Exponential regression of panel data fractional response models with an application to firm capital structure*

Esmeralda A. Ramalho and Joaquim J.S. Ramalho

Department of Economics and CEFAGE-UE, Universidade de Evora

Luís M.S. Coelho

Faculty of Economics, Universidade do Algarve and CEFAGE-UE

September 2015

Abstract

New panel data estimators are proposed for logit and complementary loglog fractional regression models. The standard specifications of these models are transformed into a form of exponential regression with multiplicative individual effects and time-variant heterogeneity, from which six alternative estimators that do not require assumptions on the distribution of the unobservables are developed. Most of the proposed estimators are robust to both time-variant and time-invariant heterogeneity, can be applied to dynamic panel data models and accommodate endogenous explanatory variables without requiring the specification of a reduced form model. A Monte Carlo study and an application to firm capital structure choices illustrates the usefulness of the suggested estimators.

Keywords: fractional responses, exponential regression, panel data, fixed effects, heterogeneity, endogeneity, dynamic models.

JEL Classification: C25, C23.

*Financial support from Fundacao para a Ciencia e a Tecnologia (grant PTDC/EGE-ECO/119148/2010) is gratefully acknowledged. Aspects of this research were presented at the 2nd Annual Conference of the International Association for Applied Econometrics (IAAE) and the 11th Econometric Society World Congress. Address for correspondence: Joaquim J.S. Ramalho, Department of Economics, Universidade de Évora, Largo dos Colegiais, 7000-803 ÉVORA, Portugal (e-mail: jsr@uevora.pt).

1 Introduction

Formal models for response variables defined on the unit interval were first suggested by Papke and Wooldridge (1996); see Ramalho, Ramalho and Murteira (2011) for a recent survey. Quasi-maximum likelihood (QML) estimators of the most well known models, the fractional logit and probit, which require only the specification of the conditional mean of the response variable, have since been applied in numerous empirical cross-sectional studies. Examples include the proportion of debt in the financing mix of firms (Ramalho and Silva, 2009), the proportion of deaths caused by traffic accidents across districts (Ospina and Ferrari, 2012), employer 401(k) contribution match rates (Papke and Wooldridge, 1996) and data envelopment analysis efficiency scores (Ramalho, Ramalho and Henriques, 2010).

The extension of cross-sectional fractional response models to a panel setting is not trivial. Papke and Wooldridge (2008), for balanced panels, and Wooldridge (2010), for unbalanced panels, developed correlated random effects probit estimators that allow for consistent estimation of partial effects conditional on observables. However, their approaches do not allow for consistent estimation of structural parameters and require normality of individual effects with specific functional forms for their mean and variance. Loudermilk (2007), for balanced panels, and Elsas and Florysiak (2015), for unbalanced panels, suggested the use of doubly-censored tobit models, which also require similar distributional assumptions for the individual effects. In addition, their estimators require the assumption of a normal, homoscedastic distribution for the error term of the latent specification underlying the tobit model and can only be applied when the sample comprises observations at both the boundary values of zero and one. The latter feature reduces greatly the usefulness of these estimators, since, while it is common to observe a substantial proportion of boundary values in samples of fractional data, most samples cluster only at zero or one.[1] Finally, some of these estimators allow the inclusion of a lagged dependent variable (Loudermilk, 2007; Elsas and Florysiak, 2015) or an endogenous covariate (Papke and Wooldridge, 2008) in the model, but require strict exogeneity for all the other explanatory variables and, in the latter case, the assumption of a reduced form

[1] For instance, in the first two examples given at the end of the previous paragraph the samples cluster at zero but there are no observations at one, and in the last two examples it occurs the opposite.

model for the endogenous covariate.

Instead of specifying the distribution of individual effects, in linear settings it is common to deal with them by estimating individual-specific intercepts or by applying some transformation to eliminate them. Both approaches have also been used in the fractional framework. The first approach was considered by Hausman and Leonard (1997), who proposed a fixed effects logit QML estimator. However, because using logit or other fractional models in a panel data setting produces an incidental parameters problem (the QML estimates of the explanatory variables' parameters depend on the estimates of the individual-specific intercepts), this estimator is only appropriate for long panels. The second approach is based on the linearization of the fractional logit model using the log-odds transformation. However, this linearized model is not appropriate for models that have a fractional dependent variable with observations at zero and/or one, as in the examples provided above.

This paper develops new panel data estimators for fractional responses, which are based on a transformation of logit and complementary loglog (cloglog) fractional regression models into a form of exponential regression with multiplicative fixed effects. The proposed exponential-fractional regression model (EFRM) generates two sets of estimators. The first set comprises three fixed effects estimators that eliminate the time-invariant individual effects using alternative quasi difference and mean differenced transformations. The second set also includes three estimators and is based on direct estimation of the EFRM under a variety of assumptions (fixed effects, random effects and correlated random effects). All the suggested estimators: (i) do not require distributional assumptions on unobservables; (ii) under the assumptions stated, allow consistent estimation of structural parameters; (iii) accommodate the zero values of the response variable; (iv) do not require a long temporal dimension for the panel; and (v) are robust to time-variant heterogeneity. In addition, all the proposed estimators have at least two of the following properties: (vi) are robust to time-invariant heterogeneity correlated with the covariates; (vii) are applied in exactly the same way with balanced and unbalanced panels; (viii) may be applied to dynamic panel data models; (ix) require only contemporaneous or weak exogeneity of all regressors; and (x) accommodate cases of endogenous explanatory variables without requiring the specification of their reduced form. In particular, one estimator exhibits simultaneously these ten properties.

This paper is organized as follows. Section 2 describes the proposed EFRM. Section 3 addresses a small Monte Carlo simulation study. Section 4 presents an application of the EFRM to firm capital structure decisions. Section 5 concludes.

2 The exponential-fractional regression model

2.1 Alternative data generating processes for panel fractional responses

Let y_{it} denote the fractional response variable, defined on the interval $[0, 1]$, to be explained for individual i, $i = 1, ...N$, at time t, $t = 1, ...T$, and let x_{it} denote a k-vector of explanatory variables. The standard fractional regression model used in the cross-sectional context (*e.g.*, Papke and Wooldridge, 1996) is defined by the following conditional expectation:

$$E\left(y_i|x_i\right) = G\left(x_i\theta\right), \tag{1}$$

where θ is the vector of parameters of interest and $G(\cdot)$ is a (nonlinear) function bounded on the unit interval. Typically, a logit or probit functional form has been adopted for $G(\cdot)$. Alternative specifications such as the the loglog, cloglog and cauchit models were discussed by Ramalho, Ramalho and Murteira (2011).

A direct extension of model (1) to a fixed effects panel data setting yields

$$E\left(y_{it}|\alpha_i, x_{it}\right) = G\left(x_{it}\theta + \alpha_i\right), \tag{2}$$

where α_i denotes time-invariant unobserved heterogeneity. While under correct specification of $G(\cdot)$ it is straightforward to obtain consistent estimators for θ in the cross-sectional model (1), see Papke and Wooldridge (1996), the same does not happen with the panel data model (2) due to the presence of unobservables in the argument of the $G(\cdot)$ function. All the alternatives suggested, namely estimating explicitly the parameters α_i (Hausman and Leonard, 1997), assuming a distribution for α_i (Loudermilk, 2007; Papke and Wooldridge, 2008; Wooldridge, 2010; Elsas and Florysiak, 2015) or using a linearized model based on log-odds transformations, have several drawbacks, as discussed in the Introduction.

The methods proposed by the cited authors have the merit of allowing the estimation of partial effects conditional either on both observables and unobservables (Hausman and Leonard, 1997) or only on observables (the other authors). However, often applied researchers are not so much interested in the magnitude of partial effects as they are in their sign and statistical significance. Moreover, in analyses conditional on both observables and unobservables, which are the only ones that respect the notion of a *ceteris paribus* change, the structural parameters θ are of interest in its own right: on the one hand, they give the direction of the partial effects; on the other hand, testing their significance is equivalent to test the significance of partial effects.

