Pr\left(Y > 0|X\right)}{\partial X_j} = \frac{\beta_j}{\sigma}\phi\left(\frac{X\beta}{\sigma}\right) \tag{7}$$

and

$$\frac{\partial E\left(Y|X,Y>0\right)}{\partial X_j}=\beta_j\left\{1-\frac{\phi\left(\frac{X\beta}{\sigma}\right)}{\Phi\left(\frac{X\beta}{\sigma}\right)}\left[\frac{X\beta}{\sigma}+\frac{\phi\left(\frac{X\beta}{\sigma}\right)}{\Phi\left(\frac{X\beta}{\sigma}\right)}\right]\right\}. \tag{8}$$

Given the non-linearity of specifications (3), (4) and (5), the corresponding marginal effects of the explanatory variables on leverage ratios are not constant, having to be calculated for specific values of the explanatory variables. However, it is straightforward to show that, in expressions (6), (7) and (8), β_j is being multiplied by a positive term. Therefore, to examine the significance and direction of each marginal effect, it suffices to test the significance and analyze the sign of β_j. This implies that in the tobit model: (i) if an explanatory variable is relevant to explain $E\left(Y|X\right)$, it is also important to explain $\Pr\left(Y>0|X\right)$ and $E\left(Y|X,Y>0\right)$; and (ii) if an explanatory variable influences positively (negatively) $E\left(Y|X\right)$, its influence over $\Pr\left(Y>0|X\right)$ and $E\left(Y|X,Y>0\right)$ is also positive (negative).

Using the tobit model in our framework has the advantage of taking into account the existence of a mass-point at zero in the distribution of leverage ratios but still ignores their bounded nature: equation (3), despite being limited from below at zero, still has no upper bound.[4] Thus, like the linear model, the tobit model cannot represent the true data generating process of leverage ratios. However, in contrast to the linear model, the tobit model may constitute a very reasonable approximation to the true data generating process in some cases. Indeed, in practical terms, the absence of an upper bond in the tobit model may be irrelevant in many cases, in particular when the proportion of very highly leveraged firms is insignificant. A more serious problem is that the tobit model is very stringent in terms of assumptions, requiring normality and homoskedasticity of the latent dependent variable. The assumption of each covariate to influence in the same direction $\Pr\left(Y>0|X\right)$ and $E\left(Y|X,Y>0\right)$ may also be too restrictive in some cases; for an example, see the last paragraph of Section 3.1. There are some modified tobit models that could be used (e.g. the heteroskedasticity-robust tobit estimator used by Wald, 1999), but none of them would solve simultaneously all the issues associated with the use of tobit models. Anyway, if we are not interested in the latent model, instead of first specifying it in order to find a model for the actual outcomes, would not it be more

[4]Note that a two-limit version of the tobit model, with limits at 0 and 1 (e.g. Johnson 1997), which would in fact restrict the predicted values of Y to the unit interval, cannot be applied, in general, to model leverage ratios, since usually there are no observations for $Y=1$.

natural simply assuming directly a model for $E(Y|X)$, as the models discussed next do?

2.3 Fractional regression models

Recently, Cook, Kieschnick and McCullough (2008) and Ramalho and Silva (2009) have used the so-called fractional regression model (FRM) (or some extension of it) to analyze the financial leverage of firms; see Ramalho, Ramalho and Murteira (2011) for a recent survey on this model. The FRM was developed by Papke and Wooldridge (1996) for dealing specifically with dependent variables defined on the unit interval and, therefore, is based on the assumption of a functional form for $E(Y|X)$ that respects the range of values that leverage ratios may take on:

$$E(Y|X) = G(X\beta), \tag{9}$$

where $G(\cdot)$ is some nonlinear function satisfying $0 \leq G(\cdot) \leq 1$.

Papke and Wooldridge (1996) suggest as possible specifications for $G(\cdot)$ any cumulative distribution function such as those that are commonly employed with binary responses. Thus, popular choices for $G(\cdot)$ are the well-known probit and logit functional forms or the asymmetric loglog and complementary loglog models, which are given by, respectively,

$$G(X\beta) = \Phi(X\beta), \; G(X\beta) = \frac{e^{X\beta}}{1+e^{X\beta}}, \; G(X\beta) = e^{-e^{-X\beta}} \text{ and } G(X\beta) = 1 - e^{-e^{X\beta}}.$$

The partial effects implied by each one of these alternative FRMs are given by

$$\frac{\partial E(Y|X)}{\partial X_j} = \beta_j g(X\beta), \tag{10}$$

where $g(X\beta) = \partial G(X\beta)/\partial(X\beta)$. Hence, for the same reasons indicated for the tobit model, the significance and the direction of the marginal effects may be analyzed simply by examining the significance and sign of β_j.

The most relevant assumption made in the FRM is the functional form adopted for $E(Y|X)$. Thus, this model requires fewer assumptions than the tobit model and similar assumptions to the linear model. Another advantage relative to the tobit model is that the predicted values of leverage ratios are guaranteed to lie in the unit interval in all circumstances. On the other hand, the main drawback of the FRM, also relative to the tobit model, is that it treats the zero observations as any other value, i.e. implicitly it is assumed that the probability of observing a specific value in the interval [0,1] is insignificant. This implies that the FRM may not be the best option for modeling leverage ratios when a large proportion of the sampled firms do not use debt. Also, for the same reason, it is not possible to obtain sensible estimates of $\Pr(Y > 0|X)$ and $E(Y|X, Y > 0)$, which may be quantities of interest for researchers in many empirical studies, given the large number of no-debt firms that are usually present in financial leverage regression analysis.

Typically, the FRM is estimated by the quasi-maximum likelihood method, using as log-likelihood function the same Bernoulli function that is used with binary responses. See Papke and Wooldridge (1996) or Ramalho, Ramalho and Murteira (2011) for details.

2.4 Two-part fractional regression models

In contrast to the FRM, the two-part FRM (2P-FRM) proposed by Ramalho and Silva (2009) takes explicitly into account that the probability of observing a no-debt firm may be relatively large. The 2P-FRM uses separate models to explain the decisions: (i) to issue or not to issue debt; and (ii) (for those firms that do decide to use debt) on how much debt to issue (in relative terms). With this model, the factors that explain the former decision are not constrained to be the same that affect the latter decision and their effect may be different in magnitude.

The 2P-FRM may be expressed as

$$\begin{aligned} E(Y|X) &= \Pr(Y > 0|X) \cdot E(Y|X, Y > 0) \\ &= F(X\beta_{1P}) \cdot M(X\beta_{2P}), \end{aligned} \tag{11}$$

where β_{1P} and β_{2P} are vectors of variable coefficients and $F(\cdot)$ and $M(\cdot)$ are typically cumulative distribution functions, i.e. they may be specified as the $G(\cdot)$ function considered in the previous section. Thus, in the first part of the 2P-FRM model, a standard binary choice model is used for explaining the probability of a firm using debt,

$$\Pr(Y > 0|X) = \Pr(Z = 1|X) = F(X\beta_{1P}), \tag{12}$$

where $Z = 1$ for $Y > 0$ and $Z = 0$ otherwise, while in the second part, a standard FRM is used to explain the magnitude of the leverage ratios of firms that do use debt:

$$E(Y|X, Y > 0) = M(X\beta_{2P}). \tag{13}$$

For simplicity, we assume that the same regressors appear in both parts of the model, but this assumption can be relaxed and, in fact, should be if there are obvious exclusion restrictions. Note that the two components of (11) are estimated separately: while model (12) is estimated by maximum likelihood using the whole sample, model (13) is estimated by quasi-maximum likelihood using only the sub-sample of firms with nonzero leverage ratios.

The marginal effects of a covariate X_j over the probability of observing a firm using debt and the conditional mean leverage ratios of leveraged firms are given by, respectively,

$$\frac{\partial \Pr(Y > 0|X)}{\partial X_j} = \beta_{1P_j} f(X\beta_{1P}) \tag{14}$$

and

$$\frac{E(Y|X, Y > 0)}{\partial X_j} = \beta_{2P_j} m(X\beta_{2P}), \tag{15}$$

where $f(X\beta_{1P})$ and $m(X\beta_{2P})$ are the partial derivatives of $F(\cdot)$ and $M(\cdot)$ with respect to $X\beta_{1P}$ and $X\beta_{2P}$, respectively. The overall marginal effects of changes in X_j on Y can be written as

$$\frac{\partial E(Y|X)}{\partial X_j} = \beta_{1P_j} f(X\beta_{1P}) M(X\beta_{2P}) + \beta_{2P_j} F(X\beta_{1P}) m(X\beta_{2P}). \tag{16}$$

To analyze the significance and direction of the marginal effects (14) and (15), it suffices to examine the significance and sign of β_{1P_j} and β_{2P_j}, respectively. Therefore, in contrast to the tobit model, as β_{1P_j} and β_{2P_j} are not constrained to be identical, each covariate is allowed to influence in opposite ways $\Pr(Y > 0|X)$ and $E(Y|X, Y > 0)$. Regarding the overall marginal effect (16), the simple analysis of β_{1P_j} and β_{2P_j} may not lead, in general, to any conclusion. Indeed, unless both parameters are significant and have the same sign, determining the overall significance and direction of a covariate in a 2P-FRM requires the full evaluation of (16). Given that (16) depends on the values of all explanatory variables, the overall marginal effect of a particular covariate may be positive for some firms, negative for others and insignificant for the remaining.

