

In search of robust methods for dynamic panel data models in empirical corporate finance

Viet Anh Dang^a, Minjoo Kim^{b,*}, Yongcheol Shin^c

^aManchester Business School, University of Manchester, Manchester, M15 6PB, UK. +44(0)161 275 0438.

^bAdam Smith Business School/Economics, University of Glasgow, Glasgow, G12 8QQ, UK. +44(0)141 330 7772.

^cDepartment of Economics and Related Studies, University of York, York, YO10 5DD, UK. +44(0)190 443 3757.

Abstract

We examine which methods are appropriate for estimating dynamic panel data models in empirical corporate finance. Our simulations show that the instrumental variable and GMM estimators are unreliable, and sensitive to the presence of unobserved heterogeneity, residual serial correlation, and changes in control parameters. The bias-corrected fixed-effects estimators, based on an analytical, bootstrap, or indirect inference approach, are found to be the most appropriate and robust methods. These estimators perform reasonably well even in models with fractional dependent variables censored at [0,1]. We verify these results in two empirical applications, on dynamic capital structure and cash holdings.

Keywords: Dynamic panel data estimation; GMM; bias correction; capital structure; cash holdings.

JEL Classification: G30, C23.

***Corresponding author:** Minjoo Kim. Address: Adam Smith Business School/Economics, University of Glasgow, G12 8QQ, UK. Email: Minjoo.Kim@glasgow.ac.uk. Tel: +44(0)141 330 7772. Fax: +44(0)141 330 4939.

Email addresses: vietanh.dang@mbs.ac.uk (Viet Anh Dang), Minjoo.Kim@glasgow.ac.uk (Minjoo Kim), yongcheol.shin@york.ac.uk (Yongcheol Shin)

1. Introduction

Many empirical studies in corporate finance use dynamic panel data models to investigate the dynamic behavior of a financial policy of interest. In the corporate payout literature, several studies have examined the degree of dividend smoothing by estimating Lintner's (1956) partial adjustment model (e.g., Brav et al., 2005; Skinner, 2008; Andres et al., 2009). In the capital structure literature, researchers have used this dynamic model extensively to study how quickly firms adjust toward their long-run target leverage ratios.¹ The use of dynamic panels is also common in other areas of corporate finance.²

Despite the growing popularity of dynamic panel data models, they are difficult to estimate due to the likely presence of firm fixed effects and several complexities in empirical corporate finance, such as unobserved heterogeneity and endogeneity, residual serial correlation, and the fractional nature of the dependent variable. To begin with, due to the correlation between the fixed effects and the lagged dependent variable, the pooled OLS (hereafter POLS) estimator is biased and inconsistent. The fixed-effects (hereafter FE) method also estimates these models with a finite-sample bias (Nickell, 1981). Previous simulation results (e.g., Judson and Owen, 1999) suggest that this bias is likely to be substantial for corporate finance studies, which typically analyze annual company data over a relatively short period. The econometrics literature has advanced two main approaches to deal with this bias. The first involves using instruments for the lagged dependent variable, and comprises five methods: the just-identified instrumental variable estimator (hereafter AH-IV) (Anderson and Hsiao, 1981), the first-difference generalized methods of moments estimator (hereafter FD-GMM) (Arellano and Bond, 1991), the system GMM estimator (hereafter SYS-GMM) (Blundell and Bond, 1998), and the long-difference GMM estimator (hereafter LD-GMM or LDP-GMM, depending on long-difference parameters used) (Hahn et al., 2007; Huang and Ritter, 2009). The second approach, consisting of three estimators, corrects for the estimation bias either analytically, or by simulation. Specifically, these

¹See Jalilvand and Harris (1984), Ozkan (2001), de Miguel and Pindado (2001), Fama and French (2002), Flannery and Rangan (2006), Kayhan and Titman (2007), Antoniou et al. (2008), Lemmon et al. (2008), Byoun (2008), Huang and Ritter (2009), Öztekin and Flannery (2012), Faulkender et al. (2012), and Warr et al. (2012).

²See, for example, studies of cash holdings by Opler et al. (1999), Ozkan and Ozkan (2004), and Bates et al. (2009), studies of corporate investment by Aivazian et al. (2005a,b), studies of debt maturity by Ozkan (2000) and Antoniou et al. (2006), and studies of corporate governance by Wintoki et al. (2012).

estimators either develop bias correction formulas in the (fixed-effects) least-squares dummy variable model (hereafter LSDVC) (Kiviet, 1995; Bruno, 2005), or approximate the bias function and search for unbiased estimates using an iterative bootstrap-based correction procedure (hereafter BC) (Everaert and Pozzi, 2007), or a simulation-based indirect inference method (hereafter II) (e.g., Gouriéroux et al., 2010). Although these advanced methods should, in theory, reduce the POLS and FE bias, little is known about their performance in the presence of the complex issues listed above. In what follows, we briefly discuss the possible sources of those issues and their effects on the IV/GMM and bias-corrected estimators.

First, unobserved heterogeneity and endogeneity, caused by non-zero correlation between the firm fixed effects and a regressor, is a common problem in corporate research (Roberts and Whited, 2011; Wintoki et al., 2012). This problem may affect the performance of the estimators that assume the strict exogeneity of the explanatory variables. The second issue is the likely presence of residual autocorrelation, which violates one of the most important assumptions of the IV/GMM estimators, and renders their instruments invalid (Arellano and Bond, 1991). In empirical corporate research, serial correlation may be caused by the persistence of the financial variable (Lemmon et al., 2008), the presence of measurement errors (Welch, 2011; Roberts and Whited, 2011), or the use of an incorrect functional form (e.g., non-linear versus linear models).³ The third problem concerns the measurement of the dependent financial variable. The financial policy variable under consideration (e.g., leverage or debt maturity) can be a ratio bounded by the unit interval $[0, 1]$. Since the IV/GMM and bias-corrected approaches listed above were originally developed for continuous, unbounded dependent variables, their properties may be affected if the dependent variable is fractional (Loudermilk, 2007).

In this paper, we examine which of the existing estimators are most appropriate and robust for dynamic panels in empirical corporate finance, especially in the likely presence of unobserved heterogeneity, autocorrelation, and fractional dependent variables. As mentioned above, we consider five IV/GMM estimators, AH-IV, FD-, SYS-, LD-, and LDP-GMM, as well as three bias-corrected estimators, LSDVC, BC, and II. We also examine an augmented

³There is evidence of residual autocorrelation in empirical research. The results of our studies on capital structure and cash holdings confirm that the test for no serial correlation is frequently rejected.

doubly-censored Tobit estimator (termed DPF by Elsas and Florysiak, 2014) that accounts for the fractional nature of the dependent variable (Loudermilk, 2007). We conduct Monte Carlo simulation studies and empirical applications in order to examine the performance of these estimators.

Our simulation studies show that the bias-corrected estimators, LSDVC, BC, and II, are generally the most appropriate and robust methods for dynamic panel data models in empirical corporate finance. These estimators estimate the coefficient on the lagged dependent variable (i.e., the autoregressive coefficient) and those on the explanatory variables with the most accuracy and efficiency. Among the three, BC performs well in regressions with residual autocorrelation and in specifications with high lag orders. In a special case where the dependent variable is a ratio, censored at 0 and 1, LSDVC, BC, and II may still provide reasonable estimates with a moderate amount of bias, although, at a high percentage of censoring (e.g., more than 20%), DPF emerges as the most robust method.

Our results further suggest that the IV/GMM estimators are outperformed by the bias-corrected methods. The IV/GMM estimates, especially those for the autoregressive coefficient, tend to be unreliable in most of our simulation experiments. Crucially, these methods are very sensitive to the presence of unobserved heterogeneity and serially correlated errors where their instruments become invalid. There are only a few conditions under which these methods would be useful. For example, SYS-GMM could be used for regressions without unobserved heterogeneity, endogeneity, and autocorrelation. In empirical research, however, these conditions are rather restrictive and unlikely to be met. In short, our paper highlights the potential drawbacks of using the IV/GMM estimators in empirical corporate finance.

We verify our simulation results using two empirical applications to dynamic capital structure and cash holdings. In the first application, the dependent variable, leverage, is a ratio bounded by the unit interval, while, in the second, the dependent variable, cash holdings, is measured by the natural logarithm of the cash-to-net-assets ratio, i.e., a continuous, unbounded variable. In these applications, one of our main objectives is to estimate the speeds with which firms adjust toward their target leverage and cash holdings, respectively. We find that LSDVC, BC, and II produce the most plausible estimates of the speeds of dynamic leverage and cash

adjustments, as well as the most reasonable coefficients on the explanatory variables, consistent with prior theoretical predictions. In the leverage application, these bias-corrected methods obtain similar speeds of leverage adjustment, ranging between 24% and 28%. DPF, the only method that explicitly accounts for the fractional nature of leverage, yields a similar estimate of 27%. The estimates obtained using FD-, SYS-, and LD-GMM vary between 15% and 18%, putting them very close to the (biased) POLS estimate; these estimates are unreliable because their fundamental assumptions of instrument validity and no autocorrelation are both violated according to our diagnostic tests. In the cash application, we find similar results regarding the performance of the estimators: LSDVC, BC, and II again produce similar estimates of the speed of cash adjustment, at 48-49%.

Our study is related to previous simulation studies in the econometrics literature (Kiviet, 1995; Judson and Owen, 1999; Bun and Kiviet, 2003; Bruno, 2005; Bun and Carree, 2006; Everaert and Pozzi, 2007). Our simulations maintain the rigor of these studies in terms of properly controlling for two key parameters in dynamic panel data models, namely the magnitude of the fixed effects relative to that of the idiosyncratic error (i.e., the loading factor) and the explanatory power of the regressors relative to that of the disturbances (i.e., the signal-to-noise ratio).⁴ However, we extend these general simulation studies by considering issues relevant to empirical corporate finance. We use data-generating processes that mimic actual company panel data and, moreover, explicitly allow for unobserved heterogeneity, endogeneity, and residual serial correlation. We also conduct additional simulation experiments to examine the properties of the alternative econometric methods when the dependent financial variable of interest is fractional and bounded by the unit interval. Hence, our findings and conclusions are directly applicable to empirical research in corporate finance.

Finally, our study is related to a recent simulation study by Flannery and Hankins (2013, hereafter FH).⁵ However, our study differs from, and improves on, their analysis in many im-

⁴In empirical capital structure research, Lemmon et al. (2008) show that the variation in leverage is mainly explained by the firm fixed effects (60%), as opposed to the independent variables (18%). Hence, it is important to examine the impact of the relative magnitude of the fixed effects on the properties of the estimators.

⁵In a contemporaneous study, Zhou et al. (2014) examine a method of (linear) bias correction for the estimate of the speed of adjustment (SOA) in dynamic capital structure models. They further propose a global minimum variance (GMV) combined estimator to approximate a consensus SOA estimate, which is a GMV-weighted average of the (bias-corrected) estimates obtained using six popular baseline estimators, including OLS, FE, FD-,

portant respects. First, our benchmark experiment explicitly controls for the loading factor and the signal-to-noise ratio, two fundamental control parameters in dynamic panel data models (Kiviet, 1995). Under a more rigorous simulation design, we show that the IV/GMM estimators are sensitive to the control parameters and perform less well than in FH's simulations. Second, while FH only examine a bias-corrected estimator based on an analytical approach (LSDVC), we further consider two recently developed bias correction methods based on iterative sampling algorithms, namely BC and II. Compared to LSDVC, these methods rely on less restrictive assumptions, and are computationally less demanding, especially in the case of BC. We find that both perform well with BC being least affected by autocorrelation. Third, we compare the IV/GMM and bias-corrected estimators with DPF in an important simulation exercise allowing for the fractional nature of the dependent variable (Loudermilk, 2007; Elsas and Florysiak, 2014). Inconsistent with FH's conjecture, we show the non-negligible effect of severe censoring and the relevance of DPF in this case. Fourth, while FH conclude that LSDVC and SYS-GMM are the most appropriate methods, we observe poor performance from SYS-GMM and the related IV/GMM estimators (especially in the presence of autocorrelation). We, thus, caution against the use of these methods and recommend the more robust bias-corrected estimators. Finally, we are able to corroborate our simulation results using empirical applications to two highly relevant areas of corporate finance, capital structure and cash holdings. In sum, our study provides new simulation and empirical results with general and relevant implications for future empirical corporate research.

The remainder of our paper is organized as follows. In Section 2, we discuss the motivation for using dynamic panel data models in empirical corporate finance and review the existing methods for estimating such models. In Section 3, we conduct Monte Carlo simulation studies in order to examine the properties of these estimators. In Section 4, we conduct two empirical applications, on dynamic capital structure and cash holdings. Section 5 concludes.

SYS-, LD-GMM, and LSDVC. This pooled estimation approach is, however, different from ours, which is to evaluate the relative performance of single estimation methods. Further, it mainly focuses on reducing the bias in the autoregressive coefficient (i.e., the SOA) and does not consider three recently developed bias-corrected estimators, namely BC, II, and DPF.

2. Dynamic panel data models in empirical corporate finance

2.1. Model specifications

Empirical studies in corporate finance often examine the dynamics of financial policy variables through the estimation of dynamic panel data models. In such models, the lagged dependent (financial) variables are included as regressors to capture firms' adjustments toward their optimal financial policies or the degree of persistence in the policies. In particular, in studies on capital structure, cash holdings, and dividends, many researchers employ a dynamic partial adjustment model to study firms' target adjustment behavior (refer to the studies cited in the Introduction for further details). Formally, this model is specified as:

$$\Delta y_{it} = \delta (y_{it}^* - y_{i,t-1}) + \eta_i + v_{it}, \quad (1)$$

where y_{it} and y_{it}^* denote the actual and optimal financial policies for firm i at time t , respectively. η_i is the (unobserved) time-invariant fixed effect, and v_{it} is the idiosyncratic error term.

In the dynamic model (1), a firm adjusts partially toward its optimal financial policy due to the presence of positive adjustment costs. The speed of adjustment (hereafter SOA), δ , should vary between 0 and 1, with a higher δ indicating a more rapid adjustment. Although the optimal financial policy, y_{it}^* , is not directly observed, it is typically modeled as a function of some firm-specific characteristics, i.e., $y_{it}^* = \theta' \mathbf{x}_{it}$, where \mathbf{x}_{it} denotes a $k \times 1$ vector of explanatory variables.

In estimating (1), most corporate finance studies adopt a one-stage procedure (e.g., Flannery and Rangan, 2006; Antoniou et al., 2008), which involves substituting $y_{it}^* = \theta' \mathbf{x}_{it}$ into (1) to yield the following model:

$$y_{it} = \gamma y_{i,t-1} + \beta' \mathbf{x}_{it} + \eta_i + v_{it}, \quad (2)$$

where $\gamma = 1 - \delta$ and $\beta = \delta \theta$. Researchers using the dynamic model (2) are able to simultaneously estimate the autoregressive coefficient, γ , and the long-run relationships between the optimal financial policy and the determining factors in equilibrium, θ (i.e., the long-run coefficients) (e.g., Antoniou et al., 2008). The majority of the empirical research estimating (2) is interested in the autoregressive coefficient, γ , or alternatively the adjustment speed, $\delta (= 1 - \gamma)$, because these coefficients shed light on the important questions of (i) how quickly firms adjust toward their optimal financial policies and (ii) how sticky the policies are.