The focus of this paper is on the estimation of structural parameters. Given the difficulty of obtaining consistent estimators for θ under (2), this paper proposes new estimators for cases where the data generating process (DGP) of y_{it} may be described in a slightly different way. In particular, the regression model proposed in the next section requires the DGP to be given by

$$y_{it} = G\left(x_{it}\theta + \alpha_i + v_{it}\right), \tag{3}$$

where v_{it} denotes time-varying unobserved heterogeneity and $G(\cdot)$ is assumed to have a logit, $G(\cdot) = \frac{\exp(\cdot)}{1+\exp(\cdot)}$, or cloglog, $G(\cdot) = 1 - \exp\left[-\exp(\cdot)\right]$, specification. This DGP is somewhat peculiar in the sense that all heterogeneity appears in the argument of the $G(\cdot)$ function and no other random component is allowed. However, it mirrors the typical economic theory formulation, treating observables and unobservables in a symmetrical way. In particular, (3) may be interpreted as the DGP of a 'well-posed economic model', that is a 'model that satisfies all of the input processes, observed and unobserved by the analyst, and their relationship to outputs' (Heckman 2000, p. 47). See Heckman (2000, 2001) for a discussion on these so-called Marshalian causal functions and Ramalho and Ramalho (2015) for a related DGP used in the cross-sectional context.

2.2 The model

The DGP (3) may be transformed into a form of exponential regression with multiplicative fixed effects when $G(\cdot)$ has a logit or cloglog specification. Indeed, a common feature of these models is that the observed and unobserved characteristics are included in an exponential function in such a way that (3) can be written as

$$y_{it} = G_1 \left[\exp\left(x_{it}\theta + \alpha_i + v_{it}\right)\right], \tag{4}$$

where $G_1(c) = \frac{c}{1+c}$ or $G_1(c) = 1 - \exp(-c)$ for, respectively, logit and cloglog models. Let $H_1(\cdot) = G_1(\cdot)^{-1}$. Then, (4) can be expressed as an exponential model with a transformed dependent variable:

$$H_1(y_{it}) = \exp\left(x_{it}\theta + \alpha_i + v_{it}\right), \tag{5}$$

where $H_1(y_{it}) = \frac{y_{it}}{1-y_{it}}$ (logit model) or $-\ln(1 - y_{it})$ (cloglog model) and y_{it} is restricted to the interval $[0, 1)$.

An important feature of the EFRM (5) is its ability to handle zero observations for the response variable, unlike the relatively popular linear-fractional model $H(y_{it}) = x_{it}\theta + \alpha_i + v_{it}$, where $H(y_{it}) = \ln[H_1(y_{it})]$ is the so-called link function widely used in the generalized linear models literature.[2] This is a major advantage, since in many cases samples of fractional data cluster only at zero or one, see the examples provided in the Introduction. Note also that cases of fractional responses defined on the interval $(0, 1]$ may be dealt with in this framework by modeling their complementary.

Another advantage of the EFRM is that dealing with fixed effects is easier than in any other nonlinear framework. On the one hand, the individual effects α_i may be eliminated from equation (5) using techniques similar to those available to deal with fixed effects in exponential models for non-negative data. For example, the difference and mean differenced transformations reviewed by Windmeijer (2008) for standard fixed ef-

[2] Typically, this limitation of the linear-fractional model has been circumvented by adding an arbitrary constant to all observations of y_{it} or dropping observations with $y_{it} = 0$. However, as shown by Ramalho and Ramalho (2015) for a cross-sectional setting, both approaches may yield large biases in the estimation of the parameters of interest (similar conclusions were achieved by Santos Silva and Tenreyo, 2006, for the case of a log-linear model for non-negative data). Furthermore, with panel data, the consequences of dropping observations with $y_{it} = 0$ are likely to be even more serious, since such a procedure gives rise to an incomplete panel governed by complicated attrition mechanisms.

fects exponential regression models may be adapted for the EFRM in a straightforward manner. On the other hand, both pooled fixed effects Poisson- and Exponential-based QML of exponential regression models do not suffer from the incidental parameters problem (see Dhaene and Jochmans, 2015), allowing consistent estimation of θ in (5) without requiring a long panel.

A final advantage of the EFRM is that it can be written in such a way that (functions of) the explanatory variables and the unobservables become additively separable, simplifying the process of dealing with endogenous covariates in this framework. In particular, dividing both sides of (5) by $\exp(x_{it}\theta + \alpha_i)$ or $\exp(x_{it}\theta)$, the EFRM may be equivalently expressed as, respectively:

$$\frac{H_1(y_{it})}{\exp(x_{it}\theta + \alpha_i)} = \exp(v_{it}) \tag{6}$$

or

$$\frac{H_1(y_{it})}{\exp(x_{it}\theta)} = \exp(\alpha_i + v_{it}). \tag{7}$$

These alternative representations of the EFRM are extensions of the model used by Mullahy (1997) for cross-sectional count data to deal with endogenous covariates and of the model considered by Ramalho and Ramalho (2015) for cross-sectional fractional data.

2.3 Alternative estimators

The structural parameters of the EFRM may be estimated in a number of ways. In this section, six alternative estimators are proposed, all of them being defined by a set of orthogonality conditions between a function of the time-varying unobservables, say $v^* \equiv v^*(y_{it}, x_{it}, \alpha_i; \theta)$, and a set of m instrumental variables, denoted by z:

$$E(z'v^*) = 0. \tag{8}$$

Depending on the assumptions made, the vector z may include the explanatory variables x_{it}, their lags and leads and/or external instruments.

To estimate the structural parameters of models defined by moment conditions, as in (8), the generalized method of moments (GMM) may be employed. The (two-step) GMM estimator $\hat{\theta}$ is obtained by minimizing

$$\left[\frac{1}{N}\sum_{i=1}^{N} Z_i' v_i^*(\theta)\right]' \hat{\Omega}\left(\tilde{\theta}\right)^{-1}\left[\frac{1}{N}\sum_{i=1}^{N} Z_i' v_i^*(\theta)\right], \tag{9}$$

where Z_i is a block diagonal matrix of instruments, $\hat{\Omega}$ is a consistent estimator of the $(m \times m)$ symmetric, positive definite matrix $\Omega \equiv E\left[Z_i' v_i^*(\theta) v_i^*(\theta)' Z_i\right]$ and $\tilde{\theta}$ is some preliminary estimator defined by an equation similar to (9), but with $\hat{\Omega}\left(\tilde{\theta}\right)$ replaced by the identity matrix. See Hansen (1982) for more details on GMM estimation of moment condition models.

To simplify the exposition, the proposed estimators are grouped into two categories, each one comprising three estimators. In one group, the EFRM (5) is first transformed in order to eliminate the individual effects. In the other group, the EFRM defined by (5), or by its alternative representations (6) or (7), is directly estimated. In each group, the estimators differ on the assumptions about the unobservables that are required for consistent estimation of the structural parameters. In all cases, α_i and v_{it} are assumed to be not correlated and the latter is assumed to be not serially correlated. For most estimators, but not all, α_i is allowed to be correlated with x_{it}. Regarding the correlation between the explanatory variables and the time-variant heterogeneity, x_{it} may be:[3]

- strictly exogenous, i.e. there is no correlation between any of the $\exp(v_{is})$ terms, $s = 1, ..., T$, and any of the x_{it}, $t = 1, ..., T$:

$$E\left[\exp(v_{it})\middle|\, \alpha_i, x_{i1}, ..., x_{iT}\right] = 1; \tag{10}$$

- weakly exogenous, i.e. x_{it} is not correlated with current and future shocks, but may be correlated with past shocks:

$$E\left[\exp(v_{it})\middle|\, \alpha_i, x_{i1}, ..., x_{it}\right] = 1; \tag{11}$$

- contemporaneously exogenous, i.e. x_{it} is not correlated with current shocks, but

[3] In all cases, because time dummies can be included in x_{it} and z_{it}, the assumption that $E\left[\exp(v_{it})\right] = 1$ is without loss of generality.

may be correlated with past and future shocks:

$$E\left[\exp\left(v_{it}\right)\mid\alpha_i, x_{it}\right] = 1; \tag{12}$$

- endogenously determined, i.e. x_{it} is correlated with current shocks and may be also correlated with past shocks,

$$E\left[\exp\left(v_{it}\right)\mid\alpha_i, x_{i1}, ..., x_{i,t-1}\right] = 1, \tag{13}$$

  or not:

$$E\left[\exp\left(v_{it}\right)\mid\alpha_i, x_{i1}, ..., x_{i,t-1}, x_{i,t+1}, ..., x_{iT}\right] = 1. \tag{14}$$

  In the former case, lagged regressors may be used as instruments for x_{it}. In the latter case, both lags and leads may be used as instruments. In both cases, it is also possible to use external instruments.