Clearly, the 2P-FRM is much more flexible than the tobit model. In fact, an expression similar to (16) may be also written for the tobit model, see McDonald and Moffitt (1980), but with β_{1P} and β_{2P} constrained to be identical and $F(\cdot)$ and $M(\cdot)$ required to be based on normal distribution functions. Another model that is more flexible than the tobit model and closely related to the two-part model is the bivariate sample selection model (also known as type II tobit model), namely Heckman's (1979) two-step procedure. While the two-part model assumes that the level of use, if any, is conditionally independent of the decision to use (implying that each part of the model is modelled independently of the other), Heckman's (1979) approach is based on a joint model for both the censoring mechanism and outcome, where the error terms of the participation and amount debt decision equations are assumed to be related (implying that the equation estimated in the second step has an additional regressor that was estimated in the first step). To the best of our knowledge, Heckman's (1979) two-step procedure has not been adapted to the fractional response framework. See Leung and Yu (1996) for a generic comparison between two-part and sample selection models.

3 Which regression model to use?

In this section, we first discuss why some capital structure theories are best represented by specific regression models. Then, we review some econometric tests that may be used for assessing the specification of each regression model and discriminating between the competing models and, hence, theories.

3.1 Theoretical reasoning

From the analysis in Section 2, it is clear that the four regression models analyzed may be divided in two main groups. On the one hand, we have the linear, the tobit and the fractional regression models, termed from now on 'one-part models', which imply that each covariate has a unique type of effect on leverage ratios. On the other hand, we have the 2P-FRM, which allows the zero and the positive leverage ratios to be explained differently. Therefore, when choosing a suitable regression model for describing a specific capital structure theory, a first issue to consider is whether the theory provides or not a single explanation for the zero and the positive leverage ratios, i.e. for the participation and amount debt decisions. Then, when it seems preferable to use one-part models, a further issue arises, again related to the interpretation placed upon the observed zeros: is it possible to interpret them as caused by a censoring mechanism? In case of a positive answer, then the tobit model is potentially the most suitable representation of the theory. Otherwise, the FRM should be used (and the linear model could perhaps provide a reasonable approximation for the true data generating process).

Up to date, most capital structure empirical studies have used one-part models to explain leverage ratios, which follows directly from the fact that most capital structure theories provide a single explanation for all possible values of leverage ratios, including the value zero. This is the case, for example, of the two most popular explanations of capital structure decisions, the trade-off and the pecking-order theories. For details on both theories, see the recent survey by Frank and Goyal (2008).

The trade-off theory claims the existence of an optimal capital structure that firms have to reach in order to maximize their value. The focus of this theory is on the benefits and costs of debt. The former include essentially the tax deductibility of interest paid, while the latter are originated by an excessive amount of debt and the consequent potential bankruptcy costs. Thus, firms set a target level for their debt-equity ratio that balances the tax advantages of additional debt against the costs of possible financial distress and bankruptcy. This optimization problem may generate for leverage ratios any value in the unit interval, including the value zero.

The pecking-order theory, on the other hand, argues that firms do not possess an optimal capital structure, although the financing decisions of each firm are not irrelevant for its value. Indeed, due to information asymmetries between firms' managers and potential outside financiers, which limit access to outside finance, firms tend to adopt a perfect hierarchical order of financing: first, they use internal funds (retained earnings); in case external financing is needed, they issue low-risk debt; only as a last resort, when the firm exhausts its ability to issue safe debt, are new shares issued. In the absence of investment opportunities, firms retain earnings and build up financial slack to avoid having to raise external finance in the future. Hence, the firm leverage at each moment merely reflects its external financing requirements, which may be null or any positive amount, without a tendency to revert to any particular capital structure.

As stated, the trade-off and the pecking-order theories seem to imply the use of the FRM, since null leverage ratios result from an optimization problem, being, therefore, a consequence of individual choices and not of any type of censoring. However, it is straightforward to incorporate in those theories plausible justifications for the use of the tobit model. For example, we may assume that the firms with zero debt would really like to have negative debt (e.g. own short term debt securities or loans) but accounting conventions do not allow the entry of negative debt. Therefore, both FRM and tobit models may be used in empirical work based on the trade-off and the pecking-order theories.

In contrast to these classical capital structure theories, Kurshev and Strebulaev (2007), Strebulaev and Yang (2007) and Ramalho and Silva (2009) have recently argued that zero-leverage behaviour is a persistent phenomenon and that standard capital structure theories are unable to provide a reasonable explanation for it. In particular, they found that while larger firms are more likely to have some debt, conditional on having debt, larger firms are less levered, that is, firm size seems to affect in an inverse way the participation and amount debt decisions. According to Kurshev and Strebulaev (2007), these opposite effects of firm size on leverage may be explained by the presence of fixed costs of external financing, and the consequent infrequent refinancing of firms, since smaller firms are much more affected in relative terms then larger firms. Thus: (i) small firms choose higher leverage at the moment of refinancing to compensate for less frequent rebalancing, which explains why, conditional on having debt, they are more levered than large firms; (ii) as they wait longer times between refinancings, small firms, on average, have lower levels of leverage; and (iii) in each moment, there is a mass of firms opting for no leverage, since small firms may find it optimal to postpone their debt issuances until their fortunes improve substantially relative to the costs of issuance. Clearly, in this framework, the 2P-FRM is the best option for modelling leverage ratios, since the variable size and other variables are allowed to influence each decision in a different fashion.

3.2 Specification tests

From the previous discussion, it is clear that when analyzing financial leverage decisions, we cannot establish *a priori*, using only theoretical arguments, whether one- or two-part models should be used, since some of the competing theories imply the use of one-part models and others favour the use of two-part models. Moreover, although the financial theory suggests the type of regression model that should be used, it does not provide, in general, any indication about the specific model functional form that offers the best representation for the relationship between leverage ratios and explanatory variables. Therefore, in order to increase the reliability of empirical results on capital structure decisions, it is essential to apply statistical tests to discriminate between one- and two-part models and between alternative specifications for each class of models. However, despite the availability of a number of tests that can be used to that end, such tests have been rarely applied in empirical work. In this section, we discuss some of those econometric tests. In particular, given that the main practical difference between alternative one-part and two-part regression models relates to the functional form assumed for $E(Y|X)$, see (1), (3), (9) and (11), next we focus on tests for conditional mean assumptions.

One way of assessing the specification of $E(Y|X)$ is to use tests appropriate for detecting general functional form misspecifications, such as the well-known RESET test. Indeed, using standard approximation results for polynomials, it can be shown that any index model of the form $E(Y|X) = L(x\theta)$, for unknown $L(\cdot)$, can be arbitrarily well approximated by $S\left(X\theta + \sum_{j=1}^{J} \gamma_j (X\theta)^{j+1}\right)$ for J large enough. Therefore, testing the hypothesis $E(Y|X) = S(X\theta)$ is equivalent to testing $\gamma = 0$ in the augmented model $E(Y|X, W) = S(X\theta + W\gamma)$, where $W = \left[\left(X\hat{\theta}\right)^2, ..., \left(X\hat{\theta}\right)^{J+1}\right]$. The first few terms in the expansion are the most important, and, in practice, only the quadratic and cubic terms are usually considered. Note that the RESET test cannot be directly applied to assess (11), the functional form assumed for two-part models. Instead, it has to be separately applied to their two components, given by (12) and (13).

Alternatively, because all competing specifications for $E(Y|X)$ are non-nested, we may apply standard tests for non-nested hypotheses, where the alternative specifications for $E(Y|X)$ are tested against each other. An example of this type of test is the P test proposed by Davidson and MacKinnon (1981), which is probably the simplest way of comparing nonlinear regression models. To our knowledge, only Ramalho, Ramalho and Henriques (2010) have applied the P test for choosing between linear, tobit and one-part and two-part FRMs.

Suppose that $H(X\alpha)$ and $T(X\eta)$ are competing functional forms for $E(Y|X)$. As shown by Davidson and MacKinnon (1981), testing $H_0 : H(X\alpha)$ against $H_1 : T(X\eta)$ is equivalent to testing the null hypothesis $H_0 : \delta_2 = 0$ in the following auxiliary regression:

$$\left(y - \hat{H}\right) = \hat{h}X\delta_1 + \delta_2\left(\hat{T} - \hat{H}\right) + error, \tag{17}$$

where $h = \partial H(X\alpha) / \partial(X\alpha)$, δ_2 is a scalar parameter and $\hat{}$ denotes evaluation at the estimators $\hat{\alpha}$ or $\hat{\eta}$, obtained by separately estimating the models defined by $H(\cdot)$ and $T(\cdot)$, respectively. To test $H_0 : T(X\eta)$ against $H_1 : H(X\theta)$, we need to use another P statistic, which is calculated using a similar auxiliary regression to (17) but with the roles of the two models interchanged. As is standard with tests of non-nested hypotheses, three outcomes are possible: one may reject one model and accept the other, accept both models or reject both.