2.2. Alternative estimators for dynamic panel data models

2.2.1. Least-squares and fixed-effects estimators

The pooled OLS (POLS) and fixed-effects (FE) estimators are not appropriate for estimating the dynamic model (2). Applying POLS to (2) produces biased and inconsistent estimates because the lagged dependent variable, $y_{i,t-1}$, is correlated with the firm fixed effect, η_i , i.e., $\mathbf{E}[y_{i,t-1}\eta_i] \neq 0$. It is well established that POLS tends to overestimate the autoregressive coefficient, γ , and underestimate the SOA (Baltagi, 2008). To deal with this bias, several studies turn to the FE estimator, which eliminates the fixed effects through the within-group transformation. Nevertheless, for a fixed T , FE is biased because the error term is correlated with the lagged dependent variable in the transformed equation. Specifically, FE underestimates the autoregressive coefficient, γ , and overestimates the SOA (Nickell, 1981). Most corporate studies that estimate dynamic models with FE are subject to this bias because they typically examine annual company data covering a short time period.

2.2.2. Instrumental variable (IV) and generalized method of moments (GMM) estimators

To deal with the finite-sample bias of the FE estimator, corporate finance researchers have turned to advanced methods developed in the panel data literature. The earliest of such methods is based on Anderson and Hsiao's (1981, 1982) instrumental variable approach (AH-IV), which removes the fixed effects via the first-difference transformation of (2):

$$\Delta y_{it} = \gamma \Delta y_{i,t-1} + \beta' \Delta x_{it} + \Delta v_{it}, \quad i = 1, \dots, N; t = 3, \dots, T. \quad (3)$$

This method employs $y_{i,t-2}$ (or $\Delta y_{i,t-2}$) as the only instrument for $\Delta y_{i,t-1}$ in (3). The main limitation of this just-identified estimator is a lack of efficiency due to the ignoring of additional valid instruments.

To address this lack of efficiency, researchers have considered using longer lagged dependent variables as additional instruments. This approach involves employing the generalized method of moments (GMM) to deal with the over-identification problem resulting from the use of those additional instruments. Arellano and Bond (1991) propose the first-difference GMM estimator (FD-GMM), which considers the moment conditions, $\mathbf{E}[y_{i,t-s}\Delta v_{it}] = 0$, with $t = 3, \dots, T$ and $s = 2, \dots, t - 1$, and uses a vector $(y_{i1}, \dots, y_{i,t-2})$ as the GMM instruments for $\Delta y_{i,t-1}$ in the first-differenced equations (3). Blundell and Bond (1998) further develop the

system GMM estimator (SYS-GMM), which improves the efficiency of FD-GMM by utilizing additional moment conditions in the level equations (2). It considers $(\Delta y_{i,2}, \dots, \Delta y_{i,t-1})$ as instruments for $y_{i,t-1}$ under the following moment conditions: $\mathbf{E}[\Delta y_{i,t-s} v_{it}] = 0$ for $t = 3, \dots, T$, and $s = 1, \dots, t - 2$ (Arellano and Bover, 1995; Blundell and Bond, 1998).

Hahn et al. (2007) propose the long-difference estimator (LD-GMM) to address the weak instruments problem typically associated with the above GMM estimators. This method relies on a smaller set of moment conditions and involves taking the difference between the first and last observations in (2). The long-differenced equation is then estimated by a two-stage least squares (2SLS) in which $(y_{i,T-1} - y_{i1})$ is instrumented by y_{i1} . Next, the residuals from this step are used as additional IVs in the next regression. A number of iterations (at least three) are often required to ensure that the parameters can be estimated efficiently (Hahn et al., 2007). Huang and Ritter (2009) propose a modified version of the LD-GMM estimator for unbalanced panels (LDP-GMM) by specifying a shorter differencing parameter, $k = 4$.

Note that the IV/GMM approaches are based on the assumption of no residual autocorrelation, i.e., $\mathbf{E}[v_{it} v_{is}] = 0 \forall t \neq s$. The GMM estimators further require valid, optimal instruments to be chosen. In theory, both of these conditions can be evaluated by using the AR(2) and Sargan tests (Arellano and Bond, 1991).⁶

2.2.3. Bias-corrected estimators

Recent research has followed another route to correct for the bias of FE. Kiviet (1995) develops a bias-corrected least-squares dummy variable estimator (LSDVC), which performs an analytical correction of the FE bias. Kiviet (1999), Bun and Kiviet (2003), Bruno (2005), and Bun and Carree (2006) extend this estimator to cases with heteroskedasticity and unbalanced panels.⁷ Although simulation studies show that LSDVC can outperform IV/GMM estimators for balanced panels of all lengths (Judson and Owen, 1999), this method has two potential drawbacks. First, it relies on a strong assumption that the true model is an AR(1) model. Second, this approach faces a severe computational challenge due to non-trivial matrix manip-

⁶Note, however, that the Sargan test is severely under-sized and has extremely low power in long panels (Bowsher, 2002).

⁷To approximate the finite-sample bias, an initial matrix of coefficient starting values must be specified (e.g., the FD-GMM's or SYS-GMM's initial estimates). See, for example, Bruno (2005).

ulations being required in the bias-correction formulas. To date, only a few studies in empirical corporate finance have used LSDVC (Öztekın and Flannery, 2012; Wintoki et al., 2012).

Everaert and Pozzi (2007) develop a bias-corrected FE estimator (BC) based on an iterative bootstrap that simulates the distribution of the FE estimator using the original (biased) FE estimates. This method then corrects the bias iteratively until unbiased estimates of the true parameters are found; see our Internet Appendix for details of the estimation algorithm. BC has a number of advantages over the above approaches to bias reduction. First, unlike the IV/GMM estimators, it does not require the specification of optimal, valid instruments, on which the properties of the IV/GMM estimators crucially depend. Second, compared to analytical approaches (e.g., LSDVC), BC performs a bias correction by means of non-parametric bootstrapping and as a result does not rely on restrictive parametric distributional assumptions. In practice, BC is computationally less demanding than LSDVC. Everaert and Pozzi (2007) demonstrate that BC outperforms GMM in terms of bias reduction and efficiency, while yielding better inferences than LSDVC in short panels. However, to the best of our knowledge, this method has not yet been applied in corporate finance research.

Recently, Gouriéroux et al. (2010) have proposed another approach to bias correction using the technique of indirect inference (II) (see also Smith, 1993; Gouriéroux et al., 1993, 2000).⁸ Similar to the bootstrap procedure above, this method also involves simulating data iteratively, estimating the model to approximate the bias function, and searching for unbiased estimates using the inverse function of the bias. One immediate advantage of II is that it does not rely on an explicit form for the bias function but calibrates it via simulation. Thus, similar to BC, this method is computationally less involved and more feasible than LSDVC, which requires explicit, and often complex, analytic derivations of the bias function or its expansion. Compared to BC, however, II is, at least in theory, a more general and versatile approach applicable to different model specifications and estimation methods, including both FE and IV/GMM estimators.⁹ Further, the II procedure can be implemented in conjunction with an appropriate

⁸We are extremely grateful to the Associate Editor and the reviewers for encouraging us to follow this line of inquiry.

⁹In principle, this technique can also be applied to non-linear dynamic panel data models proposed by Dang et al. (2012).

baseline estimator to take advantage of some of its useful properties. Gouriéroux et al. (2010), for example, propose to use an MLE method such as FE as the baseline estimator because it has a small variance, meaning that only a small number of simulated paths is required to ensure an accurate calibration of the bias function. The authors show via simulations that this II method is highly effective in the univariate autoregressive panel data model. We adopt the same approach in our simulation and empirical studies below; see our Internet Appendix for details of the estimation algorithm.¹⁰

3. Monte Carlo simulation studies

We now perform Monte Carlo simulation studies to investigate the finite-sample performance of the estimators. Our simulation design draws on previous simulations in the panel data literature (Kiviet, 1995; Judson and Owen, 1999; Bun and Kiviet, 2003; Bruno, 2005; Bun and Carree, 2006; Everaert and Pozzi, 2007), with several important extensions catered to corporate finance research. First, unlike those studies, we specify the key simulation parameters based on actual company data. Thus, while our simulations maintain the rigor of those studies, our findings and conclusions are directly applicable to empirical corporate finance.

Second, we explicitly account for unobservable heterogeneity in which an explanatory variable and the unobserved fixed effects are correlated. In empirical corporate finance, this problem is common: an explanatory variable determining the financial policy variable may proxy for both observed and unobserved factors, the latter being correlated with the fixed effects (Roberts and Whited, 2011). In dynamic cash models (Opler et al., 1999; Bates et al., 2009), for example, leverage, typically included as an explanatory variable, is potentially endogenous thanks to its likely correlation with the fixed effects. In studies on the relation between board structure and firm performance, the fixed effects affect firm performance but are also correlated with board structure, an explanatory variable in the performance equation (Wintoki et al., 2012).

¹⁰While theoretically II can be applied to the IV/GMM estimators, in practice, it is likely to be computationally demanding, especially in models with a large number of regressors. In our simulation analysis (unreported), we have implemented an II procedure using FD-GMM as the baseline estimator. We find that this approach does not outperform the II procedure using FE as the baseline estimator described above, while facing a significant computational burden.

Third, we also investigate the impact of (unaccommodated) residual serial correlation on the relative performance of the estimators. This analysis is important because both FH and our empirical analysis in Section 4 reveal that second-order autocorrelation is a potential problem in Compustat company data, one that can violate the validity of moment conditions underlying the IV/GMM estimators.

3.1. Data generating processes and simulation design

We consider the following data-generating process (hereafter, DGP):

$$y_{it} = \gamma y_{i,t-1} + \beta x_{it} + \eta_i + v_{it}, \quad i = 1, \dots, N; t = 2, \dots, T, \quad (4)$$

where both x_{it} and v_{it} follow an AR(1) process:

$$x_{it} = \rho x_{i,t-1} + \xi_{it}, \quad (5)$$

$$v_{it} = \phi v_{i,t-1} + \varepsilon_{it}, \quad (6)$$

where $\xi_{it} \sim N(0, \sigma_\xi^2)$, $\varepsilon_{it} \sim N(0, \sigma_\varepsilon^2)$, and they are distributed independently of each other.¹¹

Throughout our simulation experiments, we specify $\gamma = 0.8$, $\beta = 0.2$, $\rho = 0.5$, and $\sigma_v = 1$. Our choice of $\gamma = 0.8$ is guided by existing studies in both corporate finance and econometrics. Empirically, the choice of the parameter value $\gamma = 0.8$ (and the corresponding SOA, $\delta = 0.2$) is in line with the SOA estimated in recent corporate finance studies, especially those in the capital structure literature (Lemmon et al., 2008; Huang and Ritter, 2009). Previous Monte Carlo simulations typically consider the important case where the dependent variable is highly persistent, with $\gamma = 0.8$, i.e., where the finite-sample bias is more pronounced, thus greatly affecting the performance of the estimators (Kiviet, 1995; Judson and Owen, 1999; Bruno, 2005).¹² Next, we choose $\beta = 1 - \gamma$ so that a change in γ only affects the short-run dynamics and not the long-run relationship, which is normalized to unity, i.e., $\theta \equiv \beta / (1 - \gamma) = 1$. Our choice of $\rho = 0.5$ is reasonable given the range (0.085–0.844) of the AR(1) coefficient estimates for Compustat explanatory variables reported in Table 2 of FH.

¹¹Unlike FH's simulations, we do not consider the case with several regressors because we are more interested in the conditional distribution of the dependent variable than the unconditional joint distribution of the regressors. A minor advantage of considering multiple regressors is that it helps to account for multicollinearity among the regressors. However, provided that the underlying specification is correct, multicollinearity does not affect the bias in the estimates, but simply leads to larger standard errors.

¹²In unreported simulations, we examine another case where $\gamma = 0.2$ and find that, as γ decreases, the estimators (except POLS) generally perform better, which is in line with previous findings in the literature (e.g., Judson and Owen, 1999). Importantly, our conclusions regarding the performance of these methods remain unchanged.

As mentioned, our experiments explicitly allow for a correlation between the fixed effect, η_i , and the explanatory variable, x_{it} , i.e., $\mathbf{E}[x_{it}\eta_i] \neq 0$.¹³ To model this type of endogeneity, we generate η_i using $\eta_i = \mu(1-\gamma)\sigma_v\sigma_{z_i}^{-1}z_i$, where $z_i = \bar{x}_i - \bar{x}$, with $\bar{x}_i = T^{-1}\sum_{t=1}^T x_{it}$ being the within-group means, and $\bar{x} = (NT)^{-1}\sum_{i=1}^N\sum_{t=1}^T x_{it}$ being the overall mean.¹⁴ Since the variance of the fixed effect is $\sigma_\eta^2 = \mu^2(1-\gamma)^2\sigma_v^2$, the correlation between η_i and x_{it} is $Cov(\eta_i, x_{it}) = \mu(1-\gamma)\sigma_v\sigma_{z_i}^{-1}Cov(x_{it}, z_i)$ where the analytic forms of $\sigma_{z_i}^2$ and $Cov(x_{it}, z_i)$ are derived in our internet appendix. In the special case where we assume away any correlation between η_i and x_{it} , i.e., $\mathbf{E}[x_{it}\eta_i] = 0$, we simply draw η_i from $N(0, \sigma_\eta^2)$.

Following existing simulation studies (Kiviet, 1995; Judson and Owen, 1999; Bruno, 2005; Everaert and Pozzi, 2007), we examine the effects of two key control parameters, namely the loading factor, denoted μ , and the signal-to-noise ratio, denoted ϑ , on the finite-sample performance of the estimators. First, the loading factor measures the impact of the fixed effect, η_i , relative to that of the idiosyncratic error, v_{it} , on the dependent variable, y_{it} , as follows:

$$\mu = \frac{(1-\gamma)^{-1}\sigma_\eta}{\sigma_v}. \quad (7)$$

When the impact on y_{it} of the fixed effect, $(1-\gamma)^{-1}\sigma_\eta$, and that of the error term, σ_v^2 , have equal magnitudes, the loading factor is equal to unity, i.e., $\mu=1$. However, when the impact of the fixed effect is bigger than that of the disturbance, the loading factor increases, which could lead to greater estimation bias. It is established that the loading factor significantly affects the finite-sample performance of the estimators (e.g., Judson and Owen, 1999; Hayakawa, 2012).

The second key control parameter, the signal-to-noise ratio, measures the variance ratio of the signal contained in y_{it-1} and x_{it} (i.e., the explanatory power of these regressors), to the noise, v_{it} . The higher the signal-to-noise ratio, the more useful is the regressor in explaining y_{it} and the better will be the performance of the estimators (Judson and Owen, 1999; Bruno,

¹³Here we follow previous research and specify the fixed effect as a function of x_{it} (e.g., Chamberlain, 1982).

¹⁴Using a minimum-variance benchmark, Baker and Ruback (1999) show via simulations that the harmonic mean of a multiple (e.g., a financial performance ratio such as profitability) for an industry, or more generally, for a group of firms is more appropriate than alternative estimators, including the arithmetic mean (i.e., the simple mean used in our study). The latter estimator tends to put greater weights on firm-year observations with extreme values. Empirically, however, this concern is often negated because multiples are usually winsorized in order to reduce the potential effects of outliers.