Throughout the paper it is also assumed that in incomplete panels data are missing at random. Under this assumption, all the estimators proposed next may be used with unbalanced panels.

2.3.1 Fixed effects estimators based on alternative quasi difference and mean differenced transformations

The quasi difference transformations usually employed to eliminate fixed effects from multiplicative models may also be applied with the same purpose to the EFRM in (5). A quasi first-difference transformation particularly useful is obtained by performing the following steps: (i) divide both sides of (5) by $\exp\left(x_{it}\theta\right)$; (ii) write the lagged version of (5) and divide it by $\exp\left(x_{i,t-1}\theta\right)$; and (iii) subtract the latter equation from the former in order to obtain

$$\frac{H_1\left(y_{it}\right)}{\exp\left(x_{it}\theta\right)} - \frac{H_1\left(y_{i,t-1}\right)}{\exp\left(x_{i,t-1}\theta\right)} = \exp\left(\alpha_i\right)\left[\exp\left(v_{it}\right) - \exp\left(v_{i,t-1}\right)\right] \equiv v_{it}^{ww}. \tag{15}$$

Clearly, assuming that x_{it} is weakly exogenous, then $E\left(v_{it}^{ww}\mid\alpha_i, x_{i1}, ..., x_{i,t-1}\right) = 0$. Alternatively, assuming that x_{it} is endogenously determined as in (13), then $E\left(v_{it}^{ww}\mid x_{i1}, ..., x_{i,t-2}\right) =$

0 or a set of external instruments may be used. In this framework, x_{it} may contain lagged values of the response variable. Despite its generality, the counterpart of v_{it}^{ww} in the standard exponential regression has been rarely considered in the literature. It was first mentioned by Wooldridge (1997, endnote 2) and then revisited by Windmeijer (2000), who proposed a solution to the computational difficulties identified by the former author for cases where x_{it} include variables that assume only non-negative or non-positive values.[4]

A much more popular quasi-differencing transformation for exponential regression models has been proposed by Chamberlain (1992). The corresponding version for the EFRM is obtained by multiplying (5) by $\exp(x_{i,t-1}\theta)/\exp(x_{it}\theta)$ and then subtracting (5) lagged one period:

$$\frac{\exp(x_{i,t-1}\theta)}{\exp(x_{it}\theta)}H_1(y_{it}) - H_1(y_{i,t-1}) = \exp(x_{i,t-1}\theta + \alpha_i)\left[\exp(v_{it}) - \exp(v_{i,t-1})\right] \equiv v_{it}^c. \quad (16)$$

Again, the weak exogeneity assumption (11) suffices to ensure that $E\left(v_{it}^c \middle| \alpha_i, x_{i1}, ..., x_{i,t-1}\right) = 0$ and x_{it} may contain lagged dependent variables. However, this transformation is not appropriate to deal with cases where x_{it} and v_{it} are correlated, since in that case there is non-separability of $\exp(x_{i,t-1}\theta)$ and $\exp(v_{i,t-1})$ in the right-hand side of (16).

Finally, a third alternative is the mean differenced transformation used by Blundell, Griffith and Windmeijer (2002) for exponential regression models with additive time-varying error term. In the EFRM framework, where $\exp(v_{it})$ appears multiplicatively in the model, a similar transformation is obtained by averaging (5) over time for each individual, then multiplying it by $\exp(x_{it}\theta)/\overline{\exp(x_i\theta)}$ and finally subtracting the resultant equation from (5):

$$H_1(y_{it}) - \frac{\overline{H_1(y_i)}}{\overline{\exp(x_i\theta)}}\exp(x_{it}\theta) = \exp(x_{it}\theta + \alpha_i)\left[\exp(v_{it}) - \frac{\overline{\exp(x_i\theta + v_i)}}{\overline{\exp(x_i\theta)}}\right] \equiv v_{it}^{bgw}, \quad (17)$$

where $\overline{H_1(y_i)}$ is the mean over t of $H_1(y_{it})$ and $\overline{\exp(x_i\theta)}$ and $\overline{\exp(x_i\theta + v_i)}$ are defined in a similar way. In this case, only under the strict exogeneity assumption (10) has v_{it}^{bgw} a zero conditional expectation: $E\left(v_{it}^{bgw} \middle| \alpha_i, x_{i1}, ..., x_{iT}\right) = 0$. Therefore, endogenous covariates

[4] When x_{it} includes such variables, the corresponding estimates for θ go to infinity. A simple way around this problem is to transform x_{it} in deviations from its overall mean.

and lagged dependent variables are not allowed in x_{it}.

These three transformed regression models imply a set of moment conditions of the type $E\left(v^{*}\middle| z\right)=0$, where v^{*} is given by the left-hand side of the equations defining v_{it}^{ww}, v_{it}^{ww} and v_{it}^{ww}. From these moment conditions, a set of orthogonality functions of the type (8) may be generated and the model parameters θ may be estimated by GMM.

2.3.2 Estimators based on direct estimation of the EFRM

The second group of estimators proposed in this paper is based on direct, pooled estimation of the EFRM model given by (5), or its alternative representations (6) and (7). As in the previous case, in general, GMM estimation has to be performed. However, under exogeneity, the estimators may also be obtained by the Exponential-based QML considered by Santos Silva and Tenreyro (2006) as alternative to the most commonly applied Poisson QML.[5] Indeed, subtracting one to both sides of (6) and (7) and multiplying by x_{it}, the left-hand side of those equations are identical to the first-order conditions defining Exponential QML estimators.

The first estimator is a 'pooled random effects' estimator, assuming that x_{it} and α_i are independently distributed. Treating $\alpha_i + v_{it}$ as a single error term and subtracting one to both sides of equation (7), it follows that:

$$\frac{H_1\left(y_{it}\right)}{\exp\left(x_{it}\theta\right)} - 1 = \exp\left(\alpha_i + v_{it}\right) - 1 \equiv v_{it}^{pre}. \tag{18}$$

The contemporaneous exogeneity assumption (12) suffices for consistent estimation of θ. Alternatively, under endogeneity of x_{it}, external instruments may be used.

The second estimator is a 'pooled fixed effects' estimator, allowing x_{it} and α_i to be correlated and interpreting α_i as a vector of individual-specific intercepts to be estimated simultaneously with θ. Estimates for θ and α_i may be obtained by Exponential QLM estimation of (5), which is equivalent to GMM estimation based on orthogonality conditions between the regressors (x_{it} and individual-specific dummies) and the left-hand side

[5] In models with multiplicative heterogeneity, Poisson QML is not well suited to deal with endogenous covariates (Windmeijer and Santos Silva, 1997). Instead, this method would be the appropriate choice for models of the type $H_1\left(y_{it}\right) = \exp\left(x_{it}\theta + \alpha_i\right) + v_{it}$, which, however, do not respect the DGP considered in this paper.

of (6) minus one, that is:

$$\frac{H_1\left(y_{it}\right)}{\exp\left(x_{it}\theta+\alpha_i\right)}-1=\exp\left(v_{it}\right)-1\equiv v_{it}^{pfe}. \tag{19}$$

As shown by Dhaene and Jochmans (2015), there is no incidental parameters problem in this case. The log-likelihood function based on the exponential distribution is defined by

$$LL=\sum_{i=1}^{N}\sum_{t=1}^{T}\left\{\left[-\frac{H_1\left(y_{it}\right)}{\exp\left(x_{it}\theta+\alpha_i\right)}\right]-\log\left[\exp\left(x_{it}\theta+\alpha_i\right)\right]\right\}, \tag{20}$$

implying that the first-order conditions for θ are given by

$$\sum_{i=1}^{N}\sum_{t=1}^{T}x_{it}\left[\frac{H_1\left(y_{it}\right)}{\exp\left(x_{it}\hat{\theta}\right)\widehat{\exp\left(\alpha_i\right)}}-1\right]=0 \tag{21}$$

and for $\exp\left(\alpha_i\right)$ by:

$$\widehat{\exp\left(\alpha_i\right)}=\overline{\left[\frac{H_1\left(y_i\right)}{\exp\left(x_i\hat{\theta}\right)}\right]}. \tag{22}$$

The last result may be used to produce a conditional log-likelihood that is free of nuisance parameters.