In contrast to the RESET test, the P test may be applied to test directly the full specification of two-part models, i.e. $H(\cdot)$ (and $T(\cdot)$) may be given by (11). Thus, the P test based on (17) may be used for choosing between: (i) alternative specifications for one-part models; (ii) alternative specifications for two-part models; and (iii) one-part and two-part models. In addition, $H(x\alpha)$ and $T(x\eta)$ may represent alternative functional forms for $\Pr(Y > 0|x)$ or $E(Y|X, Y > 0)$, in which case the P test may be used to select between competing specifications for the first or the second component of a two-part model, respectively.

As fractional data is intrinsically heteroskedastic, heteroskedasticity-robust versions of the RESET and P tests must be computed in all cases.

4 Empirical comparison of alternative regression models for leverage ratios

In order to explore some of the functional form issues that affect empirical capital structure studies, in this section we compare the results produced by several alternative regression models for leverage ratios using the data set considered previously by Ramalho and Silva (2009). These authors analyzed the financial leverage decisions of Portuguese firms using a 2P-FRM model based on a logistic specification for the two levels of the model. Here, we consider also the tobit model and loglog specifications for the one- and (both components of) two-part FRMs. We could have also considered other specifications for the FRM but, as shown by Ramalho, Ramalho and Murteira (2011), in general, the most distinct results are obtained when we contrast symmetric specifications (e.g. logit, probit) with asymmetric ones (e.g. loglog, complementary loglog). Given that the number of Portuguese firms that do not use (long-term) debt is very large (see Table 1 below), using a loglog specification is clearly the best option for an asymmetric FRM. We consider also a linear specification for the fractional component of 2P-FRMs in order to examine whether the linear model is a better approximation for the true data generating process of leverage ratios when the analysis is conditional on using debt.

Next, we first provide a brief description of the data used in the analysis. Then, we illustrate the usefulness of the specification tests discussed in section 3 for selecting appropriate regression models for leverage ratios. Finally, we compare the results of each estimated model in the following respects: (i) the significance, direction and magnitude of marginal effects; and (ii) the prediction of leverage ratios.

4.1 Data and variables

The sample used by Ramalho and Silva (2009) were drawn from the *Banco de Portugal* Central Balance Sheet Data Office, which contains some information about balance sheets, income statements and other characteristics of many Portuguese firms for the year 1999. We excluded from the analysis the following types of firms: (i) non-financial firms, since the capital structure of financial corporations is not strictly comparable with those of other firms - see Rajan and Zingales (1995); (ii) 15 firms with zero sales, in order to exclude firms which were temporarily unoperational or in the very early or very late stages of business operations; (iii) 283 firms with negative earnings before interest, taxes and depreciation (EBITDA), because our regression model uses the ratio between depreciation and EBITDA as a proxy for the explanatory variable non-debt tax shields (NDTS) - the inclusion of firms with negative earnings would create a discontinuity in the NDTS measure at zero euros of EBITDA (see e.g. Jensen, Solberg and Zorn 1992, p. 253, footnote 9); (iv) 334 firms with negative book values of equity, because such firms lack economic interpretation (e.g. Baker and Wurgler, 2002; Byoun, 2008; Lemmon, Roberts and Zender, 2008); and (v) 4 firms with huge outliers for the variable NTDS. This selection criteria produced a final sample of 4692 firms.

In accordance with the latest definitions adopted by the European Commission (recommendation 2003/361/EC), each firm was assigned to one of the following four size-based group of firms: micro firms, small firms, medium firms and large firms. Taking into account the conclusions achieved in Ramalho and Silva (2009), in this paper we perform a separate regression analysis for each one of the following size-based group of firms: (i) micro firms; (ii) small firms; and (iii) medium/large firms. In order to save space, only the results obtained for the first and third groups are reported below.[5]

[5] The results for small firms are relatively similar to those obtained for micro firms. Full results are available from the author upon request.

We consider as a measure of financial leverage the ratio of long-term debt (LTD, defined as the total company's debt due for repayment beyond one year) to long-term capital assets (defined as the sum of LTD and the book value of equity). As reported in Table 1, which contains the breakdown of our sample by group, a very high proportion of firms do not use LTD to finance their businesses: almost 90% of micro firms and about half of medium and large firms. On the other hand, very few firms display leverage ratios close to one. Clearly, in this framework, a very relevant issue is in fact how to deal with the lower bound of leverage ratios. The much larger proportion of zero leverage ratios in our sample than in those referred to in Footnote 2 may be explained as follows: (i) the papers cited in Footnote 2 define a zero-leverage firm as a firm which has no outstanding short-term and long-term debt in a given year, while we focus only on LTD; and (ii) the databases (*Compustat* or *Worldscope*) used by those authors cover essentially large (and publicly traded) firms, which are well known to use debt more frequently.

Table 1 about here

In all alternative regression models estimated next, we used the same explanatory variables as those employed by Ramalho and Silva (2009): non-debt tax shields (NDTS), measured by the ratio between depreciation and EBITDA; tangibility (TANGIB), the proportion of tangible assets and inventories in total assets; size (SIZE), the natural logarithm of sales; profitability (PROFITAB), the ratio between earnings before interest and taxes and total assets; growth (GROWTH), the yearly percentage change in total assets; age (AGE), the number of years since the foundation of the firm; liquidity (LIQUIDITY), the sum of cash and marketable securities, divided by current assets; and four industry dummies: MANUFACTURING, CONSTRUCTION, Wholesale and Retail Trade (TRADE) and Transport and Communication (COMMUNICATION). Table 2 reports some descriptive statistics of the continuous explanatory variables for the two size-based groups of firms considered in this analysis.

Table 2 about here

4.2 Model selection

We start our empirical analysis by applying to each alternative formalization the RESET and the P tests. In the latter case, we considered, one by one, all estimated models as the alternative hypothesis. Tables 3 and 4 summarize the results obtained for one- and two-part models, respectively.

Table 3 about here

Table 4 about here

The results reported in Table 3 clearly indicate that using the linear model to describe the financial leverage decisions of Portuguese firms is not appropriate at all. In fact, both for micro and medium/large firms, the specification of the linear model is rejected in all cases, irrespective of the test applied and of the alternative hypothesis used in the implementation of the P test. For micro firms, the other one-part models do not seem also to be suitable representations of capital structure decisions, since all of them are often rejected when the alternative hypothesis considered in the application of the P test is a two-part model. In contrast, for medium/large firms, the correct specification of tobit, FRM-logit and FRM-loglog is never rejected.

With regard to the specification of two-part models, see Table 4, we performed two types of tests. First, we applied separately the RESET and the P tests to each level of 2P-FRMs. Then, we used the P test to assess the full specification of 2P-FRMs against many alternative models. While the former set of tests did not provide any evidence against the correct specification of any of the 2P-FRMs estimated, the latter confirmed that these models are particularly adequate for micro firms. Indeed, for this group of firms, the correct specification of four (two) 2P-FRM is never rejected at the 5% (10%) level. In contrast, in the case of medium/large firms, all two-part models are rejected at least once at the 10% and only two are never rejected at the 5% level. Note also that using a linear model in the second component of 2P-FRMs does not seem to produce any additional problems relative to other alternatives, i.e. once no-debt firms are dropped, linear models seem to provide a much better approximation for the data generating process of leverage ratios.

Combining the results reported in Tables 3 and 4, we find that two-part models are clearly preferable for micro firms, while there is some evidence that one-part models are better for medium/large firms. These conclusions are not surprising and we conjecture that they are directly related to the proportion of zero-debt firms in each group. In fact, in the micro firm group, zero leverage ratios occur with too large a frequency than seems to be consistent with a simple, one-part model. Indeed, given their reduced size, the theory put forward by Kurshev and Strebulaev (2007), see Section 3.1, applies particularly to them.[6]

4.3 Marginal effects: statistical significance, direction and magnitude

The results obtained in the previous section show clearly that the same regression model is not suitable, in general, to explain the capital structure decisions of all size-based groups of firms. However, in most empirical studies, only one type of regression model is estimated and no specification tests are applied. In this Section, we investigate whether the conclusions, in terms of the significance, direction and magnitude, produced by alternative models, some of which are naturally misspecified, are substantially different or not.

In Tables 5 and 6, we report for both one- and two-part models, respectively, the estimation results obtained for micro and medium/large firms. For each explanatory variable, we report the values of the associated estimated coefficient and t-statistic. For each model, we report also the value of an R^2-type measure, which we call Pseudo-R^2, that was calculated as the square of the correlation between actual and predicted leverage ratios and, thus, is comparable across models.[7] For the linear model, we report also the percentage of predictions outside the unit interval. Note that, based on these tables, we can compare the regression coefficients (and, hence, the marginal effects of covariates) in terms of their significance and sign but not their magnitude, since each model implies a different functional form for $E(Y|X)$, $\Pr(Y>0|X)$ and $E(Y|X,Y>0)$.