2005). Formally, we follow Kiviet (1995) and rewrite (4) as:

$$y_{it} = \beta \varphi_{it} + \psi_{it} + \frac{\eta_i}{1 - \gamma}, \quad (8)$$

where y_{it} is decomposed into the weighted sum of two stationary AR(2) processes, $\varphi_{it} = (\gamma + \rho) \varphi_{i,t-1} - \gamma\rho \varphi_{i,t-2} + \xi_{it}$ and $\psi_{it} = (\gamma + \phi) \psi_{i,t-1} - \gamma\phi \psi_{i,t-2} + \varepsilon_{it}$, and the (long-run) unobserved fixed effect $\eta_i / (1 - \gamma)$, which are mutually independent of one another. Defining the latent variable, $w_{it} = \beta \varphi_{it} + \psi_{it} = y_{it} - \eta_i / (1 - \gamma)$, we are able to derive the signal-to-noise ratio as $\vartheta = \sigma_s^2 / \sigma_v^2$, where σ_s^2 is the variance of the signal $s_{it} = w_{it} - v_{it}$; see our internet appendix for detailed derivations and the exact form of σ_s^2 .¹⁵

In order to mimic actual company data, we set the values of the two main control parameters to $\mu = 3$ and $\vartheta = 6$, based on the empirical findings from the full- and sub-sample analyses of Compustat company data for the period from 1967 to 2006. Specifically, they are equal to the median values of the estimates from our rolling regressions of corporate leverage on a set of explanatory variables estimated using the IV/GMM and bias-corrected estimators.¹⁶

In sum, we consider five simulation experiments. Experiment 1 is our benchmark case with parameter values $\gamma = 0.8$, $\beta = 0.2$, $\rho = 0.5$, $\mu = 3$, $\vartheta = 6$, with $\mathbf{E}[x_{it} \eta_i] \neq 0$, and $\phi = 0$. Next, we examine the impact of varying the loading factor, to $\mu = 6$ or $\mu = 1$, in Experiment 2, and the impact of increasing the signal-to-noise ratio, to $\vartheta = 12$, in Experiment 3. Experiment 4 investigates a setting similar to Experiment 1, but assumes no correlation between the regressor and the fixed effect, i.e., $\mathbf{E}[x_{it} \eta_i] = 0$. Finally, Experiment 5 assesses the impact of serial correlation in the error term, v_{it} , with $\phi = -0.3$ or $\phi = 0.3$.¹⁷ In all experiments, we set $N = 400$ and use 1,000 replications. We consider the case with $T = 10$ as this length approximates the median and mean lengths of the (unbalanced) Compustat company data often used in empirical

¹⁵In our simulations, it is straightforward to control both parameters, ϑ and μ , i.e., we can control ϑ by adjusting σ_ξ^2 given σ_v^2 , and μ by adjusting σ_η given σ_v in (7).

¹⁶Following Flannery and Rangan (2006) and subsequent capital structure studies (Byoun, 2008; Faulkender et al., 2012), in those regressions we regress leverage on its lagged value and eight explanatory variables, namely profitability, market-to-book, depreciation, firm size, tangibility, research and development and its dummy variable, and the industry median of leverage. We consider 11 subsamples with the longest panel length of 30 years. See Section 4 for details. Based on the rolling regressions for those subsamples, we find that $\hat{\mu}$ varies between 2.5 and 3.7, while $\hat{\vartheta}$ varies between 3.6 and 8.

¹⁷Recently, Davidson and Monticini (2014) have proposed a (robust) dependent wild bootstrap procedure in the presence of heteroskedasticity and autocorrelation of unknown form. Given that existing studies have generally ignored the issue of time-varying heterogeneity, it would be useful to investigate this issue in future research.

corporate research and is in line with previous simulation studies (Judson and Owen, 1999).¹⁸ Finally, following Kiviet (1995), we generate the data carefully by deriving the initial conditions for x_{i0} and y_{i0} to avoid any slow convergence problems and a waste of random numbers; see our internet appendix for further details.¹⁹

3.2. Main simulation results

To evaluate the relative performance of each estimator, we use three conventional criteria, namely the bias, the standard error (SE), and the RMSE of the estimates. We are interested in the estimates of the autoregressive coefficient, γ , and the coefficient on the explanatory variable, β . For the IV/GMM estimators, we also report the rejection frequencies of the AR(2) and Sargan tests, enabling us to evaluate the validity of these estimators. We assess the accuracy of the estimate of the long-run target financial policy, $\hat{\theta}_{x_{it}}$, via the mean absolute error (MAE).

3.2.1. Experiment 1: Benchmark simulation

Table 1 reports the benchmark simulation results. First, with respect to γ , the POLS and FE estimators are most biased (by 0.097 and -0.126, respectively). In economic terms, this finding suggests that using POLS (FE) in research on target adjustment behavior potentially underestimates (overestimates) the adjustment speed, δ , by 10% (13%). Thus, POLS and FE are clearly not suitable for estimating the SOA in empirical corporate finance. Among the IV/GMM estimators, AH-IV and FD-GMM are both downwardly biased, by -0.074 and -0.064 respectively, while SYS-GMM is upwardly biased by 0.080; this bias is sizable, both statistically and economically. The LD- and LDP-GMM estimators perform better, with a moderate amount of bias (-0.021 and -0.016, respectively). However, the bias-corrected estimators, LSDVC, BC, and II, are the most favorable methods for estimating γ , with negligible bias (-0.007, 0.009, and 0.009, respectively). In terms of efficiency, most of the IV/GMM estimators (except for SYS-GMM)

¹⁸To examine the asymptotic performance of the estimators, we have also considered a longer time period, $T = 30$, which approximates the longest Compustat panel length. Unreported results suggest that, as T increases, estimation bias decreases and efficiency improves, consistent with theory.

¹⁹We perform our Monte Carlo simulations in Stata. For POLS and FE, we use the Stata procedures *reg* and *xtreg*, respectively. In terms of the IV/GMM estimators, we use Roodman's (2009b) Stata procedure *xtabond2*. Following Arellano and Bond (1991) and Blundell and Bond (1998), we include all lagged instruments in our GMM regressions; however, see Roodman (2009a) and Wintoki et al. (2012) for an alternative approach using fewer instruments. For the LD- and LDP-GMM methods, we implement a Stata code based on Huang and Ritter's (2009) approach. Regarding the bias-corrected estimators, we use Bruno's (2005) Stata procedure *xtlsdvc* for LSDVC and our own Stata codes for BC and II. To implement the DPF estimation, we follow Elsas and Florysiak (2014) and use *xttobit*. All our Stata codes are available upon request.

have large SEs.²⁰ In contrast, LSDVC, BC, and II are highly efficient and, moreover, have the smallest RMSEs. Taking these findings together, LSDVC, BC, and II outperform the other estimators in estimating γ and δ .

Second, in terms of β , most of the estimators, except for POLS and SYS-GMM, estimate this coefficient reasonably well. They all have a fairly small amount of bias, and low SEs and RMSEs. Third, turning to the estimates of the target financial variable, POLS, FE, and the IV/GMM estimators (except for LD- and LDP-GMM) are highly inaccurate: their MAEs vary between 0.91 and 3.72. In contrast, LSDVC, BC, and II are considerably more reliable, with the smallest MAEs (0.16, 0.28, and 0.26, respectively).²¹ Next, regarding the diagnostic tests, the Sargan test strongly rejects the null of valid instruments in the GMM estimations (except for LDP-GMM). In particular, it is frequently rejected (38.5%) in the SYS-GMM regressions, implying that, in practice, this method is likely to suffer from the problem of invalid instruments.

In sum, our benchmark simulation shows that POLS and FE are severely biased and inefficient and, thus, are inappropriate for estimating dynamic panel data models in corporate research. The IV/GMM estimators are highly unreliable while the bias-corrected estimators, LSDVC, BC, and II, are the most appropriate methods. In what follows, we examine whether the performance of the alternative estimators is robust to changes in the key control parameters. We focus our discussion on the IV/GMM estimators and bias-corrected estimators.

3.2.2. *Experiment 2: Changes in the relative magnitude of the fixed effect*

In Table 2, we examine how varying the magnitude of the fixed effect relative to that of the idiosyncratic error, measured by the loading factor μ , affects the performance of the estimators. First, we find that the IV/GMM estimators are highly sensitive to changes in μ , which is in line with the panel data literature (e.g., Kiviet, 1995; Judson and Owen, 1999) but is inconsistent with FH's remark that their simulation results are marginally affected by varying the loading factor. Specifically, in Panel A, when we increase μ from 3 to 6, the performance of both

²⁰FH use the RMSE to evaluate the performance of the estimators. However, using the RMSE as the sole criterion is likely to favor SYS-GMM among the IV/GMM estimators. For example, we find that SYS-GMM has a smaller RMSE than both AH-IV and FD-GMM, but is more biased than the latter two methods.

²¹Our results (untabulated) also show that LSDVC, BC, and II estimate the long-run coefficient with smallest amount of bias.

AH-IV and SYS-GMM, especially with respect to γ , deteriorates seriously: the amount of bias they exhibit, their SEs, and their RMSEs all increase considerably. Although FD-, LD-, and LD-GMM yield relatively precise estimates of γ and β , they still produce large SEs and RMSEs, especially with respect to γ . These estimators (except for FD-GMM) also produce their estimates of the target financial policy with rather large MAEs. In Panel B, we decrease μ from 3 to 1, assuming that the effects of the fixed effect and the error on y_{it} have equal magnitudes. Here, as expected, the IV/GMM estimators show some improvements. However, AH-IV, FD-, and SYS-GMM still estimate γ with a relatively large amount of bias (ranging between -0.045 and 0.031), although they perform reasonably well with respect to β . These methods seem to be outperformed by LD- and LDP-GMM, which estimate γ and β reasonably well. Notice however that FD-, SYS-, and LD-GMM are still likely to suffer from the problem of invalid instruments, as indicated by the high rejection frequencies of the Sargan test reported in both panels.

Second, the three bias-corrected estimators are robust to the above changes in the relative magnitude of the fixed effect. While BC and II are insensitive to the loading factor, LSDVC's performance changes little when $\mu = 6$ (Panel A), and is only slightly affected when $\mu = 1$ (Panel B). Moreover, all the three bias-corrected methods outperform the IV/GMM estimators with respect to γ and β , whether $\mu = 6$ or 1. Overall, these findings suggest that the bias-corrected estimators remain the most favorable methods, regardless of the relative magnitude of the fixed effect.

3.2.3. *Experiment 3: Changes in the relative explanatory power of the regressor*

In Table 3, we investigate the impact of varying the explanatory power of the regressor relative to that of the error term, as measured by the signal-to-noise ratio, ϑ , on the properties of the estimators. Specifically, we increase ϑ from 6 to 12 to consider a favorable setting where the regressor, x_{it} , has greater explanatory power than in the benchmark simulation. As expected, for most of the estimators, performance with respect to γ and β generally improves. However, although the IV/GMM estimators (except for AH-IV) estimate β more accurately and efficiently, they still show substantial bias in relation to γ and fairly large MAEs with respect to their target estimates. Further, as ϑ increases, the rejection frequencies of the Sargan

test in the GMM regressions become remarkably high, especially for SYS-GMM, even though its performance has generally improved. Overall, the most appropriate GMM estimators in this setting are LD- and LDP-GMM, although these methods are still outperformed by the bias-corrected methods, LSDVC, BC, and II. Compared to the benchmark experiment, the estimation accuracy and efficiency of LSDVC and BC improve slightly with higher signal-to-noise ratios. II shows the most remarkable improvement in its performance as it estimates both γ and β with negligible bias, and very small SE, RMSE, and MAE. In short, LSDVC, BC, and II are generally robust to changes in the relative explanatory power of the explanatory variable.

3.2.4. Experiment 4: No correlation between the regressor and the fixed effect

Our experiments have so far explicitly controlled for correlation between the explanatory variable and the unobserved fixed effect.²² In Table 4, we investigate a restrictive setting where x_{it} is independent of the fixed effect, $\mathbf{E}[x_{it}\eta_i] = 0$. We first find that the IV/GMM estimators improve markedly, producing more accurate and efficient estimates of γ and β . The AH-IV, SYS-, and LDP-GMM estimators perform very well, with SYS-GMM emerging as the most appropriate method. The latter estimator estimates both γ and β with negligible bias, and very small SEs and RMSEs. It also produces an estimate of the target financial policy with a small MAE. The finding that SYS-GMM is the most favorable method under this setting is consistent with FH’s main simulation results.

Regarding the bias-corrected estimators, we find that the properties of BC and II stay almost the same as in the benchmark simulation. Importantly, these two methods still perform very well with respect to γ and β , and are comparable to SYS-GMM, the most favorable method in the present experiment. However, the performance of LSDVC with respect to γ deteriorates slightly, although LSDVC and II exhibit the smallest MAE.

²²In further analysis (unreported), we also consider the case with two-way endogeneity, due to the correlation of the explanatory variable, \mathbf{x}_{it} , with both the fixed effect, η_i , and the idiosyncratic error term, v_{it} . This scenario, in theory, favors the GMM estimators, which can alleviate the second source of endogeneity via the use of lagged instruments. However, we find that the GMM estimators do not perform well, despite their advanced treatment of the endogeneity due to $\mathbf{E}[\mathbf{x}_{it}v_{it}] \neq \mathbf{0}$. The bias-corrected estimators, especially LSDVC, remain generally robust, although their performance deteriorates compared with the benchmark experiment. For empirical applications, we thus recommend using the lagged values of the explanatory variables ($\mathbf{x}_{i,t-1}$) to avoid this endogeneity problem.

3.2.5. Experiment 5: Residual serial correlation

In Table 5, we consider the impact of residual serial correlation on the performance of the estimators. We set the AR(1) coefficient in (6) to be $\phi = 0.3$ (i.e., positive autocorrelation) or $\phi = -0.3$ (i.e., negative autocorrelation).²³ As expected, the performance of the IV/GMM estimators deteriorates considerably when their critical assumption of no serial correlation is violated. These methods now estimate γ with a very large amount of bias, varying between 0.068 and 0.551 (0.062 and -0.582) in the presence of positive (negative) serial correlation. The IV/GMM estimators, especially AH-IV and FD-GMM, also tend to become much less efficient, as they produce relatively large SEs and RMSEs. Further, their target estimates contain substantial estimation errors (MAEs). As expected, the rejection frequencies of the AR(2) and of the Sargan tests reach almost 100% in the IV/GMM regressions, confirming that these methods are invalid in the presence of residual serial correlation. Note, however, that most of the IV/GMM estimators (except for AH-IV and FD-GMM) still estimate β reasonably well, although this coefficient is estimated more accurately by the FE estimator.

Turning to the bias-corrected estimators, BC's performance when $\phi = -0.3$ is, surprisingly, marginally better than in the benchmark case. This method becomes slightly worse when $\phi = 0.3$ but still outperforms the remaining methods: its bias in relation to γ is slightly on the high side (0.035) but still acceptable. However, as expected, LSDVC deteriorates in the presence of autocorrelation, a violation of the assumption underlying this estimator. Its bias regarding γ is non-negligible (0.070 and -0.068), while its MAE with respect to the target estimate is rather large, especially when $\phi = 0.3$. The performance of II is heavily influenced by autocorrelation simply because the simulated sample paths in the II procedure do not reflect this feature of the DGP. Notably, this method now estimates γ with a large amount of bias (0.099 and -0.062) and produces a highly inflated MAE (2.93), especially when $\phi = 0.3$.

In sum, our results show that among the bias-corrected estimators, BC is the only method robust to the presence of residual serial correlation. On the other hand, the IV/GMM estimators become less reliable for dynamic panel data models with autocorrelation.²⁴

²³These values are chosen according to our subsample analysis, which is discussed in Section 4; they are significantly higher than the trivial AR(2) coefficients of 0.1 and 0.05 considered by FH.

²⁴A solution to residual serial correlation involves extending the ARDL(1,0) model (2) to an ARDL(2,1) model.

3.3. Further extensions: Fractional dependent variables

In empirical corporate research, the financial policy variable of interest (e.g., leverage) can be a ratio bounded by 0 and 1. Loudermilk (2007) argues that the fractional property of the dependent variable may lead to estimation bias in dynamic panels. The author develops an augmented Tobit estimator, in which the latent dependent variable can take values outside the unit interval but is doubly censored at the corner outcomes of 0 and 1. Moreover, based on Wooldridge (2005), Loudermilk (2007) models the distribution of the unobserved fixed effects, η_i , conditional on the initial value, y_{i1} , and the exogenous explanatory variable, \bar{x}_i . This estimation is effectively a combination of the Tobit estimation and the correlated random-effects (maximum likelihood) estimation (hereafter CRE), with two additional regressors, y_{i1} and \bar{x}_i . Elsas and Florysiak (2014, hereafter EF) further extend Loudermilk's (2007) method to the case of unbalanced panels and name this estimator DPF.