A moment indicator for GMM estimation that produces identical estimators to those defined by v_{it}^{pfe} in (19), but does not require estimation of the individual effects, is available when the value zero is not observed for y_i. This alternative formulation is obtained by performing the following steps: (i) multiply both sides of the EFRM (7) by $\overline{\left[\frac{H_1(y_i)}{\exp(x_i\theta)}\right]}^{-1}$; (ii) in the right-hand side of the resulting expression, replace $H_1\left(y_{it}\right)$ by $\exp\left(x_{it}\theta+\alpha_i+v_{it}\right)$ in order to simplify $\overline{\left[\frac{H_1(y_i)}{\exp(x_i\theta)}\right]}^{-1}$ to $\left[\exp\left(\alpha_i\right)\overline{\exp\left(v_i\right)}\right]^{-1}$; and (iii) subtract one to both sides, producing

$$\frac{H_1\left(y_{it}\right)}{\exp\left(x_{it}\theta\right)}\overline{\left[\frac{H_1\left(y_i\right)}{\exp\left(x_i\theta\right)}\right]}^{-1}-1=\frac{\exp\left(v_{it}\right)}{\overline{\exp\left(v_i\right)}}-1\equiv v_{it}^{pfe'}. \tag{23}$$

Consistency of the GMM estimator based on v_{it}^{pfe} or $v_{it}^{pfe'}$ requires the strict exogeneity assumption (10), the endogeneity assumption (14) or the use of appropriate external instruments.

The third estimator is a 'correlated random effects' estimator similar in spirit to the estimators proposed by Papke and Wooldridge (2008) and Wooldridge (2010). In fact, the same flexible functional forms for representing the relationship between α_i and x_{it} may be used. For example, assuming balanced panel data, α_i may be modelled as a linear function of all exogenous variables:

$$\alpha_i = \psi_0 + \bar{z}_i \psi_1 + \eta_i, \tag{24}$$

where $\bar{z}_i \equiv T^{-1} \sum_{t=1}^{T} z_{it}$, $\psi = (\psi_0, \psi_1)$ is a vector of parameters to be estimated and η_i is a disturbance term that is uncorrelated with all the other variables and error terms. Plugging (24) into (6), it follows that[6]

$$\frac{H_1 (y_{it})}{\exp (x_{it}\theta + \psi_0 + \bar{z}_i \psi_1)} - 1 = \exp (\eta_i + v_{it}) - 1 \equiv v_{it}^{cre}. \tag{25}$$

Consistent estimates for θ and ψ may be obtained by using pooled methods and the augmented vector of covariates $(x_{it}, \bar{z}_i)$ and assuming similar assumptions to the 'pooled fixed effects' estimator on the correlation between the covariates and the time-varying heterogeneity. A similar analysis may be carried out in the case of unbalanced data, but based on

$$\alpha_i = \sum_{r=2}^{T} \delta_{T_i,r} \psi_{0r} + \sum_{r=2}^{T} \delta_{T_i,r} \bar{z}_i \psi_{1r} + \eta_i, \tag{26}$$

where T_i is the number of observations available for individual i and $\delta_{T_i,r}$ is a dummy variable which is equal to unity if $T_i = r$ and data exist on the full set of variables. See Wooldridge (2010) for details and alternative expressions for α_i.

Similarly to the previous set of estimators, equations (18), (19) (or (23)) and (25) imply a set of moment conditions of the type $E (v^* | z) = 0$, where v^* is now given by the left-hand side of the equations defining v_{it}^{pre}, v_{it}^{pfe} and v_{it}^{cre}. The resultant GMM estimators are pooled versions, in the last two cases based on a suitable augmented vector of covariates, of the GMM_x (x_{it} exogenous) and GMM_z (x_{it} endogenous) estimators proposed by Ramalho and Ramalho (2015) for dealing with unobserved heterogeneity in the cross-sectional framework.

[6] Without loss of generality, since (24) contains a constant term, it is assumed that $E [\exp (\eta_i)] = 1$.

Although the estimation of partial effects is not the focus of this paper, it is worth mentioning that these three estimators also allow the estimation of those effects under additional assumptions. In particular, it is necessary to assume that the dependence between x_{it} and v_{it} is restricted to the conditional mean, i.e. the right-hand side of (18), (23) and (25) may depend on the covariates but other functions of v_{it} not. Under this new assumption, then the smearing method considered by Ramalho and Ramalho (2015) may be used to consistently estimate partial effects conditional on both α_i and x_{it} (correlated random effects or pooled fixed effects) or only on x_{it} (pooled random effects). See Ramalho and Ramalho (2015) for details on the smearing method, which does not require any distributional assumption on the unobservables.

3 Monte Carlo simulation study

To illustrate the potential bias of each estimator when used in inappropriate settings, the results of a small-scale Monte Carlo simulation study based on 5000 replications are now provided. This study is based on panel logit fractional responses generated from $y_{it} = \frac{\exp(\theta_1 x_{it}+\alpha_i+v_{it})}{1+\exp(\theta_1 x_{it}+\alpha_i+v_{it})}$, where $x_{it} = \pi_1\alpha_i+\pi_2 v_{i,t-1}+\pi_3 v_{it}+\pi_4 z_{it}+\epsilon_{it}$, $\alpha_i \sim \mathcal{N}(-0.5\sigma^2_\alpha, \sigma^2_\alpha)$, $v_{it} \sim \mathcal{N}(-0.5\sigma^2_v, \sigma^2_v)$, $z_{it} \sim \mathcal{N}(0, \sigma^2_z)$, $\epsilon_{it} \sim \mathcal{N}(0, \sigma^2_\epsilon)$, $\theta_1 = 1$ and $\sigma^2_\alpha = \sigma^2_z = \sigma^2_\epsilon = 0.5$. Four experimental designs are considered. Two of them produce a strictly exogenous regressor x_{it} by setting $\pi_2 = \pi_3 = \pi_4 = 0$: design 1 considers $\pi_1 = \{-1, -0.5, -0, 25, 0, 0.25, 0.5, 1\}$ and $\sigma^2_v = 1$, and design 2 considers $\sigma^2_v = \{0.05, 0.5, 1, 1.5, 2, 2.5, 3\}$ and $\pi_1 = 0.5$. Design 3 produces a weakly exogenous x_{it} by setting $\pi_3 = \pi_4 = 0$ and considers $\pi_2 = \{0, 0.15, 0.3, 0.45, 0.6, 0.8, 1\}$, $\sigma^2_v = 1$ and $\pi_1 = 0.5$. Finally, design 4 defines x_{it} as an endogenous regressor with $\pi_3 = \{0, 0.25, ..., 1.5\}$ and sets $\sigma^2_v = 1$, $\pi_1 = 0.5$, $\pi_2 = 0$ and $\pi_4 = 1$. For each design, $(N, T) = (50, 5)$.

In all cases, the six alternative panel data estimators for the structural parameter θ_1 that are based on (15), (16), (17), (18), (23) and (25) are computed, which are designated as, respectively, GMM_{ww}, GMM_c, GMM_{bgw}, GMM_{pre}, GMM_{pfe} and GMM_{cre}. For comparative purposes, it is also computed: the pooled QML logit estimator (QML_p), which is the typical estimator used in the cross-sectional context and does not allow for any kind of heterogeneity; and the pooled fixed effects QML logit estimator of Hausman and Leonard (1997) (QML_{pfe}), which was designed for long panels and does not allow for time-varying unobservables in the index function. GMM_{ww} and GMM_c use as instruments $x_{i,t-1}$ (designs 1-3) or $z_{i,t-1}$ (design 4), while GMM_{bgw}, GMM_{pre}, GMM_{pfe} and GMM_{cre} use $x_{i,t}$ (designs 1-3) or $z_{i,t}$ (design 4). The R package *frmpd*, produced by the authors, is available at the Comprehensive R Archive Network (CRAN) web page (http://cran.r-project.org/web/packages/frmpd/) and may be used to compute all the estimators proposed in the paper.