Table 5 about here

Table 6 about here

Considering first one-part models, note that these new results provide further evidence about the low ability of the linear model to explain leverage ratios. Indeed, this model

[6] Alternatively, we could conjecture that most micro firms with zero debt would really like to have positive debt but face borrowing constraints, which may be accommodated by a two-part model but not by a one-part model: under that assumption, we may interpret the first level of a two-part model as explaining the probability of a firm overcoming possible borrowing constraints. In contrast, we expect that most of no-debt large firms do not use debt simply because it is not advantageous for them, which, as discussed in Section 3, is straightforwardly accommodated by one-part models.

[7] In a linear regression model, this Pseudo-R^2 equals the traditional R^2. See Cameron and Trivedi (2005), section 8.7.1., for a discussion of alternative goodness-of-fit measures for nonlinear models.

displays the lowest Pseudo-R^2 of all models for both groups of firms. Moreover, if we use the linear model for predicting leverage ratios for the sampled firms, in the two regressions carried out we obtain some predictions outside the unit interval (below zero). As could be expected, the higher the percentage of zero-debt firms in each group, the higher the percentage of negative predicted leverage ratios. In contrast, given that most observed leverage ratios are very far away from one, ignoring the upper bound of leverage ratios does not cause any special problem for the tobit model in these examples. Finally, note that the FRMs display the largest Pseudo-R^2's in all cases. The differences between the alternative models are more important for micro firms, in which case the Pseudo-R^2 of the linear and tobit models are, respectively, about 28% and 14% smaller than that of the FRM-logit. Interestingly, in spite of treating the zero observations as any other value, the FRMs seem to fit the data better than the tobit model.

The clear econometric inappropriateness of the linear model does not seem to jeopardize its ability to examine the significance of the regression coefficients and to calculate the type of effect (positive/negative) of the explanatory variables, particularly if we base our decisions on the 10% significance level. Indeed, in such a case, the linear model produces exactly the same conclusions as FRM-logit in all the regressions performed and differs from tobit and FRM-loglog only on the analysis of the effects of the variables SIZE (medium/large firms) and AGE (micro firms), respectively. When decisions are based on the 5% or 1% significance levels, the conclusions achieved by each model, although not so analogous, are still very similar in most cases.

The results produced by the various specifications considered for two-part models, see Table 6, are also very similar, both in terms of the significance and sign of the parameters of interest, in all alternative specifications considered for each component of the 2P-FRM. Moreover, the Pseudo-R^2 displayed by all specifications in each level of the model is almost identical in all cases. This similarity of results includes the linear specification used in the second component of some 2P-FRMs. Thus, as had already been suggested by the tests applied in the previous section, when only leveraged firms are used in the regression analysis, the differences between the various models are attenuated and, hence, linear models may provide a reasonable approximation for the true data generating process of leverage ratios. Nevertheless, note that, even in this case, the linear model yields some negative predicted leverage ratios for both groups of firms.

In contrast to the comparisons involving only one class of regression models, we find some important differences in the comparison of one- and two-part models. As discussed earlier, while the tobit model assumes that each covariate affects in the same direction $E(Y|X)$, $\Pr(Y>0|X)$ and $E(Y|X,Y>0)$, the 2P-FRM allows the covariates to affect in independent ways each one of those quantities. Analyzing Tables 5 and 6, we find that, in fact, tobit and 2P-FRMs often lead to opposite conclusions. For example, according to the tobit model, the variable SIZE has a positive effect over those three quantities, while all 2P-FRMs indicate that SIZE influences positively the probability of a firm raising debt but has no (micro firms) or a negative (medium/large firms) effect on $E(Y|X,Y>0)$.[8] Also, according to any of the estimated 2P-FRMs, most of the explanatory variables used in this study are relevant for explaining $\Pr(Y>0|X)$ but not $E(Y|X,Y>0)$. Note also that for micro firms the highest Pseudo-R^2's are displayed by 2P-FRMs (namely, those based on a logit specification for the first level of the model), while for medium/large firms one-part FRMs present slightly higher values. Both findings conforms with the conclusions achieved in the model selection stage described in the previous section.

Overall, the results obtained in this section suggest that if our main interest is simply the determination of which factors affect capital structure choices, then we should take a special care in deciding between the use of one- and two-part models. For this option, the specification tests applied in the previous section may be especially useful. Which particular specification should be used in each class of models seems to be a less relevant issue. However, in some cases, we may also be interested in the magnitude of the effects that each variable exerts over the proportion of debt used by firms. As discussed before, apart from the linear model, the marginal effects yielded by the other models are not constant, depending on the values of the explanatory variables. Thus, in order to investigate whether the magnitude of marginal effects differs substantially across models, we need to evaluate them at specific values of the covariates. This is also the only way of determining the statistical significance and direction of the overall marginal effects produced by

[8]Note the opposite effects that SIZE has over the two levels of the 2P-FRMs estimated for medium and large firms. These effects are in accordance with the two-part capital structure theory put forward by Kurshev and Strebulaev (2007), see Section 3.1, and accommodate recent findings by Cassar (2004) and Faulkender and Petersen (2007), which found that, conditional on having debt, larger firms are less prone to use debt.

2P-FRMs.[9]

In applied work, the two standard measures of marginal effects in nonlinear regression models are the average sample effect, which is the mean of the partial effects calculated independently for each firm in the sample, and the population partial effect, which is calculated for specific values of the covariates. In Table 7 we report the latter type of effect for a firm belonging to the manufacturing industry (most firms in our sample are in this industry) and evaluate each non-binary covariate at its sample mean. We report the marginal effects of non-binary covariates on $E(Y|X)$, $\Pr(Y>0|X)$ and $E(Y|X,Y>0)$.

Table 7 about here

Regarding the overall marginal effects of covariates, we found that the significance and direction of those effects in 2P-FRMs is similar to those of one-part models. Moreover, the estimates of those effects are of a comparable magnitude across models in most cases, with a few exceptions occurring mainly in the analysis of the effects of the variable PROFITABILITY. A similar conclusion can be achieved when we compare the marginal effects of covariates over the probability of a firm using debt, especially for the micro firm case. In contrast, the estimates produced by the tobit model for the effects of covariates on $E(Y|X,Y>0)$ differ substantially from those yielded by 2P-FRMs. In fact, it seems that in the tobit model the effects on $E(Y|X,Y>0)$ are confounded by the effects of covariates on the participation decision: the former effects are clearly biased in the direction of the latter, especially for micro firms. Thus, as already found above for the significance and direction of marginal effects, in terms of their magnitude it is also in the estimation of effects on $E(Y|X,Y>0)$ that tobit and 2P-FRMs produce more distinct results.

The similarities and divergences found in this section between the tobit and the 2P-FRMs may be explained as follows. In the tobit case, the parameters β that appear in $\Pr(Y>0|X)$ and $E(Y|X,Y>0)$, see expressions (4) and (5), are estimated using both the censored as well as the uncensored observations. In contrast, with 2P-FRMs the whole sample is used only in the estimation of the parameters β_{1P} in $\Pr(Y>0|X)$ of (12), while only the uncensored observations are used to identify the parameters β_{2P} in $E(Y|X,Y>0)$ of (13). Hence, while marginal effects for $\Pr(Y>0|X)$, being based on the same sample, tend to be similar across models, those for $E(Y|X,Y>0)$, especially

[9]We use the delta method to test the statistical significance of the overall effects of covariates in the 2P-FRMs.

when the percentage of censored observations is large, as is typical in our and most capital structure studies, may be very distinct. Therefore, when the mechanisms that explain the participation and amount debt decisions are different, using the tobit model to estimate effects on $E(Y|X, Y > 0)$ can produce misleading results in terms of significance, direction and magnitude.

4.4 Prediction of leverage ratios

Finally, we may also be interested in using the estimated models for predicting leverage ratios for specific firms. We have already found that, in general, the linear model gives rise to predicted outcomes outside the unity interval. In this section, we provide a comprehensive comparison of the magnitude of predicted outcomes produced by each alternative regression model.

Table 8 reports the estimated correlations between the leverage ratios predicted by each model for the sampled firms in each size-based group. All correlations are high, being above 0.8 in all cases and 0.9 if we exclude the linear model. Indeed, the lowest correlations are those that involve the linear model, which is not surprising, since this is the only model that produces predictions outside the unit interval. Also as expected, the linear model is less correlated with the other models in the case of micro firms, suggesting again that the performance of the linear model effectively deteriorates as the proportion of zero-debt firms in the sample increases. Given the results obtained in previous sections, it was also expected that the correlations between the outcomes predicted by the six variants of the 2P-FRMs were very high and, in fact, they are at least 0.987 in all cases.