We now compare the performance of the IV/GMM and bias-corrected estimators with that of the DPF estimator in an additional Monte Carlo simulation where the dependent variable is generated as a ratio bounded by 0 and 1. We design our simulation as follows. First, based on EF's simulation, we account for the fractional nature of the dependent variable, y_{it} , by generating a latent variable, $y_{it}^\#$:

$$y_{it}^\# = \gamma y_{i,t-1} + \beta x_{it} + \eta_i + v_{it}, \quad i = 1, \dots, N; \quad t = 2, \dots, T, \quad (9)$$

$$\eta_i = \alpha_0 + \alpha_1 y_{i1} + \alpha_2 \bar{x}_i + \alpha_i, \quad (10)$$

$$y_{i1}^\# = \beta x_{i1} + \alpha_0 + \alpha_2 \bar{x}_i + \alpha_i + v_{i1}, \quad (11)$$

where $x_{it} \sim U(x_L, x_U)$, $\bar{x}_i = T^{-1} \sum_{t=1}^T x_{it}$, $\alpha_i \sim N(0, \sigma_\alpha^2)$, $v_{it} \sim N(0, \sigma_v^2)$, and α_i , x_{it} , and v_{it} are independently generated. We then doubly censor $y_{it}^\#$ at the corner outcomes, 0 and 1, to obtain y_{it} . That is, we set $y_{it} = 0$ if $y_{it}^\# \leq 0$, $y_{it} = y_{it}^\#$ if $0 < y_{it}^\# < 1$, and $y_{it} = 1$ if $y_{it}^\# \geq 1$. Following Loudermilk (2007) and EF, we also include within the fixed effect, η_i , two additional terms, y_{i1} and \bar{x}_i . In this DGP, we set $(N, T) = (400, 10)$, $\gamma = 0.8$, and $\beta = 0.2$, as in the main simulation studies in Sections 3.1 and 3.2. Next, we select the parameter values as follows:

Our analysis (unreported) shows that BC is also appropriate for the ARDL(2,1) model.

$x_L = -0.5$, $x_U = 1$, $\alpha_0 = 0$, $\alpha_1 = 0.1$, $\alpha_2 = -0.25$, $\sigma_v = 0.05$, and $\sigma_\alpha = (0.05, 0.10)$. We vary σ_α in order to investigate how changing the magnitude of the fixed effect relative to that of the idiosyncratic error affects the performance of the estimators. By varying σ_α , we can also control for the percentage of firm-year observations that lie outside the unit interval and are hence to be doubly censored. We find that the percentage censored is 9% when $\sigma_\alpha = 0.05$, and 21% when $\sigma_\alpha = 0.10$.

The DGP in (9) assumes a dynamic structure whereby the (true, but unobserved) dependent variable, $y_{it}^\#$, depends recursively on the past realizations of $y_{i,t-1}$, rather than those of $y_{i,t-1}^\#$. However, this process has two potential shortcomings. First, it may not be consistent with the dynamics of the financial policy variable. Consider an example where the fractional dependent variable is leverage. In this case, the latent dependent variable, $y_{it}^\#$, is often interpreted as the firm's target leverage ratio or debt capacity, which can take values outside the unit interval (see EF). However, to the extent that $y_{it}^\#$ proxies for target leverage, it should depend on its past realizations, as modeled by the partial adjustment framework (2). Second, by assuming that $y_{it}^\#$ depends on y_{it-1} , not $y_{i,t-1}^\#$, this DGP favors the Tobit method and guarantees that DPF, an augmented Tobit estimator, will perform well. To address these two potential limitations of the DGP in (9), we consider an alternative DGP whereby the latent dependent variable, $y_{it}^\#$, is generated by the following partial adjustment model (see also Hu, 2002):

$$y_{it}^\# = \gamma y_{i,t-1}^\# + \beta x_{it} + \eta_i + v_{it}, \quad i = 1, \dots, N; \quad t = 2, \dots, T. \quad (12)$$

The difference between (9) and (12) is that the latter includes $y_{i,t-1}^\#$ rather than $y_{i,t-1}$ in the updating equations. To generate the data, we first update $y_{it}^\#$ using (12) and then collect $\left\{ \left\{ y_{it}^\# \right\}_{t=1}^T \right\}_{i=1}^N$, to which we apply double censoring to obtain $\left\{ \left\{ y_{it} \right\}_{t=1}^T \right\}_{i=1}^N$.

In Table 6, we report the simulation results for the two DGPs, (9) in Panels A and B, and (12) in Panels C and D. First, we find from Panel A that DPF estimates both γ and β highly accurately and efficiently, with the least amount of bias and the smallest SEs.²⁵ Nevertheless, several other estimators also perform reasonably well. AH-IV and LD-GMM produce their

²⁵We provide more comprehensive simulation results than EF, who only consider the relative performance of POLS, FE, SYS-GMM, LD-GMM, LSDVC, and DPF, and report the simulation results for γ (see their Table 2).

estimates of γ and, to a lesser extent, β with negligible bias, while FD-GMM, LSDVC, BC, and II yield a small to moderate amount of bias, and moderate SEs and RMSEs. However, in Panel B, when we increase the percentage of censoring, the results become more favorable toward DPF. This estimator now outperforms most of the other methods, except for LD-GMM, which still estimates γ reasonably well.

Next, the results for the second DGP, (12), reported in Panel C, show that DPF is no longer the most favorable method for estimating γ and β . It estimates γ with non-negligible bias, in fact even more than that produced by AH-IV, LDP-GMM, and LSDVC. Moving to Panel D, as expected, most of the estimators exhibit deteriorating performance as the percentage of censoring increases (33% compared to 15% in Panel C). Surprisingly, although DPF performs well in this setting, it does not dominate the other estimators. It estimates γ less accurately and efficiently than LD-GMM and BC. However, DPF still produces the most accurate estimate of β and the target long-run relationship with the smallest MAE. Our results are inconsistent with EF's finding that DPF outperforms the other estimators in a simulation based on a similar censoring scheme (in which the authors first generate the complete set of panel data before introducing censoring; see their footnote 19). Our simulations show that DPF is sensitive to the DGP used and that it does not always dominate the other estimators.

In sum, we find that DPF is the most robust estimator in the presence of severe censoring at the corner outcomes and a relatively high magnitude of the fixed effect.²⁶ However, if censoring is not a serious issue and the relative magnitude of the fixed effect is moderate, the other estimators, such as LD-GMM, LDP-GMM, LSDVC and, in some cases, BC, can also produce fairly accurate results, with respect to both γ and β . In capital structure research, for example, recent studies show that about 9% of firms have zero leverage (Strebulaev and Yang, 2013).²⁷ With this mild level of censoring, our simulations suggest that the average effects for fractional dependent variables in linear models can still be estimated with a small amount of bias; see also Papke and Wooldridge (2008).

²⁶The strong and robust performance of DPF is not simply due to its unique double-censoring mechanism, but is also due to the inclusion of the additional regressors, y_{i1} and \bar{x}_i , which provide a near-perfect solution to the unobserved heterogeneity and endogeneity (Chamberlain, 1984; Wooldridge, 2005, 2010).

²⁷In our sample covering the period between 1967 and 2006, we find that 9,387 out of a total of 103,556 firm-year observations (also about 9%) have zero leverage, while no observations have leverage above 1.

4. Empirical applications

In this section, we evaluate the performance of the estimators in an empirical application in corporate finance. We examine one of the most interesting, yet contentious, questions in empirical capital structure research, i.e., how quickly firms adjust toward their target leverage ratios. To this end, we estimate the following dynamic model:

$$y_{it} = \gamma y_{i,t-1} + \beta' \mathbf{x}_{i,t-1} + \eta_i + \tau_t + v_{it}, \quad (13)$$

which is similar to the baseline dynamic specification (2). The model further includes year dummies, τ_t , except in the LD-GMM (cross-sectional) regression, and the LDP-GMM and LSDVC regressions, which do not allow for time effects.²⁸ We measure the dependent variable, y_{it} , by book leverage.²⁹ In terms of the explanatory variables, we follow Flannery and Rangan (2006) and include the first lagged values of profitability (EBIT), market-to-book (MB), depreciation (DEP), firm size (TA), tangibility (FA), research and development (RD) and its dummy variable (RDD), and the industry median of leverage (IND). The use of lagged values addresses a potential endogeneity problem in which the contemporaneous values of the explanatory variables may be correlated with the idiosyncratic error term, i.e., $\mathbf{E}[x_{it}v_{it}] \neq 0$.

We collect annual accounting data for US firms from the Compustat database, for the period 1967–2006. Our sample of panel data is unbalanced with the mean and median lengths of 15 and 14 years, respectively. Following standard practice, we exclude financial firms (SIC codes 6000-6999) and utility firms (SIC codes 4900-4999). We winsorize all independent variables at the 1st and 99th percentiles as in previous studies (e.g., Flannery and Rangan, 2006). Table 7 provides variable definitions and summary statistics. To preserve space, we will not discuss these statistics here because they are generally consistent with those reported in earlier studies on capital structure (e.g., Flannery and Rangan, 2006).

²⁸To sidestep the issue with year dummies, in a robustness check, we demean the variables and then apply the LDP-GMM and LSDVC estimators. The results (untabulated) are qualitatively similar.

²⁹Several studies use book leverage as the main measure of leverage (Fama and French, 2002; Chang and Dasgupta, 2009). We also consider market leverage and obtain qualitatively similar results for the SOA.

4.1. Speed of Adjustment

Table 8 reports the regression results for the dynamic partial adjustment model of leverage (13). First, the results clearly show that the SOA varies considerably, depending on the estimation method employed. The POLS and FE estimates are 15% and 40%, respectively, consistent with previous evidence in the literature (e.g., Lemmon et al., 2008). There is a large difference between these estimates because POLS has downward bias, while FE has upward bias, which is in line with our simulation results. Turning to the IV/GMM methods, the LDP-GMM estimate of the SOA is 7%, which seems inappropriate because it does not fall into the range of the POLS and FE estimates (15%–40%).³⁰ The LDP-GMM estimate is sensitive to the choice of the long-difference parameter, k : when $k = 8$, the SOA is estimated at 15.5%. The estimates of the SOA obtained using AH-IV, FD-, SYS-, and LD-GMM vary between 15% and 18%, close to the (biased) POLS estimate. On closer inspection of the Sargan and AR(2) test results, these estimates are unreliable because their fundamental assumptions of valid instruments and no autocorrelation are violated in most cases. Specifically, there is evidence of negative autocorrelation in the AH-IV and FD-GMM regressions, which, as shown in our simulation experiments, has a serious effect on the performance of these estimators. In short, the IV/GMM estimators seem to produce unreliable estimates of the SOA in capital structure research.

The bias-corrected estimators, LSDVC and BC, obtain very similar adjustment speeds, of 26% and 28%, respectively. II produces a slightly lower estimate of 24%. DPF, the only estimator that explicitly accounts for the fractional nature of leverage, yields an estimate of 27%. Our earlier simulation results suggest that, among the methods reviewed in this paper, LSDVC, BC, and II are the most appropriate estimators in our overall simulation experiments, while DPF is the most robust method in the special case with a fractional dependent variable. However, if the percentage of censoring is small, as is indeed the case in this application (see footnote 27 above), the performance of LSDVC, BC, and II is also acceptable.³¹ Our empirical

³⁰We follow a general consensus in the literature (e.g., Flannery and Rangan, 2006; Huang and Ritter, 2009) and conjecture that an unbiased estimate should lie between the POLS and FE estimates, which are biased downwardly and upwardly, respectively. Note, however, that Zhou et al. (2014) argue that such conjecture may no longer hold under certain conditions involving model misspecifications in which the direction of the POLS bias becomes ambiguous, although the FE estimate continues to provide an upper bound on the SOA.

³¹EF further contend that DPF can distinguish mechanical mean reversion (hereafter MMR) from real target leverage adjustment; see Chang and Dasgupta (2009) for a discussion of MMR. They argue that when the data

results are thus in line with these simulation findings because all three methods yield similar estimates of the SOA.

Given that the SOAs estimated by LSDVC, BC, II, and DPF vary in a tight range between 24% and 28%, we conjecture that these estimates are closest to the true SOA for our Compustat sample. An adjustment speed in the range of 24%-28% implies half-lives of approximately 2.2 years, which, compared to the 14-year median life of firms in our sample, provides evidence of moderate adjustment.³² In other words, firms seem to adjust toward their target leverage ratios at moderate rates. Overall, our analysis demonstrates the importance of using appropriate methods in empirical capital structure research. For example, the POLS, AH-IV, LD-, and FD-GMM estimates of 15% (with half-lives of 4.3 years) would indicate sluggish adjustment, while the FD estimate of 40% (with a half-life of 1.36 years) would suggest very rapid adjustment. Researchers using these unreliable results would run the risk of reaching incorrect conclusions about target adjustment behavior.

4.2. Coefficients on the explanatory variables

We next study whether the alternative estimators can reasonably estimate the coefficients on the determinants of leverage. We examine whether the signs of the coefficients are consistent with the theoretical predictions and empirical evidence in the literature. As suggested by Frank and Goyal (2009), we focus on the five most important determinants of leverage, namely profitability (EBIT), the market-to-book ratio (MB), firm size (TA), tangibility (FA), and the industry median of leverage (IND). Following previous capital structure research (Titman and Wessels, 1988; Rajan and Zingales, 1995; Flannery and Rangan, 2006; Byoun, 2008; Lemmon et al., 2008), we expect leverage to be inversely related to profitability and the market-to-book ratio, but positively associated with firm size, tangibility, and the industry median of leverage; see also Parsons and Titman (2009) for a review. Specifically, profitable firms are likely to have low leverage, as predicted by the pecking order theory (Myers and Majluf, 1984; Myers,

are simulated under random financing such that there is no target adjustment behavior, the SOA estimated by DPF is equal to zero, while those estimated by the other methods are positive. In our internet appendix, we conduct another simulation study on leverage processes generated under random financing. However, we do not find that DPF outperforms the other estimators.

³²The “half life” represents the number of years it takes a company to move halfway toward its target leverage after a leverage shock. Formally, it is calculated as $\ln 0.5 / \ln(1 - \delta)$.

1984) and by dynamic trade-off models (e.g., Leary and Roberts, 2005; Strebulaev, 2007).³³ Within a broad trade-off framework, agency theories predict a negative relation between the market-to-book ratio and leverage because high- (low-) growth firms use less (more) leverage to mitigate under- (over-) investment incentives (Myers, 1977; Jensen and Meckling, 1976). Firm size and leverage are positively related because large firms have low bankruptcy, agency, and transaction costs, hence strong incentives to utilize debt. Tangibility has a positive impact on leverage because tangible assets can be used as collateral to mitigate the asset substitution problem, thus reducing the agency costs of debt (Jensen and Meckling, 1976). The industry median of leverage is often used as the benchmark leverage ratio, or as a control for common, but otherwise omitted, factors affecting all firms within the industry (Hovakimian et al., 2001); this variable is expected to have a positive impact on leverage (Flannery and Rangan, 2006).