Figure 1 presents the results obtained for the four designs in terms of: mean parameter estimates across replications; root mean squared error (RMSE); and empirical coverage for a 95% confidence interval based on cluster-robust standard-errors, which was estimated by taking the proportion of the replications where the confidence interval covers the true value of θ_1. To facilitate the reading of the last indicator, the grey area marks an empirical coverage between 93 and 97%.

Figure 1 about here

The first row of Figure 1 shows that the only estimator that is approximately unbiased across all designs is GMM_{ww}. As expected, GMM_c exhibits important biases for large correlations between x_{it} and v_{it}, while GMM_{bgw} only performs well under strict exogeneity. The advantage of GMM_{ww} over GMM_c in terms of RMSE and empirical coverage is also apparent in all designs.

Regarding the pooled GMM estimators, the bias performance of GMM_{pfe} and GMM_{cre} was very similar in all designs. They are biased only in Design 3 (which imposes weak exogeneity) and deal very well with contemporaneous endogeneity. On the other hand, in Designs 1, 2 and 4, GMM_{pfe} displays smaller RMSE and higher empirical coverage than GMM_{ww}. Clearly, whenever strict exogeneity, with or without contemporaneous endogeneity, can be assumed, GMM_{pfe} seems to be a good alternative to GMM_{ww}.

The GMM_{pre} estimator, naturally, is only consistent under the assumption of random effects ($\pi_1 = 0$ in Design 1). Nevertheless, notice how this estimator accommodates very well time-varying heterogeneity: on the one hand, its bias performance is independent on the value of σ_v^2 (Design 2); on the other hand, as the importance of the time-varying heterogeneity increases relative to the fixed effects (π_3 increases in Design 4, while π_1 is kept fixed), it becomes approximately unbiased.

Regarding the QML-based estimators, the QML_p estimator displays very large biases in all experiments (apart from a particular situation where α_i and v_{it} seem to compensate each other) and its empirical coverage is often zero. On the other hand, QML_{pfe} accommodates well the fixed effects (its performance is independent on the value of π_1 in Design 1) but, unless σ_v^2 is very close to zero, cannot cope with time-varying heterogeneity.

Figure 2 considers further experiments based on Design 3, focussing on the $\pi_2 = 1$ case and considering $T = (2, 3, 5, 10, 15, 20, 25, 30, 40, 50)$. The results show that, under weak exogeneity, the bias of estimators that require strict exogeneity decreases substantially as the time dimension of the panel gets large, particularly in the case of the GMM_{pfe} and GMM_{cre}. Nevertheless, even for $T = 50$ these estimators still display some bias. On the other hand, the bias of the QML_p, QML_{pfe} and, after an initial improvement, GMM_{pre}, does not change as T increases, implying that their empirical coverages converge to zero.

Figure 2 about here

4 Empirical application

This section illustrates the usefulness of the GMM estimators proposed in the paper in conducting empirical work. In particular, the focus is on firms' capital structure decisions, which is one of the most debated topics in corporate finance. The balance of this section is as follows. First, the main theories available in this context are summarized. Next, the data is presented. In the last subsection, the empirical results using various OLS, QML and GMM estimators are discussed.

4.1 Capital structure theory

How do firms finance their operations? How should firms finance their operations? What factors influence these choices? How do these choices affect the rest of the economy? These questions have been around for more than 60 years. In particular, historically, the modern theory of capital structure started when Modigliani and Miller published their seminal paper on the topic in 1958. The key insight from their work is that capital structure decisions' are irrelevant when rational investors operate in efficient capital markets. Yet, real-world investors are hardly rational, and capital markets around the world are not always efficient (e.g. Barberis and Thaler, 2005). As such, not surprisingly, soon after the publication of the Modigliani and Miller's (1958) influential paper, a voluminous research on the capital structure topic emerged.

Currently, two main theories seem to provide some understanding on how corporations decide on the use of internal funds, debt finance, and equity issuance. The first, known as the trade-off theory, was originally presented in 1973 by Kraus and Litzenberger. Such theory revolves around the idea that firms examine the costs and benefits of debt when choosing their capital structure. As such, according to the trade-off theory, firms have an "optimal" capital structure (i.e., a target debt ratio), which is defined as a trade-off between tax and other benefits of debt against financial distress and other costs that are consequences of the use of debt (Bradley, Jarrell and Kim, 1984; Graham and Harvey, 2001; Harris and Raviv, 1991). In contrast, the pecking order theory, which stems from the work of Myers (1984), is rooted on the idea that managers are better informed about the true value of their firm's assets-in-place and respective investment opportunities than outside investors. Under this set up, raising external financing must be costly, since it generates an adverse selection problem. As a result, if the pecking order theory holds, a firm's debt ratio simply reflects a hierarchy of financing sources whereby internal financing is preferred over debt, and debt is preferred over equity (Myers, 1984; Myers and Majluf, 1984).

In 2002, Baker and Wurgler wrote an influential paper that shaped the market timing theory, which has become an important alternative for explaining firms' capital structure decisions. Their work draws on the earlier contributions of Lucas and McDonald (1990) and Korajczyk, Lucas and McDonald (1992), who show that the decision to issue equity is contingent on market performance. In particular, under the market timing theory, managers monitor current market conditions and opt to issue debt or equity depending on which is overvalued in the short-run. Thus, in the end, firms' observed capital structure is simply the cumulative outcome of a long series of incremental financing decisions.

Over the years, many empirical researchers have employed different econometric techniques to study the main determinants of capital structure choices. The most common procedure has been to use panel data linear regression models, in which the dependent variable is a ratio that compares some form of debt (e.g., total debt, long-term debt, or short-term debt) to the firm's total assets or financial capital (debt plus equity).[7] As a result, the dependent variable in these studies is typically a proportion, defined and observed only on the standard unit interval. Moreover, it assumes often the value zero and rarely the value one. Thus, Ramalho and Silva (2009, 2013) question the use of linear models in empirical capital structure research, and present clear evidence on the biases that may arise from such practice.

The present paper adds to this early work by considering panel data fractional response models to study the determinants of firms' capital structure choices, something not addressed by Ramalho and Silva's (2009, 2013) cross-sectional models. For comparison reasons, other classes of models are also estimated: panel data linear regression models, which are the prevalent models in empirical capital structure studies; linear-fractional models based on log-odds and other transformations, which take into account the fractional nature of y_{it} but require ad-hoc solutions for the zero-debt firms problem; and standard pooled QML estimators for fractional responses, which do not acommodate time-varying-heterogeneity.

4.2 Data

The data used in this study comes from the Amadeus database and comprises financial information and other characteristics of 620 Portuguese large firms for the 2007-2011 period. All firms in five industries (*Manufacturing*, *Construction*, *Wholesale and Retail Trade*, *Transportation and Storage* and *Accommodation and Food Service Activities*) with non-negative equity and without missing data were included in the sample, which yields an unbalanced panel of 1843 observations. The number of observations per firm ranges from one to five.

As a measure of financial leverage (*Leverage*), the ratio of long-term debt (defined as the total company's debt due for repayment beyond one year) to long-term capital assets

[7] Tests of the market timing theory are somewhat different. For instance, Baker and Wurgler (2002) employ the external finance weighted-average of their sample firm's past market-to-book values as their main dependent variable. Alti (2006), Leary and Roberts (2005), and Kayhan and Titman (2004) also resort to particular methodological choices when exploring the market timing theory.