Table 8 about here

A very different picture is given by Table 9, where we report the correlations between the predictions for $\Pr(Y > 0|X)$ and $E(Y|X, Y > 0)$ produced explicitly by each 2P-FRM and implicitly by the tobit model. While the correlations between alternative 2P-FRMs are again very high in all cases, those involving the tobit model are much lower when the aim is predicting $E(Y|X, Y > 0)$, which is in accordance with the findings of the previous section.

Table 9 about here

A high correlation between outcomes predicted by different models does not automatically imply that the magnitude of those outcomes is similar. Therefore, in Figures 1 and 2 we compare the magnitude of predicted leverage ratios for some specific cases. We compute both unconditional ($E(Y|X)$) and conditional on using debt ($E(Y|X, Y > 0)$) predictions. In the former case, we consider predictions from all one-part models and two of the 2P-FRMS included in this empirical study. In the latter case, we consider tobit and all 2P-FRMs based on different specifications for $E(Y|X, Y > 0)$. In Figure 1, we analyze the case of a firm belonging to the manufacturing industry, representing for each model the corresponding predicted leverage ratio as a function of SIZE. In this representation we consider for SIZE 1000 equally-spaced values between its 1% and 99% sample quantiles and evaluate the remaining non-binary explanatory variables at their median values. Figure 2 considers a similar experiment, but in this case the predicted leverage ratios are calculated as a function of PROFITABILITY. In both figures, the grey area represents a 95% confidence interval, constructed using the delta method, for the predictions yielded by the model that produces the most distinct results in each case: the linear model for $E(Y|X)$ predictions and the tobit model for $E(Y|X, Y > 0)$ predictions.

Figure 1 about here

Figure 2 about here

Figure 1 shows that all models produce very similar predictions for $E(Y|X)$ for most values of SIZE. Extreme values of SIZE may, however, produce predictions somewhat different across models and originate negative predictions in the linear model. Note also that only for extreme vales of SIZE the 95% confidence interval for the linear model does not cover the point predictions of the other models. On the other hand, when the interest lies on predicting $E(Y|X, Y > 0)$, while the 2P-FRMs yield indistinguishable predictions, the tobit model gives rise to very distinct results. This is particularly true for medium and large firms as a consequence of the opposite signs found for the SIZE coefficients in the tobit model (see Table 5) and the second-part of the 2P-FRMs (see Table 6). The analysis of Figure 2 leads to similar conclusions. Thus, overall, we may conclude that, similarly to the analysis of covariate marginal effects performed in the previous section, also when making predictions the biggest issue in the modelling of capital structure choices is deciding whether one- or two-part models should be used, in particular if those predictions are conditional on firms using debt. In contrast, choosing the right functional form for each type of model seems to be important only if we are interested in making predictions for extreme values of the covariates.[10]

5 Concluding remarks

In this paper we analyzed the main regression models that may be used to study the determinants of capital structure choices. We argued that the most commonly used functional form for modeling leverage ratios, the linear model, is not well suited to data that is bounded in the unit interval. Instead, (one- or two-part) fractional regression models seem to be the most natural way of modeling proportional response variables such as leverage ratios. The censored-at-zero tobit regression model, although do not taking into account the upper bound of leverage ratios, may be also in many cases a very reasonable approximation to the data generating process governing leverage ratios. We discussed the main econometric assumptions and features of the four classes of models analyzed, provided a theoretical foundation for all models by establishing a link between them and capital structure theories and reviewed some specification tests that may be applied to select the model (and theory) that provides the best description of financial leverage decisions of particular firms.

Using a data set previously considered in the literature, we illustrated how the proposed specification tests may be used in empirical work and investigated whether or not using different regression models may lead to conclusions substantially different. Considering the case where the only interest is how covariates affect the overall mean proportion of debt used by firms ($E(Y|X)$), we found that the significance and the direction of the marginal effects of covariates is very similar across models. This is a very reassuring result since, on the one hand, that has been the main aim of most empirical capital structure studies, and, on the other hand, most of the empirical evidence provided so far is based on (misspecified) linear models. In case researchers are also interested in the magnitude of marginal effects or in the prediction of leverage ratios, then some important differences may arise across models, although the estimates produced by the various models are of comparable magnitude in many cases.

We found also that, given the large number of firms that typically do not issue debt,

[10] Note, however, that in such a case it would probably make more sense to use the approach by Fattouh, Harris and Scaramozzino (2008), based on the use of quantile regressions.

the most relevant functional form issue in the regression analysis of leverage ratios is probably the choice between using a one- or a two-part model. In effect, this choice has two important implications. On the one hand, each one of those classes of model imply different types of capital structure theories. Therefore, rejecting the specification of one of those models imply the rejection of (at least, the standard form of) the corresponding theories. On the other hand, our empirical analysis revealed that, conditional on using debt, very distinct estimates of leverage ratios and marginal effects (in terms of significance, direction and magnitude) are produced by tobit and two-part models. The specification tests suggested in this paper, in particular the P test based on the full specification of two-part models, proved to be useful to select the best model in each application and should be routinely applied in empirical studies of capital structure.

While this paper focussed on the study of the determinants of capital structure choices, there are many other areas of the finance literature that may also benefit from the use of the fractional and two-part fractional regression models considered in this paper. Examples include studying the determinants of cash-holding decisions, corporate dividend policies, institutional equity ownership and the composition of the board of directors, where the variable of interest is typically given by, respectively, the ratio of cash and marketable securities to total assets (Opler, Pinkowitz, Stulz and Williamson, 1999), a dividend payout ratio (John, Knyazeva and Knyazeva, forthcoming), the ratio of shares held by institutional investors to total shares outstanding (Gompers and Metrick, 2001) and the fraction of independent directors on the board (Ferreira, Ferreira and Raposo, 2011). In all these cases, to the best of our knowledge, there is not a single empirical study that has taken into account the fractional nature of the dependent variable. Moreover, in some of those examples, there is often a mass-point at zero or one in the distribution of the variable of interest. For instance, in the case of corporate dividend policies, given the relatively large number of firms that often do not pay dividends (Fama and French, 2001), it would be interesting to examine whether the two-part fractional regression model is more appropriate to explain firm's payout ratios than the traditional linear and tobit models that are still predominant in this area.[11]

[11] Actually, as the dividend payout ratio sometimes exceeds the unity for some firms, in some cases it may be preferable to use some modified version of the two-part model. For example, we may use an exponential regression model (Santos Silva and Tenreyro, 2066) in the second part of the model in order to guarantee the positiveness of the dependent variable without restricting it to the unit interval. In fact,

Acknowledgements

The authors thank the referees for valuable comments that helped to improve the paper. Financial support from Fundacao para a Ciencia e a Tecnologia is gratefully acknowledged (grant PTDC/ECO/64693/2006).

References

Baker, M. and Wurgler, J. (2002), "Market timing and capital structure", *Journal of Finance*, 57(1), 1-32.

Bessler, W., Drobetz, W., Haller, R. and Meier, I. (2011), "Financial constraints and the international zero-leverage phenomenon", mimeo.

Byoun, S. (2008), "How and when do firms adjust their capital structures toward targets?", *Journal of Finance*, 63(6), 3069-3096.

Byoun, S., Moore, W.T. and Xu, Z. (2008), "Why do some firms become debt-free?", mimeo.

Cameron, A.C. and Trivedi, P.K. (2005), *Microeconometrics - Methods and Applications*, Cambridge University Press.

Cassar, G. (2004), "The financing of business start-ups", *Journal of Business Venturing*, 19, 261-283.

Cook, D.O., Kieschnick, R. and McCullough, B.D. (2008), "Regression Analysis of Proportions in Finance with Self Selection", *Journal of Empirical Finance*, 15(5), 860-867.

Dang, V.A. (2011), "An empirical analysis of zero-leverage firms: evidence from the UK", mimeo.

Davidson, R. and MacKinnon, J.G. (1981), "Several tests for model specification in the presence of alternative hypotheses", *Econometrica*, 49(3), 781-793.

the same applies to capital structure studies in cases where the firms with negative book equity are kept in the analysis and book debt/asset ratios are used as dependent variable.

Fama, E. and French, K. (2001), "Disappearing dividends: changing firm characteristics or lower propensity to pay?", *Journal of Financial Economics*, 60, 3-43.

Faulkender, M. and Petersen, M.A. (2006), "Does the source of capital affect capital structure?", *Review of Financial Studies*, 19(1), 45-79.

Fattouh, B., Harris, L. and Scaramozzino, P. (2008), "Non-linearity in the determinants of capital structure: evidence from UK firms", *Empirical Economics*, 34, 417-438.

Ferreira, D., Ferreira, M. and Raposo, F. (2011), "Board structure and price informativeness", *Journal of Financial Economics*, 99(3), 523–545.

Frank, M.Z. and Goyal, V.K. (2008), "Tradeoff and pecking order theories of debt", in B.E. Eckbo (ed.) *Handbook of Corporate Finance - Empirical Corporate Finance, Vol. 2*, Elsevier, pp. 135-202.

Gompers, P. and Metrick, A. (2001), "Institutional investors and equity prices", *Quarterly Journal of Economics*, 116(1), 229-259.