The results in Table 8 show that the coefficients on the main explanatory variables are not well estimated by the IV/GMM methods. In the AH-IV and FD-GMM regressions, EBIT is significantly positive and MB is insignificant at the 5% level, while TA, FA, and IND are all significantly negative. None of these results are consistent with either the theoretical predictions or the existing evidence in the literature. In the SYS-GMM regression, although FA is significantly positive, EBIT is significant with an unexpected sign, while MB and TA are both insignificant. In the LD-GMM model, EBIT is significantly negative, as expected, but MB, TA, FA, and IND are all insignificant. Finally, in the LDP-GMM regression, MB is insignificant, while EBIT, FA, and IND are all significant with the unexpected signs. In sum, our analysis suggests that studies using the IV/GMM methods may obtain unreasonable estimates of target leverage and, consequently, support target adjustment behavior when there is no meaningful target estimate.

The coefficient estimates obtained using LSDVC, BC, II, and DPF are in line with expectations. Specifically, the two variables EBIT and MB are always significantly negative, while TA, FA, and IND are significantly positive. The magnitudes of these coefficients are also comparable across the three models. Overall, LSDVC, BC, II, and DPF provide the most plausible

³³The static trade-off framework suggests that profitable firms should lever up in order to take advantage of the debt tax shield. However, most cross-sectional empirical evidence documents a negative effect of profitability on leverage. See Xu (2012) for an exception.

results in terms of the statistical significance and signs of the coefficients on the main explanatory variables. The results obtained with these methods suggest that firms adjust with moderate speeds toward well-defined target leverage ratios.

4.3. *Additional tests*

As a first robustness check, we conduct a subsample analysis to evaluate the performance of each estimator in a time-varying context. We examine the results obtained from 11 rolling leverage regressions, each for a 30-year (unbalanced) subsample, with the first subsample covering the period 1967–1996, and the last one covering the period 1977–2006.³⁴ Our (untabulated) results show that the IV/GMM methods are not only inappropriate for the whole sample, but are also unreliable for the time-varying subsamples. Their estimates of the SOA of leverage continue to be biased, while their estimates of the other coefficients often have the wrong sign and exhibit a large amount of variation. Meanwhile, LSDVC, BC, II, and DPF are again the most robust estimators as they produce reasonable estimates of all the coefficients that are in line with our prior expectations.

Since leverage is fractional, our findings may not be relevant to other areas in corporate research where the dependent variable is not a ratio bounded by the unit interval. To address this concern, we also conduct a second empirical study on dynamic cash holdings. Following Foley et al. (2007) and Bates, Kahle, and Stulz (2009, hereafter BKS), we measure cash holdings as the natural logarithm of the cash-to-net-assets ratio. We follow BKS and include as the explanatory variables industry cash flow volatility (IND_CFASD), market-to-book (MB), firm size (TA), cash flow (CFA), net working capital (NWC), capital expenditure (CAPEX), R&D expenditure (RDS), book leverage (BDR), acquisition expenditure (MA), and the dividend payout dummy (DIV). Our analysis, reported in Table 9, shows that LSDVC, BC, II, yield reasonable, similar estimates of the SOA of cash, at 48-49%, which are close to the midpoint between the POLS and FE estimates.³⁵ The adjustment speeds obtained using AH-IV (54%), FD-GMM (53%), and SYS-GMM (52%) are relatively closer to the biased FE estimate (63%).

³⁴The mean and median lengths of these (unbalanced) panels are 8 and 6 years, respectively.

³⁵An estimate of this size indicates a half-life of 1.1 years, suggesting that firms adjust quite rapidly toward their optimal cash holdings. While this SOA is much greater than those typically estimated in capital structure research, it is in line with recent evidence on dynamic cash adjustment (e.g., Dittmar and Duchin, 2011).

The bias in the IV/GMM results may be due to the presence of invalid instruments and (positive) residual serial correlation, as indicated by the strong rejection of the Sargan and AR(2) tests in most of these regressions. Finally, using the bias-corrected estimators, LSDVC, BC, and II, we also obtain estimates of coefficients on the explanatory variables that are generally in line with theoretical expectations.

5. Conclusions

Many studies in empirical corporate finance have employed dynamic panel data models to examine the dynamic behavior of corporate financial policy variables. However, a major difficulty in using these models is determining how to obtain consistent and efficient estimates, especially in short panels of company data, in the likely presence of (1) unobserved heterogeneity and endogeneity, (2) residual serial correlation, or (3) fractional dependent variables. Which estimators should researchers use in these contexts?

To address this important research question, we investigate two classes of advanced econometric techniques for dynamic panel data models, including (1) the instrumental variable approach, the first-difference, system, and long-difference GMM, and (2) the bias-corrected estimators, based on an analytical approach, an iterative bootstrap procedure, or a simulation-based indirect inference technique. Further, we consider an augmented Tobit estimator that accounts for the fractional nature of the dependent variable. We conduct Monte Carlo simulation experiments and present empirical applications, to capital structure and cash holdings, to examine the relative performance of the estimators.

Overall, we find that the bias-corrected estimators, LSDVC, BC, and II, are generally most appropriate and robust for estimating dynamic panel data models in empirical corporate research. These methods can estimate the autoregressive coefficient and the coefficients on the explanatory variables with the most accuracy and efficiency. In our simulations, they are also robust to changes in the key control parameters, including the relative magnitude of the fixed effects and the relative explanatory power of the regressors. Of these three methods, BC performs relatively well even in regressions with autocorrelation, and in models with higher lag orders, such as the ARDL(2,1) model; it is also computationally less demanding than both LSDVC and II. In the specific case where the dependent variable is a ratio, and censored at 0 and

1, LSDVC, BC, and II may still provide reasonable estimates with a moderate amount of bias, although at a high level of censoring researchers should consider using DPF. In our applications using capital structure and cash holdings, we find that LSDVC, BC, and II produce empirically the most plausible estimates of the speeds of dynamic leverage and cash adjustments, and reasonable estimates of the coefficients on the explanatory variables, which are consistent with prior theoretical predictions.

However, the IV/GMM estimators generally perform poorly in our simulation and empirical studies. Their estimates of the autoregressive coefficient tend to be biased and unreliable. Moreover, these methods are sensitive to the presence of unobserved heterogeneity and, in particular, serially correlated errors. We identify a few conditions (e.g., optimal instruments and no autocorrelation) under which these methods can be useful, although in our empirical applications these conditions prove to be restrictive and unrealistic. Consequently, the regression results obtained using these methods are unreliable.

In sum, our study suggests that the IV/GMM estimators should be used with extreme care, and that future empirical research in corporate finance should consider employing the highly robust bias-corrected estimators, especially BC. We also find that while II is a useful general and versatile approach to bias correction, it is computationally challenging, especially in the multidimensional case where further research on the iterative algorithm for II is warranted.

We conclude with some limitations and avenues for future research. First, while our simulations have systematically examined the most important issues in the estimation of dynamic panel data models, we cannot account for all of the specific settings and minor issues encountered in the vast empirical research literature. Further, while our simulation results should generalize to many areas of corporate finance, our empirical applications are restricted to two topics, namely capital structure and cash holdings. Hence, it would be useful for future research to verify our simulation and empirical findings in other areas of corporate finance. Second, our study does not account for a possibility that the (individual) firm fixed effects are time-varying or, more generally, that the error components are cross-sectionally correlated. Recent research shows that existing estimation procedures are significantly affected by cross-sectional dependence (e.g., Pesaran, 2006; Ahn et al., 2007; Bai, 2009). Thus, it would be important for future

studies to examine the properties of alternative methods for estimating dynamic panel data models with cross-sectionally correlated errors. Third, in empirical research, firms with different characteristics may have different response coefficients and adjust towards their long-term optimal financial policies at different rates (e.g., Byoun, 2008; Dang et al., 2012; Faulkender et al., 2012). Hence, it would be more realistic to allow for heterogeneous slope parameters (e.g., Pesaran and Smith, 1995; Pesaran et al., 1999). Finally, while our study adopts a “horse-race” approach to compare the properties of alternative estimators, it would be useful for future studies to evaluate the performance of a pooled estimation approach that combines different (bias-corrected) estimators, as proposed by Zhou et al. (2014).

Acknowledgement

We would like to thank Carol Alexander and Ike Mathur (Editors), an Associate Editor, three anonymous reviewers, seminar participants at University of Glasgow, University of Leeds, Yonsei University, and the Bank of Korea, and conference participants at the CFE Conference at University of London, December 2010, the Trans-Atlantic Doctoral Conference at London Business School, April 2011, and the FMA Annual Meeting, Colorado, October 2011, and especially Richard Baillie, Michael Binder, Charlie Cai, Mara Faccio, David Florysiak, Ian Garrett, Rongbing Huang, Daniel Hung, Krishna Paudyal, Kevin Reilly, Jay Ritter, Myunghwan Seo, Seungju Song, and Chris Veld for their helpful comments and suggestions on previous versions of this paper. Partial financial support from the Economic Social Research Council (ESRC) (grant number: RES-000-22-3161) is gratefully acknowledged. The usual disclaimer applies.

References

- Ahn, S.C., Y.H. Lee, and P. Schmidt. 2007. Stochastic frontier models with multiple time-varying individual effects. *Journal of Productivity Analysis* 27: 1-12.
- Aivazian, V.A., Y. Ge, and J. Qiu. 2005a. The impact of leverage on firm investment: Canadian evidence. *Journal of Corporate Finance* 11: 277-91.
- Aivazian, V.A., Y. Ge, and J. Qiu. 2005b. Debt maturity structure and firm investment. *Financial Management* 34: 107-19.
- Anderson, T.W., and C. Hsiao. 1981. Estimation of dynamic models with error components. *Journal of the American Statistical Association* 76: 598-606.
- Anderson, T.W., and C. Hsiao. 1982. Formulation and estimation of dynamic models using panel data. *Journal of Econometrics* 18: 47-82.
- Andres, C., A. Betzer, M. Goergen, and L. Renneboog. 2009. Dividend policy of German firms: A panel data analysis of partial adjustment models. *Journal of Empirical Finance* 16: 175-187.
- Antoniou, A., Y. Guney, and K. Paudyal. 2006. The determinants of corporate debt maturity structure: Evidence from France, Germany and the UK. *European Financial Management* 12: 161-94.
- Antoniou, A., Y. Guney, and K. Paudyal. 2008. The determinants of capital structure: Capital market oriented versus bank oriented institutions. *Journal of Financial and Quantitative Analysis* 43: 59-92.
- Arellano, M., and S.R. Bond. 1991. Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations. *Review of Economic Studies* 58: 227-297.
- Arellano, M., and O. Bover. 1995. Another look at the instrumental variable estimation of error components models. *Journal of Econometrics* 68: 29-51.
- Bai, J. 2009. Panel data models with interactive fixed effects. *Econometrica* 77: 1229-1279.
- Baker, M., and R.S. Ruback. 1999. Estimating Industry Multiples. Working paper. Harvard University.
- Baltagi, B. H., 2008. *Econometric Analysis of Panel Data*. West Sussex: John Wiley and Sons.
- Bates, T., K. Kahle, and R. Stulz. 2009. Why do U.S. firms hold so much more cash than they used to? *Journal of Finance* 64: 1985-2021.

- Blundell, R.W., and S.R. Bond. 1998. Initial conditions and moment restrictions in dynamic panel data models. *Journal of Econometrics* 87: 115-143.
- Bowsher, C. G. 2002. On testing overidentifying restrictions in dynamic panel data models. *Economics Letters* 77: 211-220.
- Brav, A., J. Graham, C. Harvey, and R. Michaely. 2005. Payout policy in the 21st century. *Journal of Financial Economics* 77: 483-527.
- Bruno, S.F. 2005. Approximating the bias of the LSDVC estimator for dynamic unbalanced panel data model. *Economics Letters* 87: 361-366.
- Bun, M.J.G., and M.A. Carree. 2006. Bias-corrected estimation in dynamic panel data models. *Journal of Business and Economic Statistics* 23: 200-210.
- Bun, M.J.G., and J.F. Kiviet. 2003. On the diminishing returns of higher order terms in asymptotic expansions of bias. *Economics Letters* 79: 145-152.
- Byoun, S. 2008. How and when do firms adjust their capital structures toward targets? *Journal of Finance* 63: 3069-3096.
- Chamberlain, G. 1982 Multivariate regression models for panel data, *Journal of Econometrics* 18: 5-46.
- Chamberlain, G. 1984. Panel data. In *Handbook of Econometrics*. Vol. 2., edited by Z. Griliches and M. Intriligator. Amsterdam: North-Holland.
- Chang, X., and S. Dasgupta. 2009. Target behavior and financing: How conclusive is the evidence? *Journal of Finance* 64: 1767-1796.
- Chen, L., and Z. Zhao. 2007. Mechanical mean reversion of leverage ratios. *Economic Letters* 95: 223-229.
- Dang, V.A., M. Kim, and Y. Shin. 2012. Asymmetric capital structure adjustments: New evidence from dynamic panel threshold models. *Journal of Empirical Finance* 19: 465-482.
- Davidson, R. and A. Monticini. 2014. Heteroskedasticity-and-Autocorrelation-consistent bootstrapping. Mimeo, McGill University.
- de Miguel, A., and J. Pindado. 2001. Determinants of capital structure: New evidence from Spanish panel data. *Journal of Corporate Finance* 7: 77-99.
- Dittmar, A., and R. Duchin. 2011. The dynamics of cash. Working Paper. University of Michigan.

- Elsas, R., and D. Florysiak. 2014. Dynamic capital structure adjustment and the impact of fractional dependent variables. *Journal of Financial and Quantitative Analysis*. In-press.
- Everaert, G., and L. Pozzi. 2007. Bootstrap-based bias correction for dynamic panels. *Journal of Economic Dynamics & Control* 31: 1160-1184.
- Fama, E.F., and K.R. French. 2002. Testing trade-off and pecking order predictions about dividends and debt. *Review of Financial Studies* 15: 1-33.
- Faulkender, M.W., M.J. Flannery, K.W. Hankins, and J.M. Smith. 2012. Cash flows and leverage adjustments. *Journal of Financial Economics* 103: 632-646.
- Flannery, M.J., and K.W. Hankins. 2013. Estimating dynamic panel models in corporate finance. *Journal of Corporate Finance* 19: 1–19.
- Flannery, M.J., and K.P. Rangan. 2006. Partial adjustment toward target capital structures. *Journal of Financial Economics* 79: 469-506.
- Foley, C.F., Z. Hartzell, S. Titman, and G.J. Twite. 2007. Why do firms hold so much cash? *Journal of Financial Economics* 86: 579-607.
- Frank, M.Z., and V.K. Goyal. 2007. Trade-off and pecking order theories of debt. In *Handbook of Corporate Finance: Empirical Corporate Finance*, Vol. 1, edited by B.E. Eckbo. Amsterdam: North Holland.
- Frank, M., and V.K. Goyal. 2009. Capital structure decisions: Which factors are reliably important? *Financial Management* 38: 1-37.
- Gouriéroux, C., A. Monfort, and E. Renault. 1993. Indirect Inference. *Journal of Applied Econometrics* 8: S85-S118.
- Gouriéroux, C., P.C.B. Phillips, and J. Yu. 2010. Indirect Inference for Dynamic Panel Models. *Journal of Econometrics* 157: 68-77.
- Gouriéroux, C., E. Renault, and N. Touzi. 2000. Calibration by simulation for small sample bias correction. In Mariano, R.S., T. Schuermann, and M Weeks. (Eds.), *Simulation-Based Inference in Econometrics: Methods and Applications*. Cambridge University Press: 328-358.
- Hahn, J., J. Hausman, and G. Kuersteiner. 2007. Long difference instrumental variables estimation for dynamic panel models with fixed effects. *Journal of Econometrics* 140: 574-617.
- Hayakawa, K. 2012. The asymptotic properties of the system GMM estimator in dynamic panel data models when both N and T are large. Mimeo. Hiroshima University.