(defined as the sum of debt and equity) is considered. The explanatory variables used are four of the most commonly cited determinants of capital structure policies: 1) *Growth* (computed as the yearly percentage change in sales); 2) *Size* (computed as the natural logarithm of sales); 3) *Profitability* (computed as the ratio of earnings before interest and taxes to total assets); and 4) *Tangibility* (computed as the proportion of tangible assets in total assets). Some of the explanatory variables are expected to have a positive impact on leverage ratios (*e.g.*, *Profitability*, in the case of the trade-off theory; *Growth*, in the case of the pecking-order theory; and *Tangibility* and *Size*, in both cases), while others are expected to have a negative effect (*e.g.*, *Growth*, in the trade-off theory; and *Profitability*, in the pecking-order theory); see inter alia Ramalho and Silva (2009) for a theoretical explanation of these expected effects. In all models a set of year dummies is included to control for changes occurring over time that are common to all firms.

Table 1 reports descriptive statistics for all variables. As the minimum value observed for the dependent variable *Leverage* is zero (16.2% of the sample correspond to zero-debt observations) but the maximum is lower than one, all six estimators proposed in the paper may be used. Actually, the characteristics of this data set (unbalanced data, many observations for $y_{it} = 0$ but none for $y_{it} = 1$) makes this an ideal application for illustrating the usefulness of the panel data EFRM.

Table 1 about here

4.3 Estimation and results

4.3.1 Linear regression models

The most common regression model for panel data used in empirical capital structure studies is given by

$$y_{it} = x_{it}\theta + \alpha_i + v_{it}, \tag{27}$$

which corresponds to the index of the exponential function in (4). The main reason for the popularity of the linear regression model (27) is the very straightforward implementation of the three most common types of panel data estimators: pooled, random effects and fixed effects. However, in regression analysis of leverage ratios, this simplicity is obtained at the cost of ignoring the bounded nature of the fractional response, which may generate predictions outside the unit interval, specially in cases where some fractions are observed close to or at the boundaries of the interval, as occurs in the dataset in analysis.

The results for the parameter estimates and respective cluster-robust standard errors obtained by the three methods are presented in Table 2. There is a clear distinction between the results produced by the models that assume no correlation between α_i and x_{it} (the pooled and the random effects models) and the only model that allows for that correlation (the fixed effects model): no variable is simultaneously statistically significant in both types of models. A Hausman test statistic of 33.91, associated with a p-value of 0.000, leads to the rejection of the null hypothesis of random effects and to the consequent selection of the fixed effect estimator over the alternatives. However, this model does not show support to any of the capital structure theories: only the variable *Size* is statistically significant, but its sign is contrary to that predicted by both the trade-off and the pecking-order theories.

Table 2 about here

4.3.2 Linear-fractional regression models

Unlike the previous estimators, linear-fractional regression models have into account the fractional nature of the response variable and are based on the same model (3) that was used to generate the EFRM. However, they employ a different transformation, using the link function $H(\cdot) = G(\cdot)^{-1}$ to obtain a linearized regression model that is very simple to estimate:

$$H(y_{it}) = x_{it}\theta + \alpha_i + v_{it}. \tag{28}$$

Any standard fractional regression model for which the link function is defined may be considered in this framework. Next, in addition to the logit and cloglog models addressed in this paper, for which $H(y_{it}) = \ln\left(\frac{y_{it}}{1-y_{it}}\right)$ (logit model) and $H(y_{it}) = \ln[-\ln(1-y_{it})]$ (cloglog model), the probit and loglog models are also considered. For the probit model, $G(\cdot) = \Phi(x_{it}\theta + \alpha_i + v_{it})$ and $H(y_{it}) = \Phi^{-1}(y_{it})$, where $\Phi(\cdot)$ and Φ^{-1} are the cumulative normal and the inverse cumulative normal distributions, respectively. For the loglog model, $G(\cdot) = \exp\{-\exp[-(x_{it}\theta + \alpha_i + v_{it})]\}$ and $H(y_{it}) = -\ln[-\ln(y_{it})]$. Because in no case are the link functions defined for the boundary values of 0 and 1, all the observations with $y_{it} = 0$ are discarded, which implies that 90 out of 620 firms (or 299 out of 1843 observations) are dropped from the sample.

The results, reported in Table 3, show that all four models lead to similar conclusions in terms of individual significance and effect direction within each class of pooled, random effects and fixed effects estimators, with the exception of the cloglog model with fixed effects. Thus, it seems that the effects found for each covariate, at least in terms of significance and sign, are robust to the specific functional form G chosen for the conditional expectation of y_{it}.

Table 3 about here

Taking into account the fractional nature of the response variable did not change at all the set of statistically significant covariates in the pooled models. Conversely, random and fixed effects models now suggest that only the variable *Profitability* is relevant for explaining capital structure decisions, which is in accordance with the pecking-order theory and in opposition to the trade-off theory. Nevertheless, the Hausman tests applied in the four models still suggest that fixed effects estimators are preferable (in all cases, the p-value of the test is 0.000).

4.3.3 Fractional regression models

Three types of fractional regression models are now considered, all of which are estimated by QML as is typical in cross-sectional applications. The first is based on direct estimation of equation (1) and corresponds to the QML_p estimator considered previously in the Monte Carlo study. The second is based on equation (2), with the individual effects α_i being explicitly estimated, and corresponds to the QML_{pfe} estimator also included in the Monte Carlo section. The third, which is designated by QML_{cre}, is the estimator proposed by Wooldridge (2010) and is also based on equation (2), but with α_i replaced by expression (26). For the first two sets of estimators G is again specified alternately as a logit, probit, cloglog and loglog function, while QML_{pfe} is only applicable to the probit case. Unlike the previous estimators (and the GMM estimators of the next section), no QML estimator allows for time-varying heterogeneity.

Table 4 reports the results obtained for the QML estimators. As in the two previous sections, the largest number of relevant explanatory variables is revealed by the simple pooled estimators: once again, *Growth*, *Profitability* and *Tangibility* are statistically significant and have the signs predicted by the pecking-order theory. Also similarly to the estimators considered in the previous sections, the number of relevant covariates decreases when fixed effects are included in the regression. Actually, with fixed effects QML estimators, either all explanatory variables are not statistically significant (logit, probit and cloglog models) or only *Size* is significant but displays an effect contrary to that predicted by both capital structure theories (loglog model). This uninteresting performance of QML_{pfe} estimators may be due to the short time dimension of the panel available, which affects their consistency, and/or to the existence of time-varying heterogeneity. Actually, given the similar conclusions also achieved for the QML_{cre} estimator, omitted time-varying factors are probably the main reason for the results obtained.

Table 4 about here

4.3.4 Exponential-fractional regression models

The GMM estimators proposed in this paper require the assumption of a logit or cloglog specification for the G function in (3). Given that the results obtained in the previous sections were clearly stable across alternative specifications of G, this limitation of GMM estimators based on the EFRM does not seem to be relevant in this particular application. Actually, for the same reason, the alternative GMM estimators developed in the paper are run only for the logit model.

The results reported in Table 5 confirm that *Profitability* has a significant negative influence on the proportion of debt issued by a firm, as predicted by the pecking-order theory. This effect had already been found for all the pooled and random effects models estimated in previous sections, but was not clear in fixed effects regressions, with linear-fractional regression models also indicating the statistical significance of *Profitability*, but linear and fractional regression models showing the opposite. Because, from the four groups of fixed effects estimators used in this study, the GMM estimators are the only that, simultaneously, have into account the fractional nature of the response variable, accommodate the value zero of the response variable and are robust to time-varying heterogeneity, it may be then concluded that, in this particular application: (i) dropping the observations with $y_{it} = 0$ from the sample does not cause any special problem, which suggests that zero and positive values of leverage ratios are generated by the same DGP; (ii) it is clearly important to have into account the bounded and fractional nature of y_{it}, since the fixed effects linear regression model is unable to detect the statistical relevance of *Profitability*; and (iii) it seems to exist time-varying heterogeneity, since not allowing for its presence leads to a similar failure of the QML estimators.

Table 5 about here

5 Conclusion

In this paper six new panel data estimators for fractional responses with boundary observations were proposed. All estimators are consistent for the structural parameters under different assumptions for the relationship between observables and unobservables. Five of them allow the covariates and the individual effects to be correlated, while the other (GMM_{pre}) produces consistent estimators only under a random effects assumption. Three estimators do not require strict exogeneity of the covariates (GMM_{ww}, GMM_c and GMM_{pre}). Four estimators (GMM_{ww}, GMM_{pre}, GMM_{pfe} and GMM_{cre}) are able to accommodate endogenous explanatory variables. Two estimators (GMM_{ww} and GMM_c) may be applied to dynamic panel data models. Finally, five estimators apply indistinctly to balanced and unbalanced panel data, while the other (GMM_{cre}) requires a minor adaptation.