Heckman, J.J. (1979), "Sample selection bias as a specification error", *Econometric Theory*, 47(1), 153-161.

Jensen, G., Solberg, D. and Zorn, T. (1992), "Simultaneous determination of insider ownership, debt, and dividend policies", *Journal of Financial and Quantitative Analysis*, 27(2), 247–263.

Johnson, S.A. (1997), "An empirical analysis of the determinants of corporate debt ownership structure", *Journal of Financial and Quantitative Analysis*, 32(1), 47-69.

John, K., Knyazeva, A., and Knyazeva, D., "Does geography matter? Firm location and corporate payout policy", *Journal of Financial Economics*, forthcoming.

Kurshev, A. and Strebulaev, I.A. (2007), "Firm size and capital structure", mimeo.

Lemmon, M.L., Roberts, M.R. and Zender, J.F. (2008), "Back to the beginning: persistence and the cross-section of corporate capital structure", *Journal of Finance*, 63(4), 1575-1608.

Leung, S.F. and Yu, S. (1996), "On the choice between sample selection and two-part models", *Journal of Econometrics*, 72, 197-229.

McDonald, J.F. and Moffitt, R.A. (1980), "The uses of tobit analysis", *Review of Economics and Statistics*, 62(2), 318-321.

Opler, T., Pinkowitz, L., Stulz, R. and Williamson, R. (1999), "The determinants and implications of corporate cash holdings", *Journal of Financial Economics*, 52(1), 3-46.

Papke, L.E. and Wooldridge, J.M. (1996), "Econometric methods for fractional response variables with an application to 401(k) plan participation rates", *Journal of Applied Econometrics*, 11(6), 619-632.

Prasade, S., Green, C.J. and Murinde, V. (2005), "Company financial structure: a survey and implications for developing economies", in C.J. Green, C.H. Kirkpatrick and V. Murinde (eds.) *Finance and Development: Surveys of Theory, Evidence and Policy*, Cheltenham: Edward Elgar, 356-429.

Rajan, R.J. and Zingales, L. (1995), "What do we know about capital structure? Some evidence from international data", *Journal of Finance*, 50(5), 1421-1460.

Ramalho, E.A., Ramalho, J.J.S. and Henriques, P. (2010), "Fractional regression models for second stage DEA efficiency analyses", *Journal of Productivity Analysis*, 34(3), 239-255.

Ramalho, E.A., Ramalho, J.J.S. and Murteira, J. (2011), "Alternative estimating and testing empirical strategies for fractional regression models", *Journal of Economic Surveys*, 25(1), 19-68.

Ramalho, J.J.S. and Silva, J.V. (2009), "A two-part fractional regression model for the financial leverage decisions of micro, small, medium and large firms", *Quantitative Finance*, 9(5), 621-636.

Santos Silva, J.M.C. and Tenreyro, S. (2006), "The log of gravity", *Review of Economics and Statistics*, 88(4), 641-658.

Strebulaev, I.A. and Yang, B. (2007), "The mystery of zero-leverage firms", mimeo.

Wald, J.K. (1999), "How firm characteristics affect capital structure: an international comparison", *Journal of Financial Research*, 22(2), 161-187.

Wooldridge, J.M. (2002), *Econometric Analysis of Cross Section and Panel Data*, The MIT Press.

Table 1: Summary statistics for leverage ratios

	Micro	Medium and large firms
Number of firms	1446	1295
Number and percentage of firms with leverage ratios:		
$= 0$	1282 (88.7%)	634 (49.0%)
$> 90\%$	6 (0.4%)	1 (0.1%)
$> 95\%$	3 (0.2%)	1 (0.1%)
Sample statistics for leverage ratios		
Mean	0.053	0.148
Median	0.000	0.015
Maximum	0.985	0.978
Standard deviation	0.172	0.199

Table 2: Summary statistics for the explanatory variables

Variable	Mean	Median	Min	Max	St. Dev.
		Micro firms			
NDTS	0.866	0.503	0.000	102.149	4.039
Tangibility	0.355	0.322	0.000	0.998	0.263
Size	12.063	12.080	6.014	17.215	1.173
Profitability	0.075	0.047	-0.486	1.527	0.118
Growth	17.547	6.436	-81.248	681.354	50.472
Age	16.172	12.000	6.000	110.000	10.003
Liquidity	0.296	0.192	0.000	1.000	0.290
		Medium and large firms			
NDTS	0.829	0.628	0.000	26.450	1.485
Tangibility	0.461	0.472	0.015	0.979	0.203
Size	15.878	15.699	12.714	22.121	1.152
Profitability	0.054	0.040	-0.134	0.984	0.070
Growth	8.909	5.005	-61.621	188.035	18.284
Age	28.769	24.000	5.000	184.000	24.287
Liquidity	0.120	0.058	0.000	0.963	0.140

Table 3: Specification tests for one-part models (p-values)

	Micro firms				Medium and large firms			
	Linear	Tobit	Logit	Loglog	Linear	Tobit	Logit	Loglog
RESET test	0.000^{***}	0.872	0.586	0.636	0.005^{***}	0.940	0.729	0.843
P test								
H_1: Linear	—	0.024^{**}	0.452	0.317	—	0.780	0.847	0.527
H_1: Tobit	0.000^{***}	—	0.979	0.679	0.006^{***}	—	0.618	0.863
H_1: Logit	0.000^{***}	0.958	—	0.316	0.005^{***}	0.662	—	0.554
H_1: Loglog	0.000^{***}	0.871	0.865	—	0.007^{***}	0.676	0.397	—
H_1: Logit+Linear	0.000^{***}	0.080^{*}	0.042^{**}	0.019^{**}	0.009^{***}	0.843	0.207	0.449
H_1: Logit+Logit	0.000^{***}	0.085^{*}	0.045^{**}	0.028^{**}	0.014^{**}	0.596	0.495	0.946
H_1: Logit+Loglog	0.000^{***}	0.076^{*}	0.042^{**}	0.017^{**}	0.010^{***}	0.950	0.275	0.606
H_1: Loglog+Linear	0.000^{***}	0.173	0.129	0.027^{**}	0.013^{**}	0.623	0.139	0.377
H_1: Loglog+Logit	0.000^{***}	0.185	0.135	0.034^{**}	0.020^{**}	0.966	0.289	0.701
H_1: Loglog+Loglog	0.000^{***}	0.170	0.127	0.026^{**}	0.014^{**}	0.748	0.171	0.464

Notes: ***, ** and * denote test statistics which are significant at 1%, 5% or 10%, respectively; heteroskedasticity-robust versions of all test statistics were computed.

Table 4: Specification tests for two-part models (p-values)

	Micro firms						Medium and large firms					
	Separate assessment of each model level											
	1st part			2nd part			1st part			2nd part		
	Logit	Loglog		Linear	Logit	Loglog	Logit	Loglog		Linear	Logit	Loglog
RESET test	0.505	0.426		0.741	0.455	0.447	0.837	0.638		0.928	0.774	0.721
P test												
H_1: Linear	—	—		—	0.248	0.776	—	—		—	0.557	0.656
H_1: Logit	—	1.000		0.391	—	0.450	—	0.983		0.924	—	0.860
H_1: Loglog	0.305	—		0.554	0.191	—	0.484	—		0.844	0.455	—
	Assessment of the model's full specification											
	Logit+ Linear	Logit+ Logit	Logit+ Loglog	Loglog+ Linear	Loglog+ Logit	Loglog+ Loglog	Logit+ Linear	Logit+ Logit	Logit+ Loglog	Loglog+ Linear	Loglog+ Logit	Loglog+ Loglog
H_1: Linear	0.330	0.440	0.239	0.652	0.223	0.734	0.041**	0.678	0.036*	0.071**	0.823	0.438
H_1: Tobit	0.945	0.538	0.300	0.060*	0.602	0.838	0.366	0.988	0.861	0.089*	0.604	0.461
H_1: Logit	0.726	0.925	0.640	0.177	0.023**	0.474	0.538	0.274	0.668	0.135	0.963	0.633
H_1: Loglog	0.788	0.788	0.559	0.242	0.837	0.319	0.878	0.268	0.018**	0.463	0.181	0.747
H_1: Logit+Linear	—	0.113	0.070*	0.433	0.066	0.216	—	0.020**	0.475	0.719	0.002**	0.099*
H_1: Logit+Logit	0.080*	—	0.299	0.866	0.042**	0.275	0.602	—	0.514	0.894	0.519	0.408
H_1: Logit+Loglog	0.036**	0.543	—	0.742	0.053*	0.236	0.838	0.010***	—	0.750	0.056*	0.964
H_1: Loglog+Linear	0.437	0.416	0.271	—	0.928	0.349	0.056*	0.014**	0.259	—	0.006***	0.157
H_1: Loglog+Logit	0.237	0.224	0.178	0.121	—	0.720	0.335	0.028**	0.896	0.786	—	0.133
H_1: Loglog+Loglog	0.146	0.319	0.181	0.418	0.085*	—	0.097*	0.009***	0.259	0.684	0.003***	—

Notes: ***, ** and * denote coefficients or test statistics which are significant at 1%, 5% or 10%, respectively; heteroskedasticity-robust versions of all test statistics were computed.