- Hovakimian, A., T. Opler, and S. Titman. 2001. The debt-equity choice. *Journal of Financial and Quantitative Analysis* 36: 1-24.
- Hu, L. 2002. Estimation of a censored dynamic panel data model. *Econometrica* 70: 2499-2517.
- Huang, R., and J.R. Ritter. 2009. Testing theories of capital structure and estimating the speed of adjustment. *Journal of Financial and Quantitative Analysis* 44: 237-271.
- Jalilvand, A., and R.S. Harris. 1984. Corporate behavior in adjusting to capital structure and dividend targets: an econometric study. *Journal of Finance* 39: 127-145.
- Jensen, M.C., and W.H. Meckling. 1976. Theory of the firm: managerial behavior, agency costs and ownership structure. *Journal of Financial Economics* 3: 305-360.
- Judson, R.A., and A.L. Owen. 1999. Estimating dynamic panel data models: a guide for macro-economists. *Economic Letters* 65: 9-15.
- Kayhan, A., and S. Titman, 2007. Firms' histories and their capital structures. *Journal of Financial Economics* 83: 1-32.
- Kiviet, J.F. 1995. On bias, inconsistency, and efficiency of various estimators in dynamic panel data models. *Journal of Econometrics* 68: 53-78.
- Kiviet, J.F. 1999. Expectation of expansions for estimators in a dynamic panel data model: some results for weakly exogenous regressors. In *Analysis of Panel Data and Limited Dependent Variables*, edited by C. Hsiao, K. Lahiri, L.-F. Lee, and M.H. Pesaran. Cambridge: Cambridge University Press.
- Leary, M., and M. Roberts. 2005. Do firms rebalance their capital structure?. *Journal of Finance* 60: 2575-2619.
- Lemmon, M.L., M.R. Roberts, and J.F. Zender. 2008. Back to the beginning: persistence and the cross-section of corporate capital structure. *Journal of Finance* 63: 1575-1608.
- Lintner, J. 1956. Distribution of incomes of corporations among dividends, retained earnings, and taxes. *American Economic Review* 46: 97-113.
- 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: 462-472.
- Myers, S.C. 1977. Determinants of corporate borrowing. *Journal of Financial Economics* 5: 145-175.

- Myers, S.C. 1984. The capital structure puzzle. *Journal of Finance* 34: 575-592.
- Myers, S.C., and N.S. Majluf. 1984. Corporate financing and investments decisions when firms have information that investors do not have. *Journal of Financial Economics* 13: 187-221.
- Nickell, S.J. 1981. Biases in dynamic models with fixed effects. *Econometrica* 49: 1417-1426.
- Opler, T., L. Pinkowitz, R. Stulz, and R. Williamson. 1999. The determinants and implications of corporate cash holdings. *Journal of Financial Economics* 52: 3-46.
- Ozkan, A. 2000. An empirical study of corporate debt maturity structure. *European Financial Management* 6: 197-212.
- Ozkan, A. 2001. Determinants of capital structure and adjustment to long run target: evidence from UK company panel data. *Journal of Business Finance and Accounting* 28: 175-198.
- Ozkan, A., and N. Ozkan. 2004. Corporate cash holdings: an empirical investigation of UK companies. *Journal of Banking and Finance* 28: 2103-2134.
- Öztekin, Ö., and M. Flannery. 2012. Institutional determinants of capital structure adjustment speeds. *Journal of Financial Economics* 103: 88-112.
- Papke, L.E., and J.M. Wooldridge. 2008. Panel data methods for fractional response variables with an application to test pass rates. *Journal of Econometrics* 145: 121-133.
- Parsons, C., and S. Titman. 2009. Empirical capital structure: A review. *Foundations and Trends in Finance* 3: 1-93.
- Pesaran, M.H. 2006. Estimation and inference in large heterogeneous panels with a multifactor error structure. *Econometrica* 74: 967-1012.
- Pesaran, M.H., and R.P. Smith. 1995. Estimating long-run relationships from dynamic heterogeneous panels. *Journal of Econometrics* 68: 79-113.
- Pesaran, M.H., Y. Shin, and R.P. Smith. 1999. Pooled mean group estimation of dynamic heterogeneous panels. *Journal of American Statistical Association* 94: 621-634.
- Rajan, R.G., and L. Zingales. 1995. What do we know about capital structure? Some evidence from international data. *Journal of Finance* 50: 1421-1461.
- Roberts, M.R., and T.M. Whited. 2011. Endogeneity in empirical corporate finance. In *Handbook of the Economics of Finance*, Vol. 2., edited by G. Constantinides, M. Harris, and R. Stulz. Amsterdam: North Holland.
- Roodman, D. 2009a. A note on the theme of too many instruments. *Oxford Bulletin of Economics and Statistics* 71: 135-158.

- Roodman, D. 2009b. How to do xtabond2: an introduction to difference and system GMM in Stata. *Stata Journal* 9: 86-136.
- Skinner, D. 2008. The evolving relation between earnings, dividends, and stock repurchases. *Journal of Financial Economics* 87: 582-609.
- Strebulaev, I.A. 2007. Do tests of capital structure theory mean what they say? *Journal of Finance* 62: 1747-1787.
- Strebulaev, I.A., and B. Yang. 2013. The mystery of zero-leverage firms. *Journal of Financial Economics* 109: 1-23.
- Smith, A.A. 1993. Estimating nonlinear time-series models using simulated vector autoregressions. *Journal of Applied Econometrics* 8: S63-S84.
- Titman, S., and R. Wessels. 1988. The determinants of capital structure. *Journal of Finance* 43: 1-19
- Warr, R.S., W.B. Elliott, J. Koëter-Kant, and Ö. Öztekin. 2012. Equity mispricing and leverage adjustment costs. *Journal of Financial and Quantitative Analysis* 47: 589-616.
- Welch, I. 2011. Two common problems in capital structure research: the financial-debt-to-asset ratio and issuing activity versus leverage changes. *International Review of Finance* 11: 1-17.
- Wintoki, M.B., J.S. Linck, and J. Netter. 2012. Endogeneity and the dynamics of internal corporate governance. *Journal of Financial Economics* 105: 581-606.
- Wooldridge, J.M. 2005. Fixed-effects and related estimators for correlated random-coefficient and treatment-effect panel data models. *Review of Economics and Statistics* 87: 385-390.
- Wooldridge, J.M. 2010, Correlated random effects models with unbalanced panels. Mimeo, Michigan State University.
- Xu, J. 2012. Profitability and capital structure: Evidence from import penetration. *Journal of Financial Economics* 106: 427-446.
- Zhou, Q., R. Faff, and K. Alpert. 2014. Bias correction in the estimation of dynamic panel models in corporate finance. *Journal of Corporate Finance* 25: 494-513.

Table 1. Performance of alternative estimators: Benchmark simulation experiment

Estimators	POLS	FE	AH-IV	FD-GMM	SYS-GMM	LD-GMM	LDP-GMM	LSDVC	BC	II
Bias (γ)	0.097	-0.126	-0.074	-0.064	0.080	-0.021	-0.016	-0.007	0.009	0.009
SE (γ)	0.004	0.011	0.097	0.065	0.010	0.044	0.026	0.011	0.016	0.013
RMSE (γ)	0.097	0.127	0.122	0.091	0.080	0.049	0.031	0.013	0.019	0.016
Bias (β)	0.022	0.001	-0.006	-0.006	0.015	0.001	0.000	0.002	0.000	0.003
SE (β)	0.005	0.005	0.009	0.008	0.006	0.014	0.006	0.005	0.005	0.005
RMSE (β)	0.023	0.005	0.011	0.010	0.016	0.014	0.006	0.005	0.005	0.005
AR(2)			3.8%	4.3%	4.4%					
Sargan				10.4%	38.5%	12.1%	5.8%			
MAE	3.72	1.24	1.92	0.91	2.57	0.60	0.33	0.16	0.28	0.26

Notes: POLS and FE refer to the pooled OLS and fixed-effects methods. AH-IV is the instrumental variable estimator. FD- and SYS-GMM are the first-difference and system GMM estimators. LD- and LDP-GMM are long-difference estimators; the former uses the longest differencing parameter, while the latter sets the differencing parameter at 4. LSDVC, BC, and II are three bias-corrected (fixed-effects) estimators, based on an analytical approach, an iterative bootstrap, and a simulation-based procedure using MLE as the baseline estimator, respectively. See our Internet Appendix for detailed estimation algorithms for BC and II. The DGP is given by (4)-(6) with the following simulation parameters: $\gamma = 0.8$, $\beta = 0.2$, $\rho = 0.5$, $\vartheta = 6$, $\mu = 3$, $\mathbf{E}[x_{it}\eta_i] \neq 0$, $\phi = 0$, and $T = 10$. The performance of each estimator is evaluated by bias (Bias), standard error (SE), and root mean square error (RMSE). AR(2) denotes the rejection frequencies of the AR(2) test, a test for the null of no residual serial correlation for AH-IV, FD-, and SYS-GMM. Sargan denotes the rejection frequencies of the Sargan test, a test for the joint null of validity of the moment conditions (over-identifying restrictions), for FD-, SYS-, LD-, and LDP-GMM. Both tests are evaluated at the 5% significance level. MAE is the mean absolute error, which measures the accuracy of the estimate of the (long-run) target financial policy.

Table 2. The impact of changes in the magnitude of the fixed effect

Estimators	POLS	FE	AH-IV	FD-GMM	SYS-GMM	LD-GMM	LDP-GMM	LSDVC	BC	II
Panel A: High loading factor ($\mu = 6$)										
Bias (γ)	0.141	-0.126	-0.503	-0.025	0.115	0.008	0.003	0.002	0.009	0.009
SE (γ)	0.002	0.011	10.482	0.071	0.008	0.098	0.089	0.011	0.016	0.013
RMSE (γ)	0.141	0.127	10.494	0.075	0.115	0.098	0.089	0.011	0.019	0.016
Bias (β)	0.020	0.001	-0.041	-0.003	0.022	-0.003	-0.001	0.004	0.000	0.003
SE (β)	0.005	0.005	0.890	0.008	0.007	0.019	0.006	0.005	0.005	0.005
RMSE (β)	0.020	0.005	0.891	0.009	0.023	0.019	0.006	0.006	0.005	0.005
AR(2)			3.5%	4.6%	4.4%					
Sargan				9.8%	45.3%	17.0%	11.2%			
MAE	8.79	1.24	4.40	0.98	5.23	2.26	1.88	0.17	0.28	0.26
Panel B: Low loading factor ($\mu = 1$)										
Bias (γ)	0.033	-0.126	-0.040	-0.045	0.031	-0.028	-0.020	-0.016	0.009	0.009
SE (γ)	0.006	0.011	0.047	0.032	0.013	0.040	0.021	0.012	0.016	0.013
RMSE (γ)	0.033	0.127	0.062	0.055	0.034	0.049	0.029	0.020	0.019	0.016
Bias (β)	0.015	0.001	-0.003	-0.004	0.007	0.002	0.000	0.001	0.000	0.003
SE (β)	0.004	0.005	0.007	0.007	0.006	0.014	0.006	0.005	0.005	0.005
RMSE (β)	0.016	0.005	0.008	0.008	0.009	0.014	0.006	0.005	0.005	0.005
AR(2)			4.0%	4.2%	4.0%					
Sargan				9.7%	27.9%	12.2%	4.8%			
MAE	0.93	1.24	0.70	0.63	0.74	0.94	0.34	0.24	0.28	0.26

Notes: The DGP is given by (4)-(6) with the following simulation parameters: $\gamma = 0.8$, $\beta = 0.2$, $\rho = 0.5$, $\vartheta = 6$, $\mathbf{E}[x_{it}\eta_i] \neq 0$, $\phi = 0$, and $T = 10$. The loading factor, μ , is set at $\mu = 6$, and then $\mu = 1$. See notes to Table 1 for abbreviations.

Table 3. The impact of changes in the explanatory power of the regressor

Estimators	POLS	FE	AH-IV	FD-GMM	SYS-GMM	LD-GMM	LDP-GMM	LSDVC	BC	II
Bias (γ)	0.077	-0.081	-0.095	-0.063	0.074	-0.037	-0.022	-0.008	-0.005	0.000
SE (γ)	0.003	0.009	0.066	0.045	0.008	0.022	0.014	0.009	0.010	0.010
RMSE (γ)	0.077	0.081	0.116	0.077	0.075	0.043	0.026	0.011	0.011	0.010
Bias (β)	0.016	0.000	-0.008	-0.006	0.009	0.003	0.000	0.001	-0.001	0.001
SE (β)	0.003	0.003	0.006	0.005	0.004	0.009	0.004	0.003	0.003	0.003
RMSE (β)	0.017	0.003	0.010	0.008	0.010	0.009	0.004	0.003	0.003	0.003
AR(2)			5.9%	6.2%	5.5%					
Sargan				16.1%	72%	24%	13.8%			
MAE	3.77	1.43	1.75	1.27	3.33	0.72	0.50	0.21	0.23	0.21

Notes: The DGP is given by (4)-(6) with the following simulation parameters: $\gamma = 0.8$, $\beta = 0.2$, $\rho = 0.5$, $\mu = 3$, $\mathbf{E}[x_{it}\eta_i] \neq 0$, $\phi = 0$, and $T = 10$. The signal-to-noise ratio, ϑ , is set at $\vartheta = 12$. See notes to Table 1 for abbreviations.

Table 4. The impact of no correlation between the regressor and the fixed effect

Estimators	POLS	FE	AH-IV	FD-GMM	SYS-GMM	LD-GMM	LDP-GMM	LSDVC	BC	II
Bias (γ)	0.113	-0.126	-0.026	-0.033	0.000	-0.040	-0.022	-0.017	0.009	0.010
SE (γ)	0.005	0.012	0.045	0.032	0.021	0.033	0.022	0.012	0.016	0.013
RMSE (γ)	0.113	0.126	0.052	0.046	0.021	0.052	0.031	0.021	0.019	0.016
Bias (β)	-0.021	0.001	-0.002	-0.003	-0.002	0.004	0.000	0.000	0.000	0.002
SE (β)	0.005	0.005	0.007	0.007	0.006	0.013	0.006	0.005	0.005	0.005
RMSE (β)	0.021	0.005	0.007	0.007	0.006	0.014	0.006	0.005	0.005	0.005
AR(2)			5.4%	5.6%	5.3%					
Sargan				6.5%	9%	9.1%	6%			
MAE	3.44	1.23	0.61	0.53	0.27	0.53	0.36	0.26	0.28	0.26

Notes: The DGP is given by (4)-(6), with the following simulation parameters: $\gamma = 0.8$, $\beta = 0.2$, $\rho = 0.5$, $\mu = 3$, $\vartheta = 6$, $\phi = 0$, and $T = 10$. Contrary to the benchmark simulation, it is assumed that there is no correlation between the fixed effect and the regressor, i.e., $\mathbf{E}[x_{it}\eta_i] = 0$. See notes to Table 1 for abbreviations.