The GMM_{ww} estimator is thus the only estimator that, simultaneously, is robust to both time-variant and time-invariant heterogeneity, accommodates endogenous explanatory variables and can be applied to dynamic panel data models. The Monte Carlo simulation study carried out illustrates its good performance across a range of alternative scenarios. Another useful estimator seems to be the GMM_{pfe} estimator, namely when strict exogeneity can be assumed, in which case it tends to display a better performance in terms of dispersion and empirical coverage than all the other estimators. The empirical application dedicated to capital structure decisions emphasized the differences between the proposed GMM estimators and a large range of alternatives.

References

Alti, A. (2006), "How persistent is the impact of market timing on capital structure", *Journal of Finance*, 61, 1681–1710.

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

Barberis, D. and Thaler, R. (2005), "A survey of behavioral finance", in R. Thaler (ed) *Advances in Behavioral Finance, Volume II*, New York: Princeton University Press.

Blundell, R., Griffith, R., and Windmeijer, F.A.G. (2002), "Individual effects and dynamics in count data models", *Journal of Econometrics*, 108, 113-131.

Bradley, M., Jarrell, G.A. and Kim, E.H. (1984), "On the existence of an optimal capital structure: Theory and evidence", *Journal of Finance*, 39, 857–878.

Chamberlain, G. (1992), "Comment: sequential moment restrictions in panel data", *Journal of Business and Economic Statistics*, 10(1), 20-26.

Dhaene, G. and Jochmans, K. (2015), "Profile-score adjustments for incidental-parameter problems", mimeo.

Elsas, R. and Florysiak, D. (2015), "Dynamic capital structure adjustment and the impact of fractional dependent variables", *Journal of Financial and Quantitative Analysis*, forthcoming.

Graham, J.R. and Harvey, C.R. (2001), "The theory and practice of corporate finance: Evidence from the field", *Journal of Financial Economics*, 60, 187–243.

Harris, M. and Raviv, A. (1991), "The theory of capital structure", *Journal of Finance*, 46, 297–355.

Hansen, L. P. (1982), "Large sample properties of generalised method of moments estimators", *Econometrica*, 50(4), 1029-1054.

Hausman, J.A. and Leonard, G.K. (1997), "Superstars in the national basketball association: economic value and policy", *Journal of Labor Economics*, 15(4), 586-624.

Heckman, J.J. (2000), “Causal parameters and policy analysis in Economics: a twentieth century retrospective”, *Quarterly Journal of Economics*, 115(1), 45-97.

Heckman, J.J. (2001), “Micro data, heterogeneity, and the evaluation of public policy: Nobel lecture”, *Journal of Political Economy*, 109, 673-748.

Kayhan, A. and Titman S. (2004), "Firms' histories and their capital structures", Louisiana State University and University of Texas, Working paper.

Kraus, A. and Litzenberger, R.H. (1973), "A state-preference model of optimal financial leverage", *Journal of Finance*, 28, 911–922.

Korajczyk, R.A., Lucas, D.J. and McDonald, R.L. (1992), "Equity issues with time-varying asymmetric information", *Journal of Financial and Quantitative Analysis*, 27, 397–417.

Leary, M. and Roberts M.R. (2005), "Do firms rebalance their capital structures?", *Journal of Finance*, 60, 2575-2619.

Loudermilk, M.S. (2007), “Estimation of fractional dependent variables in dynamic panel data models with an application to firm dividend policy”, *Journal of Business and Economic Statistics*, 25(4), 462-472.

Lucas, D.J. and McDonald, R.L. (1990), "Equity issues and stock price dynamics", *Journal of Finance*, 45, 1019–1043.

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

Mullahy, J. (1997), “Instrumental-variable estimation of count data models: applications to models of cigarette smoking bahavior”, *Review of Economics and Statistics*, 79(4), 586-593.

Myers, S.C. (1984), "The capital structure puzzle", *Journal of Finance*, 39, 574–592.

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

Ospina, R. and Ferrari, S.L.P. (2012), “A general class of zero-or-one inflated beta regression models”, *Computational Statistics and Data Analysis*, 56(6), 1609-1623.

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.

Papke, L. and Wooldridge, J. (2008), “Panel data methods for fractional response variables with an application to test pass rates”, *Journal of Econometrics*, 145(1-2), 121-133.

Ramalho, E.A. and Ramalho, J.J.S. (2015), “Moment-based estimation of nonlinear regression models with boundary outcomes and endogeneity, with applications to nonnegative and fractional responses”, *Econometric Reviews*, forthcoming (DOI: 10.1080/07474938.2014.976531).

Ramalho, E.A., Ramalho, J.J.S. and Henriques, P.D. (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.M.R. (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.

Ramalho, J.J.S. and Silva, J.V. (2013), “Functional form issues in the regression analysis of corporate capital structure”, *Empirical Economics*, 44(2), 799-831.

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

Windmeijer, F. (2000), “Moment conditions for fixed effects count data models with endogenous regressors”, *Economics Letters*, 68(1), 21-24.

Windmeijer, F. (2008), "GMM for panel count data models", in L. Mátyás and P. Sevestre (eds.) *The Econometrics of Panel Data*, 3rd Edition, Advanced Studies in Theoretical and Applied Econometrics Volume 46, pp 603-624.

Windmeijer, F.A.G. and J.M.C. Santos Silva (1997), "Endogeneity in count data models: an application to demand for health care", *Journal of Applied Econometrics*, 12, 281-294.

Wooldridge, J.M. (1997), "Multiplicative panel data models without the strict exogeneity assumption", *Econometric Theory*, 13, 667-678.

Wooldridge, J.M. (2010), "Correlated random effects models with unbalanced panels", mimeo.

Table 1: Descriptive statistics

Variable	Mean	Min	Max	St. Dev.
Leverage	0.236	0	0.982	0.230
Growth	5.379	-82.032	99.629	20.701
Size	11.058	6.359	15.013	1.092
Profitability	0.056	-0.426	0.626	0.080
Tangibility	0.265	0	0.990	0.204

Table 2: Linear regression models

	Pooled	Random Effects	Fixed Effects
Growth	0.0010***	0.0003	0.0002
	(0.0003)	(0.0002)	(0.0002)
Size	0.0006	-0.0057	-0.0584*
	(0.0071)	(0.0076)	(0.0350)
Profitability	-0.4189***	-0.2543***	-0.1185
	(0.0905)	(0.0738)	(0.0935)
Tangibility	0.2395***	0.1935***	0.0205
	(0.0428)	(0.0412)	(0.0755)
Number of observations	1843	1843	1843
Number of firms	620	620	620

Notes: cluster-robust standard errors in parentheses; ***, ** and * denote coefficients which are significant at 1%,5% or 10%, respectively.

Table 3: Linear-fractional regression models

	Logit			Probit		
	Pooled	Random Effects	Fixed Effects	Pooled	Random Effects	Fixed Effects
Growth	0.0010***	0.0030	0.0013	0.0045***	0.0011	0.0005
	(0.0028)	(0.0021)	(0.0024)	(0.0013)	(0.0009)	(0.0011)
Size	0.0634	-0.0894	-0.2705	-0.0122	-0.0289	-0.1574
	(0.0799)	(0.0809)	(0.2899)	(0.0365)	(0.0377)	(0.1452)
Profitability	-0.4052***	-3.1236***	-2.1690**	-1.8810***	-1.3575***	-0.8977**
	(1.1197)	(0.9320)	(1.0598)	(0.4967)	(0.3818)	(0.4474)
Tangibility	1.0460***	0.3933	-0.4691	0.6173***	0.2938*	-0.1765
	(0.4147)	(0.3837)	(0.5647)	(0.1915)	(0.1783)	(0.2907)
Number of observations	1544	1544	1544	1544	1544	1544
Number of firms	530	530	530	530	530	530

Notes: cluster-robust standard errors in parentheses; ***, ** and* denote coefficients which are significant at 1%,5% or 10%, respectively.