Table 5: Regression results for one-part models

	Micro firms				Medium and large firms			
	Linear	Tobit	Logit	Loglog	Linear	Tobit	Logit	Loglog
NDTS	0.000	-0.073	-0.173	-0.067	-0.009***	-0.024***	-0.134**	-0.054***
	(-0.07)	(-1.45)	(-1.12)	(-1.62)	(-2.86)	(-2.62)	(-2.24)	(-2.68)
TANGIBILITY	-0.017	0.066	0.037	0.047	0.111***	0.257***	0.925***	0.406***
	(-1.05)	(0.43)	(0.11)	(0.38)	(3.12)	(4.40)	(3.20)	(3.22)
SIZE	0.028***	0.280***	0.647***	0.219***	0.007	0.028***	0.059	0.026
	(6.67)	(7.27)	(6.67)	(6.54)	(1.47)	(3.42)	(1.56)	(1.50)
PROFITABILITY	-0.069***	-1.531***	-4.942***	-1.612***	-0.455***	-1.193***	-5.797***	-2.421***
	(-2.71)	(-2.76)	(-3.23)	(-3.61)	(-5.84)	(-6.37)	(-6.59)	(-6.57)
GROWTH	0.000	0.000	-0.001	0.000	0.001***	0.002***	0.008***	0.004***
	(-0.38)	(-0.35)	(-0.43)	(-0.19)	(3.24)	(4.49)	(3.99)	(3.78)
AGE	0.001*	0.007**	0.014*	0.005	0.000	0.000	0.000	0.000
	(1.78)	(2.07)	(1.65)	(1.56)	(0.05)	(0.18)	(0.23)	(-0.07)
LIQUIDITY	-0.039***	-0.415**	-1.011**	-0.324***	-0.168***	-0.405***	-1.791***	-0.700***
	(-3.19)	(-2.55)	(-2.34)	(-2.64)	(-4.92)	(-5.58)	(-4.26)	(-4.31)
MANUFACTURING	-0.031**	-0.265**	-0.485*	-0.180*	-0.062**	-0.118**	-0.500***	-0.226**
	(-1.96)	(-2.23)	(-1.67)	(-1.90)	(-2.28)	(-2.56)	(-2.77)	(-2.49)
CONSTRUCTION	0.006	-0.184	0.029	-0.011	-0.032	-0.086	-0.244	-0.124
	(0.35)	(-1.37)	(0.10)	(-0.10)	(-0.91)	(-1.51)	(-0.99)	(-1.06)
TRADE	-0.083***	-0.828***	-1.822**	-0.465*	-0.096**	-0.283***	-1.028*	-0.460**
	(-3.00)	(-3.43)	(-2.43)	(-1.75)	(-2.32)	(-2.71)	(-1.95)	(-2.36)
COMMUNICATION	-0.044***	-0.526***	-1.407***	-0.473***	-0.035	-0.103	-0.310	-0.143
	(-2.78)	(-3.35)	(-3.59)	(-3.92)	(-0.91)	(-1.62)	(-1.20)	(-1.12)
CONSTANT	-0.258***	-4.010***	-10.151***	-3.511***	0.085	-0.334**	-2.224***	-0.853***
	(-5.10)	(-7.47)	(-7.83)	(-8.26)	(1.06)	(-2.33)	(-3.40)	(-2.83)
Number of observations	1446	1446	1446	1446	1295	1295	1295	1295
Pseudo-R^2	0.073	0.087	0.101	0.099	0.079	0.082	0.086	0.086
% of predictions outside the unit interval	11.1	—	—	—	1.7	—	—	—

Notes: below the coefficients we report t-statistics in parentheses; for the RESET test we report p-values; ***, ** and * denote coefficients or test statistics which are significant at 1%, 5% or 10%, respectively; heteroskedasticity-robust versions of all test statistics were computed.

Table 6: Regression results for two-part models

	Micro firms					Medium and large firms				
	1st part		2nd part			1st part		2nd part		
	Logit	Loglog	Linear	Logit	Loglog	Logit	Loglog	Linear	Logit	Loglog
NDTS	-0.181	-0.080	0.027	0.110	0.069	-0.098*	-0.062*	-0.014*	-0.081*	0.046**
	(-1.39)	(-1.49)	(0.69)	(0.68)	(0.59)	(-1.87)	(-1.94)	(-1.91)	(1.69)	(2.06)
TANGIBILITY	0.266	0.099	-0.054	-0.238	-0.177	1.720***	1.189***	-0.003	-0.020	0.000
	(0.49)	(0.61)	(-0.66)	(-0.70)	(-0.75)	(4.86)	(4.84)	(-0.06)	(-0.08)	(0.00)
SIZE	0.712***	0.296***	0.013	0.057	0.037	0.275***	0.203***	-0.022***	-0.111***	-0.061***
	(7.85)	(7.78)	(0.71)	(0.72)	(0.71)	(5.40)	(5.44)	(-2.97)	(-2.91)	(-3.02)
PROFITABILITY	-3.320**	-1.445***	-0.583**	-2.666**	-1.959**	-5.684***	-3.629***	-0.588***	-3.150***	-1.812***
	(-2.35)	(-2.61)	(-2.15)	(-2.12)	(-2.55)	(-5.27)	(-5.42)	(-4.21)	(-4.01)	(-4.41)
GROWTH	-0.001	0.000	0.001	0.003	0.003	0.010***	0.008***	0.001*	0.005**	0.003**
	(-0.49)	(-0.38)	(1.36)	(1.38)	(1.61)	(3.48)	(3.50)	(1.88)	(1.97)	(1.93)
AGE	0.020**	0.009**	-0.002	-0.009	-0.006	0.001	0.000	0.000	-0.001	-0.001
	(2.47)	(2.15)	(-1.12)	(-1.12)	(-1.29)	(0.42)	(0.19)	(-0.49)	(-0.48)	(-0.58)
LIQUIDITY	-1.141***	-0.395**	-0.091	-0.414	-0.242	-2.228***	-1.349***	-0.071	-0.375	-0.181
	(-2.66)	(-2.46)	(-1.05)	(-1.10)	(-1.00)	(-5.21)	(-5.27)	(-1.00)	(-1.03)	(-0.90)
MANUFACTURING	-0.703**	-0.329**	0.036	0.156	0.081	-0.661**	-0.441**	-0.036	-0.177	-0.102
	(-2.44)	(-2.49)	(0.71)	(0.74)	(0.60)	(-2.27)	(-2.00)	(-1.01)	(-1.07)	(-1.06)
CONSTRUCTION	-0.656**	-0.319**	0.254***	1.055***	0.751***	-0.724**	-0.501*	0.039	0.175	0.104
	(-2.01)	(-2.12)	(3.66)	(3.64)	(3.66)	(-2.06)	(-1.94)	(0.84)	(0.80)	(0.82)
TRADE	-2.447***	-0.830***	0.181	0.768	0.535	-1.753***	-1.174***	-0.019	-0.101	-0.061
	(-3.64)	(-3.64)	(1.26)	(1.32)	(1.28)	(-2.87)	(-3.17)	(-0.21)	(-0.23)	(-0.25)
COMMUNICATION	-1.230***	-0.562***	-0.109	-0.491	-0.307*	-0.859**	-0.588**	0.034	0.151	0.100
	(-3.19)	(-3.41)	(-1.62)	(-1.64)	(-1.68)	(-2.17)	(-2.04)	(0.68)	(0.66)	(0.73)
CONSTANT	-9.965***	-4.007	0.284	-0.891	-0.181	-3.937***	-2.577***	0.703***	1.202*	0.949**
	(-8.06)	(-7.91)	(1.14)	(-0.84)	(-0.26)	(-4.51)	(-4.05)	(5.49)	(1.84)	(2.69)
Number of observations	1446	1446	164	164	164	1295	1295	661	661	661
Pseudo-R^2										
- each model level	0.098	0.096	0.293	0.291	0.293	0.109	0.109	0.080	0.080	0.080
- full model (Logit + ...)			0.109	0.108	0.108			0.085	0.084	0.084
- full model (Loglog + ...)			0.104	0.104	0.104			0.083	0.083	0.083
% of predictions outside the unit interval	—	—	13.4	—	—	—	—	0.6	—	—

Notes: below the coefficients we report t-statistics in parentheses; for the RESET test we report p-values; ***, ** and * denote coefficients or test statistics which are significant at 1%, 5% or 10%, respectively; heteroskedasticity-robust versions of all test statistics were computed.