Table 5. Performance of alternative estimators in the presence of residual serial correlation

Estimators	POLS	FE	AH-IV	FD-GMM	SYS-GMM	LD-GMM	LDP-GMM	LSDVC	BC	II
Panel A: Positive Residual Serial Correlation ($\phi = 0.3$)										
Bias (γ)	0.115	-0.060	0.551	-0.099	0.127	0.068	0.075	0.070	0.035	0.099
SE (γ)	0.004	0.011	0.177	0.081	0.010	0.054	0.026	0.012	0.023	0.012
RMSE (γ)	0.115	0.061	0.579	0.127	0.127	0.087	0.080	0.071	0.042	0.099
Bias (β)	0.018	0.000	0.045	0.002	0.014	-0.011	-0.002	0.004	0.000	0.009
SE (β)	0.006	0.006	0.015	0.011	0.007	0.017	0.008	0.006	0.006	0.006
RMSE (β)	0.019	0.006	0.047	0.011	0.016	0.020	0.008	0.007	0.006	0.011
AR(2)			95.8%	99.4%	99.4%					
Sargan				100%	100%	99%	100%			
MAE	4.22	0.62	5.57	0.84	5.35	2.66	2.27	1.56	0.65	2.93
Panel B: Negative Residual Serial Correlation ($\phi = -0.3$)										
Bias (γ)	0.087	-0.184	-0.582	-0.474	0.062	-0.096	-0.128	-0.068	0.004	-0.062
SE (γ)	0.003	0.012	0.117	0.164	0.013	0.031	0.017	0.012	0.013	0.013
RMSE (γ)	0.087	0.184	0.593	0.501	0.063	0.101	0.129	0.069	0.014	0.063
Bias (β)	0.025	0.001	-0.047	-0.046	0.004	0.011	0.002	0.002	-0.001	0.000
SE (β)	0.004	0.004	0.010	0.013	0.006	0.013	0.005	0.004	0.004	0.004
RMSE (β)	0.025	0.004	0.048	0.048	0.007	0.017	0.005	0.004	0.004	0.004
AR(2)			92.5%	94.9%	100%					
Sargan				100%	100%	98%	100%			
MAE	3.39	1.63	2.75	2.59	1.66	0.98	1.31	0.84	0.20	0.80

Notes: The DGP is given by (4)-(6) with the following simulation parameters: $\gamma = 0.8$, $\beta = 0.2$, $\rho = 0.5$, $\mu = 3$, $\vartheta = 6$, $\mathbf{E}[x_{it}\eta_i] \neq 0$, and $T = 10$. Two cases are considered: positive ($\phi = 0.3$) and negative ($\phi = -0.3$) residual serial correlation. See notes to Table 1 for abbreviations.

Table 6. The impact of fractional dependent variables

Estimators	POLS	FE	AH-IV	FD-GMM	SYS-GMM	LD-GMM	LDP-GMM	LSDVC	BC	DPF	II
Panel A: Scheme 1 with $\sigma_\alpha = 0.05$ (censoring percentage of 9%)											
Bias (γ)	0.099	-0.116	0.001	-0.024	0.062	0.007	-0.020	-0.021	0.039	0.003	0.046
SE (γ)	0.008	0.010	0.080	0.023	0.010	0.009	0.022	0.019	0.030	0.012	0.012
RMSE (γ)	0.099	0.116	0.080	0.033	0.063	0.012	0.029	0.028	0.049	0.012	0.048
Bias (β)	-0.031	-0.020	-0.011	-0.016	-0.018	-0.011	-0.021	-0.015	-0.011	0.001	-0.009
SE (β)	0.004	0.004	0.009	0.005	0.005	0.004	0.011	0.005	0.005	0.004	0.004
RMSE (β)	0.032	0.021	0.015	0.017	0.019	0.012	0.023	0.016	0.012	0.004	0.010
AR(2)			6.7%	7.0%	7.5%						
Sargan				5.4%	58.4%	4.1%	5.3%				
MAE	0.28	0.18	0.32	0.07	0.14	0.02	0.08	0.07	0.13	0.02	0.10
Panel B: Scheme 1 with $\sigma_\alpha = 0.1$ (censoring percentage of 21%)											
Bias (γ)	0.160	-0.133	-0.082	-0.083	0.131	-0.006	-0.059	-0.067	-0.045	0.003	0.038
SE (γ)	0.006	0.010	0.042	0.019	0.008	0.009	0.017	0.013	0.021	0.011	0.013
RMSE (γ)	0.160	0.133	0.092	0.085	0.131	0.011	0.061	0.068	0.050	0.011	0.040
Bias (β)	-0.049	-0.043	-0.040	-0.043	-0.033	-0.034	-0.043	-0.041	-0.039	0.001	-0.033
SE (β)	0.005	0.004	0.006	0.005	0.005	0.004	0.011	0.006	0.005	0.004	0.004
RMSE (β)	0.049	0.043	0.040	0.043	0.034	0.035	0.045	0.041	0.039	0.005	0.033
AR(2)			4.2%	4.4%	4.6%						
Sargan				6.1%	100.0%	9.5%	15.3%				
MAE	1.17	0.22	0.17	0.18	0.60	0.08	0.16	0.17	0.14	0.02	0.03

to be continued next page

Table 6. The impact of fractional dependent variables (continued)

Estimators	POLS	FE	AH-IV	FD-GMM	SYS-GMM	LD-GMM	LDP-GMM	LSDVC	BC	DPF	II
Panel C: Scheme 2 with $\sigma_\alpha = 0.05$ (censoring percentage of 15%)											
Bias (γ)	0.129	-0.100	-0.007	-0.022	0.095	0.030	-0.007	-0.008	0.084	0.022	0.072
SE (γ)	0.008	0.010	0.072	0.023	0.010	0.009	0.021	0.017	0.039	0.012	0.013
RMSE (γ)	0.130	0.100	0.072	0.032	0.096	0.032	0.022	0.019	0.092	0.026	0.073
Bias (β)	-0.041	-0.032	-0.027	-0.031	-0.030	-0.023	-0.031	-0.027	-0.020	-0.005	-0.021
SE (β)	0.005	0.004	0.009	0.006	0.005	0.004	0.011	0.006	0.006	0.004	0.004
RMSE (β)	0.042	0.033	0.028	0.032	0.031	0.023	0.033	0.027	0.021	0.007	0.021
AR(2)			4.0%	4.2%	4.9%						
Sargan				6.3%	87.5%	6.4%	6.9%				
MAE	0.53	0.18	0.25	0.10	0.27	0.03	0.08	0.07	0.40	0.05	0.17
Panel D: Scheme 2 with $\sigma_\alpha = 0.1$ (censoring percentage of 33%)											
Bias (γ)	0.182	-0.116	-0.077	-0.078	0.166	0.018	-0.045	-0.052	-0.017	0.031	0.074
SE (γ)	0.005	0.010	0.038	0.020	0.008	0.009	0.016	0.014	0.020	0.012	0.012
RMSE (γ)	0.182	0.116	0.086	0.080	0.166	0.020	0.047	0.054	0.026	0.033	0.075
Bias (β)	-0.066	-0.063	-0.063	-0.067	-0.057	-0.055	-0.061	-0.060	-0.059	-0.009	-0.052
SE (β)	0.005	0.005	0.006	0.006	0.007	0.005	0.011	0.006	0.005	0.005	0.005
RMSE (β)	0.067	0.063	0.063	0.067	0.057	0.055	0.062	0.060	0.059	0.010	0.053
AR(2)			5.8%	5.8%	7.2%						
Sargan				6.0%	100.0%	13.1%	16.6%				
MAE	3.06	0.24	0.21	0.22	1.44	0.08	0.18	0.18	0.14	0.06	0.08

Notes: In Scheme 1 the DGP is given by (9). In Scheme 2 the DGP is given by (12). In both, the simulation parameters are set as follows: $\gamma = 0.8$, $\beta = 0.2$, $x_L = -0.5$, $x_U = 1.0$, $\alpha_0 = 0.1$, $\alpha_1 = 0.1$, $\alpha_2 = -0.25$, and $\sigma_u = 0.1$. We consider the short panel ($T = 10$) with 400 hypothetical companies ($N = 400$). We control the percentage of (doubly) censored observations outside the unit interval by varying σ_α . In Scheme 1 the censoring percentages are 9% and 21%, respectively, for $\sigma_\alpha = 0.05$ and 0.1. In Scheme 2 the censoring percentages are 15% and 33%, respectively, for $\sigma_\alpha = 0.05$ and 0.1. See notes to Table 1 for abbreviations.

Table 7. Variable definitions and summary statistics

Variable	N	Mean	Median	Stdev.	Min.	Max.
BDR	108,139	0.2549	0.2310	0.2043	0.0000	0.9387
DEP	108,139	0.0486	0.0403	0.0358	0.0023	0.2254
EBIT	108,139	0.0239	0.0851	0.2497	-1.3312	0.3574
FA	108,139	0.3252	0.2773	0.2260	0.0086	0.9074
IND	108,139	0.2277	0.2395	0.0945	0.0000	0.6557
MB	108,139	1.5395	1.0472	1.5145	0.3076	9.9618
RD	108,139	0.0363	0.0000	0.0826	0.0000	0.5090
RDD	108,139	0.5640	1.0000	0.4959	0.0000	1.0000
SIZE	108,139	13.7966	13.7312	2.0382	9.4890	18.7517

Notes: This table reports the number of observations (N), mean, median, standard deviation (Stdev.), minimum (Min.), and maximum (Max.) of the variables considered in the empirical application. We follow Flannery and Rangan (2006) and define the explanatory variables as follows. The dependent variable, book leverage (BDR), is measured by the ratio of total debt, including debt of both long-term and short-term maturities, to total assets. Non-debt tax shields (DEP) is the ratio of depreciation to total assets. Profitability (EBIT) is the ratio of Earnings Before Interest and Taxes to total assets. Asset tangibility (FA) is the ratio of property, plant, and equipment (PPE) to total assets. The industry median book leverage (IND) is calculated based on Fama and French's (2002) 49 industry groupings. Growth opportunities (MB) or the market-to-book ratio is the ratio of total liabilities plus the market value of equity to total assets. R&D expenses (RD) is the ratio of R&D expenses to total assets. The R&D dummy variable (RDD) is equal to one if the firm did not report R&D expenses and zero otherwise. Firm size (SIZE) is the natural log of total assets, measured in 1983 dollars, where CPI is the consumer price index (1983=100).

Table 8. Summary of regression results for the dynamic partial adjustment model of leverage

Variable	POLS	FE	AH-IV	FD-GMM	SYS-GMM	LD-GMM	LDP-GMM	LSDVC	BC	DPF	II
L.BDR	0.848*** (0.003)	0.604*** (0.006)	0.846*** (0.020)	0.836*** (0.018)	0.821*** (0.011)	0.855*** (0.050)	0.929*** (0.010)	0.736*** (0.003)	0.718*** (0.003)	0.727*** (0.004)	0.757*** (0.003)
EBIT	-0.027*** (0.004)	-0.035*** (0.005)	0.119*** (0.009)	0.108*** (0.009)	0.032*** (0.005)	-0.269** (0.109)	0.012** (0.006)	-0.029*** (0.002)	-0.030*** (0.003)	-0.016*** (0.003)	-0.032*** (0.003)
MB	0.000 (0.000)	-0.002*** (0.001)	-0.002* (0.001)	-0.001 (0.001)	0.000 (0.001)	0.013 (0.010)	0.000 (0.001)	-0.001** (0.000)	-0.001*** (0.000)	-0.001*** (0.000)	-0.002*** (0.000)
DEP	-0.105*** (0.017)	-0.152*** (0.032)	-0.074 (0.064)	-0.107* (0.061)	-0.036 (0.027)	-0.123 (0.469)	-0.268*** (0.038)	-0.185*** (0.018)	-0.176*** (0.019)	-0.169*** (0.021)	-0.166*** (0.020)
TA	0.000 (0.000)	0.012*** (0.001)	-0.045*** (0.004)	-0.042*** (0.004)	0.000 (0.000)	-0.002 (0.005)	0.007*** (0.001)	0.009*** (0.001)	0.009*** (0.001)	0.009*** (0.001)	0.010*** (0.001)
FA	0.019*** (0.002)	0.034*** (0.007)	-0.157*** (0.017)	-0.138*** (0.017)	0.013*** (0.004)	0.003 (0.046)	-0.015* (0.009)	0.027*** (0.005)	0.025*** (0.004)	0.030*** (0.005)	0.030*** (0.004)
RDD	0.004*** (0.001)	-0.004* (0.002)	-0.009** (0.004)	-0.010** (0.004)	0.007*** (0.001)	-0.001 (0.013)	0.002 (0.002)	0.000 (0.002)	-0.003* (0.002)	-0.001 (0.002)	-0.003* (0.002)
RD	-0.033*** (0.011)	-0.022 (0.020)	0.105*** (0.034)	0.085*** (0.033)	0.016 (0.015)	0.029 (0.246)	0.076*** (0.026)	-0.012 (0.013)	-0.018 (0.011)	0.007 (0.013)	-0.020* (0.012)
IND	0.080*** (0.006)	0.088*** (0.014)	-0.223*** (0.022)	-0.202*** (0.022)	0.067*** (0.009)	0.119 (0.100)	-0.032** (0.012)	0.094*** (0.007)	0.071*** (0.010)	0.083*** (0.010)	0.074*** (0.011)
SOA	15%	40%	15%	16%	18%	15%	7%	26%	28%	27%	24%
AR(2) test			-2.21 [0.03]	-2.09 [0.04]	-1.59 [0.11]						
Sargan test				240.72 [0.00]	394.12 [0.00]	42.7 [0.24]	138.6 [0.00]				
Year dummies	Yes	Yes	Yes	Yes	Yes	No	No	No	Yes	No	Yes
Observations	92,517	92,517	80,149	80,149	86,870	164	39,435	92,517	92,517	92,517	92,517

Notes: The speed of adjustment (SOA) is derived from the coefficient on lagged leverage. Variable definitions are listed in Table 7. AR(2) is a test for the null of no residual serial correlation and Sargan is a test for the joint null of validity of the moment conditions (over-identifying restrictions). *, **, and *** indicate significance of the coefficients at the 1%, 5%, and 10% levels, respectively. Figures in (·) are the standard errors and those in [·] are the p-values of test statistics.