Table 3 (cont.): Linear-fractional regression models

	Cloglog			Loglog		
	Pooled	Random Effects	Fixed Effects	Pooled	Random Effects	Fixed Effects
Growth	0.0088***	0.0028	0.0011	0.0038***	0.0008	0.0005
	(0.0025)	(0.0020)	(0.0022)	(0.0012)	(0.0008)	(0.0009)
Size	-0.0721	-0.0898	-0.2079*	0.0052	-0.0157	-0.1716
	(0.0742)	(0.0743)	(0.2641)	(0.0308)	(0.0330)	(0.1273)
Profitability	-3.6556***	-2.8757***	-2.0184	-1.6029***	-1.1012***	-0.7016***
	(1.0461)	(0.8868)	(1.0068)	(0.4149)	(0.3058)	(0.3617)
Tangibility	0.8794**	0.3317	-0.3927	0.5939***	0.2757	-0.1799
	(0.3795)	(0.3462)	(0.4801)	(0.1711)	(0.1747)	(0.3159)
Number of observations	1544	1544	1544	1544	1544	1544
Number of firms	530	530	530	530	530	530

Notes: cluster-robust standard errors in parentheses; ***, ** and* denote coefficients which are significant at 1%,5% or 10%, respectively.

Table 4: Standard fractional regression models

	Logit		Probit			Cloglog		Loglog	
	QMLp	QMLpfe	QMLp	QMLpfe	QMLcre	QMLp	QMLpfe	QMLp	QMLpfe
Growth	0.0055***	0.0012	0.0032***	0.0008	0.0008	0.0047***	0.0008	0.0029***	0.0008
	(0.0015)	(0.0017)	(0.0009)	(0.0010)	(0.0008)	(0.0013)	(0.0013)	(0.0008)	(0.0009)
Size	0.0027	-0.3686	0.0017	-0.2251	-0.1929*	0.0022	-0.2721	0.0017	-0.2229*
	(0.0405)	(0.2616)	(0.0236)	(0.1486)	(0.1171)	(0.0352)	(0.2037)	(0.0211)	(0.1327)
Profitability	-2.4899***	-1.0985	-1.5048***	-0.6190	-0.4722	-2.0544***	-0.8924	-1.4381***	-0.5260
	(0.5785)	(0.7729)	(0.3324)	(0.4379)	(0.3256)	(0.5016)	(0.6372)	(0.2923)	(0.3793)
Tangibility	1.2597***	0.1998	0.7449***	0.1189	0.0415	1.0722***	0.0125	0.6905***	0.1955
	(0.2189)	(0.5084)	(0.1301)	(0.2842)	(0.2402)	(0.1836)	(0.3112)	(0.1230)	(0.2870)
Number of observations	1843	1843	1843	1843	1843	1843	1843	1843	1843
Number of firms	620	620	620	620	620	620	620	620	620

Notes: cluster-robust standard errors in parentheses; ***, ** and* denote coefficients which are significant at 1%,5% or 10%, respectively.

Table 5: Exponential-fractional regression models

	Logit				
	GMMww	GMMc	GMMbgw	GMMpre	GMMcre
Growth	-0.0002	0.0001	-0.0017	0.0020	0.0001
	(0.0033)	(0.0027)	(0.0024)	(0.0030)	(0.0018)
Size	0.4158	1.4326	-0.1478	0.0604	-0.3317
	(0.5148)	(1.0375)	(0.3371	(0.0805)	(0.2812)
Profitability	-4.5678**	-7.2847***	-1.1881	-4.2104***	-1.6001*
	(2.0394)	(1.8692)	(0.8635)	(1.1353)	(0.9449)
Tangibility	0.2916	0.1354	-0.7104	0.6750**	-0.5910
	(2.0167)	(2.5127)	(0.5685)	(0.3032)	(0.6779)
Number of observations	1157	1157	1843	1843	1681
Number of firms	441	441	620	620	458

Notes: cluster-robust standard errors in parentheses; ***, ** and* denote coefficients which are significant at 1%,5% or 10%, respectively.

Figure 1: Monte Carlo results for alternative estimators of structural parameters (N = 50 and T = 5)

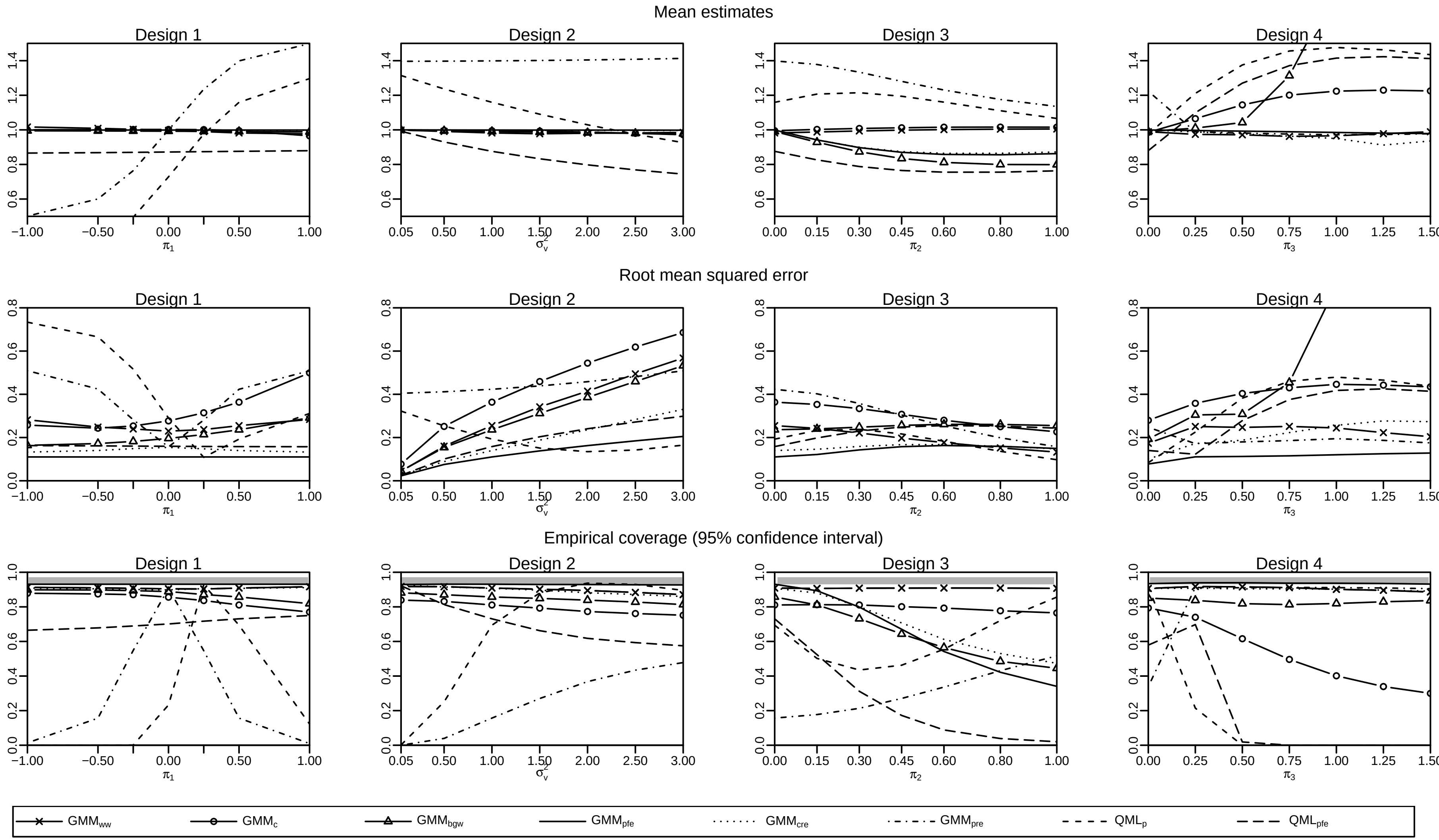

Figure 2: Monte Carlo results for Design 3 with alternative values for T