Table 7: Marginal effects

	Micro firms						Medium and large firms					
	Linear	Tobit	Logit	Loglog	Logit + Logit	Loglog + Loglog	Linear	Tobit	Logit	Loglog	Logit + Logit	Loglog + Loglog
Marginal effects based on E(Y|X)												
NDTS	0.000	-0.006	-0.005	-0.008	-0.003	-0.004	-0.009***	-0.012***	-0.015**	-0.015***	-0.014***	-0.014*
TANGIBILITY	-0.017	0.006	0.001	0.005	0.003	0.003	0.111***	0.131***	0.104***	0.109***	0.113***	0.109***
SIZE	0.028***	0.024***	0.020***	0.025***	0.022***	0.026***	0.007	0.014***	0.007	0.007	0.007*	0.007
PROFITABILITY	-0.069***	-0.131***	-0.151***	-0.181***	-0.148***	-0.183**	-0.455***	-0.607***	-0.653***	-0.652***	-0.694***	-0.665***
GROWTH	0.000	0.000	0.000	0.000	0.000	0.000	0.001***	0.001***	0.001***	0.001***	0.001***	0.001***
AGE	0.001*	0.001**	0.000*	0.001	0.001**	0.001	0.000	0.000	0.000	0.000	0.000	0.000
LIQUIDITY	-0.039***	-0.035**	-0.031**	-0.036***	-0.041***	-0.041**	-0.168***	-0.206***	-0.202***	-0.188***	-0.187***	-0.157**
Marginal effects based on Pr(Y>0|X)												
NDTS	—	-0.015	—	—	-0.013	-0.017	—	-0.030***	—	—	-0.024*	-0.021*
TANGIBILITY	—	0.013	—	—	0.020	0.021	—	0.321***	—	—	0.430***	0.404***
SIZE	—	0.056***	—	—	0.053***	0.063***	—	0.036***	—	—	0.069***	0.069***
PROFITABILITY	—	-0.308***	—	—	-0.247**	-0.309***	—	-1.492***	—	—	-1.421***	-1.233***
GROWTH	—	0.000	—	—	0.000	0.000	—	0.003***	—	—	0.003***	0.003***
AGE	—	0.001**	—	—	0.001**	0.002**	—	0.000	—	—	0.000	0.000
LIQUIDITY	—	-0.083**	—	—	-0.085***	-0.084***	—	-0.507***	—	—	-0.557***	-0.458***
Marginal effects based on E(Y|X,Y>0)												
NDTS	—	-0.012	—	—	0.026	0.025	—	-0.009***	—	—	-0.016*	-0.016**
TANGIBILITY	—	0.011	—	—	-0.056	-0.065	—	0.095***	—	—	-0.004	0.000
SIZE	—	0.045***	—	—	0.013	0.013	—	0.010***	—	—	-0.022***	-0.022***
PROFITABILITY	—	-0.246***	—	—	-0.634**	-0.720**	—	-0.439***	—	—	-0.618***	-0.640***
GROWTH	—	0.000	—	—	0.001	0.001	—	0.001***	—	—	0.001**	0.001**
AGE	—	0.001**	—	—	-0.002	-0.002	—	0.000	—	—	0.000	0.000
LIQUIDITY	—	-0.067**	—	—	-0.099	-0.089	—	-0.149***	—	—	-0.074	-0.064

Note: ***, ** and * denote marginal effects which are significant at 1%, 5% or 10%, respectively.

Table 8: Correlation between predicted leverage ratios - all firms

	Linear	Tobit	logit	loglog	logit+linear	logit+logit	logit+loglog	loglog+linear	loglog+logit
					Micro firms				
Tobit	0.809	—							
Logit	0.830	0.965	—						
Loglog	0.875	0.957	0.981	—					
Logit+Linear	0.846	0.943	0.986	0.984	—				
Logit+Logit	0.844	0.944	0.987	0.984	1.000	—			
Logit+Loglog	0.844	0.943	0.985	0.983	1.000	1.000	—		
Loglog+Linear	0.868	0.919	0.964	0.987	0.988	0.988	0.988	—	
Loglog+Logit	0.867	0.920	0.965	0.988	0.988	0.988	0.988	1.000	—
Loglog+Loglog	0.866	0.918	0.963	0.986	0.987	0.987	0.988	1.000	1.000
					Medium and large firms				
Tobit	0.938	—							
Logit	0.943	0.977	—						
Loglog	0.960	0.978	0.995	—					
Logit+Linear	0.951	0.959	0.985	0.992	—				
Lgit+Logit	0.949	0.959	0.987	0.993	0.999	—			
Logit+Loglog	0.949	0.960	0.986	0.993	1.000	0.999	—		
Loglog+Linear	0.952	0.949	0.973	0.986	0.997	0.995	0.996	—	
Loglog+Logit	0.950	0.950	0.975	0.987	0.997	0.997	0.997	0.999	—
Loglog+Loglog	0.950	0.950	0.974	0.987	0.997	0.996	0.997	1.000	0.999

Table 9: Correlation between predictions in the two components of two-part models

	1st part		2nd part		
	Logit	Loglog	Linear	Logit	Loglog
			Micro firms		
Logit	—		1.000	—	
Loglog	0.986	—	0.998	0.998	—
Tobit	0.989	0.988	0.352	0.351	0.338
			Medium and large firms		
Logit	—		0.996	—	
Loglog	0.995	—	0.998	0.998	—
Tobit	0.963	0.955	0.441	0.444	0.448

Figure 1: Predicted leverage ratios as a function of the SIZE variable

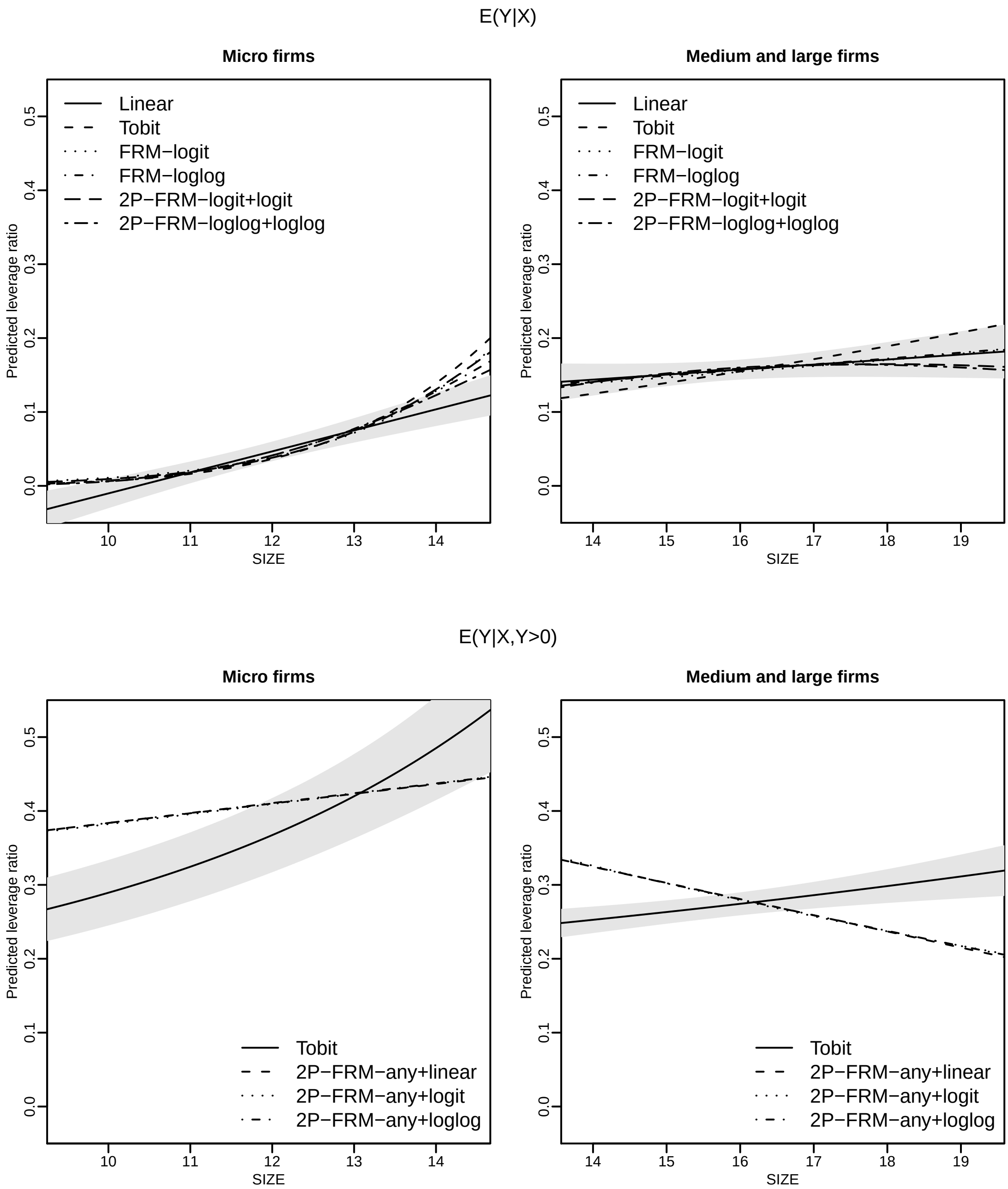

Figure 2: Predicted leverage ratios as a function of the PROFITABILITY variable