Table 9. Summary of regression results for the dynamic partial adjustment model of cash holdings

Variable	POLS	FE	AH-IV	FD-GMM	SYS-GMM	LD-GMM	LDP-GMM	LSDVC	BC	II
L.CASH	0.720*** (0.003)	0.375*** (0.005)	0.463*** (0.010)	0.468*** (0.010)	0.460*** (0.009)	0.897*** (0.067)	0.459*** (0.011)	0.521*** (0.002)	0.519*** (0.003)	0.511*** (0.003)
IND_CFASD	0.187*** (0.015)	0.079*** (0.021)	0.124*** (0.031)	0.090*** (0.031)	0.243*** (0.021)	-0.192 (0.132)	0.142*** (0.021)	0.149*** (0.019)	0.078*** (0.018)	0.202*** (0.019)
MB	0.019*** (0.002)	0.019*** (0.003)	-0.017*** (0.005)	-0.010* (0.005)	0.042*** (0.003)	-0.088*** (0.024)	0.017*** (0.004)	0.010*** (0.002)	0.013*** (0.002)	0.021*** (0.003)
SIZE	-0.005*** (0.002)	-0.129*** (0.010)	-0.403*** (0.020)	-0.379*** (0.020)	-0.034*** (0.004)	0.079 (0.057)	-0.154*** (0.010)	-0.115*** (0.008)	-0.106*** (0.006)	-0.033*** (0.007)
CFA	0.097*** (0.031)	0.129*** (0.041)	-0.098 (0.065)	-0.122* (0.067)	-0.290*** (0.046)	2.738** (1.309)	0.180*** (0.051)	0.110*** (0.037)	0.117*** (0.032)	0.047 (0.033)
NWC	-0.306*** (0.019)	-0.218*** (0.035)	0.771*** (0.056)	0.709*** (0.057)	-0.345*** (0.033)	0.195 (0.473)	-0.109*** (0.040)	-0.147*** (0.021)	-0.184*** (0.024)	-0.354*** (0.024)
CAPEX	-1.035*** (0.053)	-1.073*** (0.075)	0.629*** (0.102)	0.600*** (0.104)	-0.520*** (0.077)	-0.109 (2.171)	-1.352*** (0.086)	-1.146*** (0.075)	-1.079*** (0.060)	-1.324*** (0.062)
LEV	-0.456*** (0.020)	-0.292*** (0.035)	1.140*** (0.062)	1.080*** (0.062)	-0.792*** (0.041)	1.469*** (0.504)	-0.168*** (0.042)	-0.139*** (0.016)	-0.177*** (0.024)	-0.333*** (0.025)
RDS	0.107*** (0.007)	0.025** (0.012)	-0.039** (0.019)	-0.017 (0.019)	0.189*** (0.012)	0.963** (0.395)	0.021 (0.016)	0.005 (0.008)	0.013 (0.009)	0.019* (0.010)
DIV	-0.055*** (0.007)	-0.020 (0.015)	-0.068*** (0.021)	-0.061*** (0.021)	-0.075*** (0.014)	-0.115 (0.160)	-0.062*** (0.015)	-0.014 (0.017)	-0.021 (0.013)	-0.046** (0.013)
MA	-1.029*** (0.065)	-0.822*** (0.070)	0.742*** (0.091)	0.698*** (0.094)	-0.150* (0.080)	1.873 (2.018)	-1.107*** (0.084)	-1.013*** (0.046)	-0.920*** (0.065)	-0.856** (0.068)
SOA	28%	63%	54%	53%	54%	10%	54%	48%	48%	49%
AR(2) test			3.80 [0.00]	3.68 [0.00]	5.08 [0.00]					
Sargan test				694.87 [0.00]	798.14 [0.00]	27.65 [0.73]	24.53 [0.00]			
Year dummies	Yes	Yes	Yes	Yes	Yes	No	No	No	Yes	Yes
Observations	107,554	107,554	90,263	90,263	100,098	102	60,758	107,554	107,554	107,554

Notes: The speed of adjustment (SOA) is derived from the coefficient on the lagged cash. Our variable definitions strictly follow BKS; summary statistics are not reported but are available on request. AR(2) is a test for the null of no residual serial correlation and Sargan is a test for the joint null of validity of the moment conditions (over-identifying restrictions). *, **, and *** indicate significance of the coefficients at the 1%, 5%, and 10% levels, respectively. Figures in (·) are the standard errors and those in [·] are the p-values of test statistics.

Internet Appendix for “In search of robust methods for dynamic panel data models in empirical corporate finance”

Here we provide details of our estimation algorithms for the BC and II estimators, derivations of the formulas used in our simulations in Section 3 of the paper, and additional Monte Carlo simulation studies on leverage processes generated under random financing.

1. Estimation algorithm for BC

Consider the dynamic panel data ARDL(1,0) model:

$$y_{it} = \gamma y_{i,t-1} + \beta' \mathbf{x}_{it} + \eta_i + v_{it}, \quad i = 1, \dots, N; t = 2, \dots, T. \quad (\text{A.1})$$

Using the FE estimator eliminates the fixed effect, η_i , in (A.1) via the within-group transformation:

$$\tilde{y}_{it} = \gamma \tilde{y}_{i,t-1} + \beta' \tilde{\mathbf{x}}_{it} + \tilde{v}_{it}, \quad (\text{A.2})$$

where \tilde{y}_{it} , $\tilde{y}_{i,t-1}$, $\tilde{\mathbf{x}}_{it}$, and \tilde{v}_{it} are the deviations from the variables' respective individual means. Next, let $\hat{\pi}$ be the FE estimator of $\pi = (\gamma, \beta)'$ from (A.2). Then, for a fixed T , we have $\mathbf{E}[\hat{\pi}] \neq \pi$.

The bootstrap-based estimator, BC, corrects for the finite-sample bias of the FE estimator in the following way. First, suppose that we generate a sequence of L biased FE estimates, $\hat{\pi}^{(1)}(\pi), \dots, \hat{\pi}^{(L)}(\pi)$, by means of repeated sampling experiments. It then follows that

$$\mathbf{E}[\hat{\pi}] = \lim_{L \rightarrow \infty} \frac{1}{L} \sum_{\ell=1}^L \hat{\pi}^{(\ell)}(\pi), \quad (\text{A.3})$$

and $\bar{\pi}$ will be an unbiased estimator of π if the following condition holds:

$$\hat{\pi} = \frac{1}{L} \sum_{\ell=1}^L \hat{\pi}^{(\ell)}(\bar{\pi}). \quad (\text{A.4})$$

Hence, if we sample repeatedly from a population with parameters $\bar{\pi}$, and estimate $\hat{\pi}^{(\ell)}(\bar{\pi})$ in each sample, then $\bar{\pi}$ is an unbiased estimator of π if the average of $\hat{\pi}^{(\ell)}(\bar{\pi})$ corresponds to the FE estimate, $\hat{\pi}$, based on the original data. That is, a bias-corrected estimate of π can be

obtained by searching over the parameter space until a $\bar{\pi}$ is found that satisfies (A.4). This is implemented using a bootstrap procedure, as detailed below:

Step 1: We estimate the $N \times 1$ vector of the fixed effects, $\hat{\eta} = (\hat{\eta}_1, \dots, \hat{\eta}_N)' = (T-1)^{-1} \mathbf{D}'(\mathbf{y} - \mathbf{Z}\hat{\pi}_{(0)})$, and the residuals, $\tilde{\mathbf{v}} = (\tilde{v}'_1, \dots, \tilde{v}'_N)' = \tilde{\mathbf{y}} - \tilde{\mathbf{Z}}\hat{\pi}_{(0)}$, where $\hat{\pi}_{(0)}$ is the FE estimate based on the original data, $\mathbf{D} = \mathbf{I}_N \otimes (\mathbf{I}_{T-1} - \frac{1}{T-1} \mathbf{1}_{T-1} \mathbf{1}'_{T-1})$ is an idempotent transformation matrix that removes the fixed effects, and

$$\tilde{\mathbf{y}} = \begin{bmatrix} \tilde{\mathbf{y}}_1 \\ \vdots \\ \tilde{\mathbf{y}}_N \end{bmatrix}, \tilde{\mathbf{Z}} = \begin{bmatrix} \tilde{\mathbf{Z}}_1 \\ \vdots \\ \tilde{\mathbf{Z}}_N \end{bmatrix}, \tilde{\mathbf{y}}_i = \begin{bmatrix} \tilde{y}_{i2} \\ \vdots \\ \tilde{y}_{iT} \end{bmatrix}, \tilde{\mathbf{Z}}_i = \begin{bmatrix} \tilde{y}_{i,1} & \tilde{\mathbf{x}}_{i,2} \\ \vdots & \vdots \\ \tilde{y}_{i,T-1} & \tilde{\mathbf{x}}_{i,T} \end{bmatrix}, \tilde{\mathbf{v}}_i = \begin{bmatrix} \tilde{v}_{i2} \\ \vdots \\ \tilde{v}_{iT} \end{bmatrix}.$$

Next, following previous research (MacKinnon, 2002), we rescale the residuals as

$$\tilde{v}_{it}^* = \sqrt{\frac{T-1}{T-2}} \left(\frac{\tilde{v}_{it}}{\sqrt{m_{it}}} - \frac{1}{T-1} \sum_{s=2}^T \frac{\tilde{v}_{is}}{\sqrt{m_{is}}} \right), \quad (\text{A.5})$$

where m_{it} is the $\{(i-1)(T-1) + (t-1)\}$ th diagonal element of the idempotent matrix, $\mathbf{M} = \mathbf{I}_{N(T-1)} - \mathbf{D}'\mathbf{Z}(\mathbf{Z}'\mathbf{D}\mathbf{Z})^{-1}\mathbf{Z}'\mathbf{D}$.

Step 2: We generate the b th bootstrap sample residual, denoted $\tilde{\mathbf{v}}^{(b)}$, from the rescaled estimated residuals, $\tilde{\mathbf{v}}^*$, and generate a bootstrap sample $y_{it}^{(b)}$ using

$$y_{it}^{(b)} = \hat{\gamma}_{(0)} y_{i,t-1}^{(b)} + \hat{\beta}'_{(0)} \mathbf{x}_{it} + \hat{\eta}_i + \tilde{v}_{it}^{*(b)}, t = 2, \dots, T; b = 1, \dots, B, \quad (\text{A.6})$$

where we use $y_{i1}^{(b)} = y_{i1}$ and take \mathbf{x}_{it} as given. Next, we obtain the FE estimate, namely $\hat{\pi}^{(b)}$ ($\hat{\pi}_{(0)}$), using $\mathbf{y}^{(b)}$ and $\mathbf{Z}^{(b)}$ where

$$\mathbf{y}^{(b)} = \begin{bmatrix} \mathbf{y}_1^{(b)} \\ \vdots \\ \mathbf{y}_N^{(b)} \end{bmatrix}, \mathbf{Z} = \begin{bmatrix} \mathbf{Z}_1^{(b)} \\ \vdots \\ \mathbf{Z}_N^{(b)} \end{bmatrix}, \mathbf{y}_i^{(b)} = \begin{bmatrix} y_{i2}^{(b)} \\ \vdots \\ y_{iT}^{(b)} \end{bmatrix}, \mathbf{Z}_i^{(b)} = \begin{bmatrix} y_{i1}^{(b)} & \mathbf{x}_{i,2} \\ \vdots & \vdots \\ y_{i,T-1}^{(b)} & \mathbf{x}_{i,T} \end{bmatrix}.$$

Step 3: We repeat **Step 2** B times and calculate the sample mean of $\hat{\pi}^{(b)}(\hat{\pi}_{(0)})$ as follows:

$$\bar{\pi}_{(0)} = B^{-1} \sum_{b=1}^B \hat{\pi}^{(b)}(\hat{\pi}_{(0)}). \quad (\text{A.7})$$

Step 4: We define the difference between $\hat{\pi}_{(0)}$ and $\bar{\pi}_{(0)}$ as $\mathbf{d}_{(0)} = \hat{\pi}_{(0)} - \bar{\pi}_{(0)}$. If $\sup |\mathbf{d}_{(0)}| = 0$, then $\hat{\pi}_{(0)}$ will be an unbiased estimator of π according to condition (A.4). Otherwise, we update $\hat{\pi}_{(k+1)} = \hat{\pi}_{(k)} + \mathbf{d}_{(k)}$ for $k \geq 0$.

We iterate the bootstrap procedure outlined in **Steps 1 - 4** until condition (A.4) is satisfied. Following Everaert and Pozzi (2007), we set the number of bootstrap samples B to 1000. We use the convergence criterion, $\sup |\mathbf{d}_{(k)}| < 0.005$ and set the upper bound on the number of iterations at $k = 20$. We use the FE estimate based on the original data for $\hat{\pi}_{(0)}$.

Alternative resampling schemes. Resampling $\tilde{\mathbf{v}}^{(b)}$ in a non-parametric way, as described above, has the advantage that it does not require a distributional assumption for \mathbf{v} . Here, we consider two alternative resampling schemes. First, assuming that v_{it} is *i.i.d.* across i and over t , we resample \tilde{v}_{it} from

$$\tilde{\mathbf{v}}_i^{(b)} = (\tilde{v}_{i_1, t_2}^*, \dots, \tilde{v}_{i_N, t_T}^*), \quad i = 1, \dots, N, \quad (\text{A.8})$$

where the vectors of the indices (i_1, \dots, i_N) and (t_2, \dots, t_T) are obtained by drawing randomly with replacement from $(1, \dots, N)'$ and $(2, \dots, T)'$, respectively. Second, if we allow v_{it} to exhibit temporal dependence (e.g., conditional heteroskedasticity), we resample it by using the wild bootstrap (e.g., Goncalves and Kilian, 2004),

$$\tilde{\mathbf{v}}_i^{(b)} = (\tau_{i2} \tilde{v}_{j2}^*, \dots, \tau_{iT} \tilde{v}_{jT}^*), \quad i = 1, \dots, N, \quad (\text{A.9})$$

where the index j is drawn with replacement from $(1, \dots, N)'$, and τ_{it} is a binomial random variable (with a mean of 0 and a variance of 1) that takes the values -1 and 1 with equal probability, 0.5. This is asymptotically valid as either T , N or both become sufficiently large. We then collect all the resampled residual vectors in $\tilde{\mathbf{v}}^{(b)} = (\tilde{\mathbf{v}}_1^{(b)'}, \dots, \tilde{\mathbf{v}}_N^{(b)'})$.

2. Estimation algorithm for II

Gouriéroux et al. (2010) develop the II method for the AR(1) model. They show how to use simulations to approximate the binding function and then obtain an unbiased estimate by taking the inverse of this binding function. A similar procedure can be applied to a model with multiple regressors, although it is more complicated and the associated computation time increases exponentially. In our paper, while we only consider two parameters in the Monte Carlo simulations, we estimate nine to eleven parameters in the empirical applications.

Recall the dynamic panel data model under consideration:

$$y_{it} = \gamma y_{i,t-1} + \beta' \mathbf{x}_{it} + \eta_i + v_{it}, \quad i = 1, \dots, N; t = 1, \dots, T. \quad (\text{A.10})$$

Following Gouriéroux et al. (2000) we propose an II estimator using MLE (FE) as the baseline estimator. Further, we also choose the auxiliary model to be the true one. Let $\theta = (\gamma, \beta')$ be the vector of parameters with the parameter space, $\Theta \subset \mathbb{R}^{k+1}$. First, we estimate the model by FE and obtain the (biased) estimate, $\hat{\theta}$. We then use simulations to obtain an unbiased estimate of θ that satisfies:

$$\hat{\theta}_{II} = \underset{\theta \in \Theta}{\operatorname{argmin}} \left\| \hat{\theta} - b_{NT}(\theta) \right\|, \quad (\text{A.11})$$

where $b_{NT}(\theta)$ is the binding function defined by

$$b_{NT}(\theta) = \mathbf{E} \left[\tilde{\theta}^h(\theta) \right], \quad (\text{A.12})$$

and $\tilde{\theta}^h(\theta)$ is the FE obtained from the h -th simulated path, \tilde{y}_{it}^h in (A.13) below.

To generate the h -th simulated path of the DGPs given the parameter vector, θ , we calibrate it by

$$\tilde{y}_{it}^h = \gamma \tilde{y}_{i,t-1}^h + \beta' \tilde{\mathbf{x}}_{it} + \tilde{v}_{it}^h, \quad h = 1, \dots, H, \quad (\text{A.13})$$

where variables with tilde are the deviations from their respective individual means, e.g., $\tilde{y}_{it} = y_{it} - \bar{y}_i$ with $\bar{y}_i = \frac{1}{T} \sum_{t=1}^T y_{it}$. Without the loss of generality we assume that $\tilde{\mathbf{x}}_{it}$ is exogenous and the within transformation eliminates the individual effect, η_i from (A.10). Hence, the simulated

path in (A.13) can be generated simply by drawing \tilde{v}_{it}^h parametrically or nonparametrically. To this end we first obtain the residuals from the initial estimation:

$$\hat{v}_{it} = \tilde{y}_{it} - \hat{\gamma}\tilde{y}_{it} - \hat{\beta}'\tilde{\mathbf{x}}_{it}, \quad (\text{A.14})$$

and compute the variance, $\hat{\sigma}_v^2$. Then, we draw \tilde{v}_{it}^h from a normal distribution with a mean of zero and a variance of $\hat{\sigma}_v^2$. With a sufficient number of simulated paths, H , we can estimate the binding function by

$$\hat{b}_{NT}(\boldsymbol{\theta}) = \frac{1}{H} \sum_{h=1}^H \tilde{\boldsymbol{\theta}}(\boldsymbol{\theta}). \quad (\text{A.15})$$

In principle, $\hat{\boldsymbol{\theta}}_{II}$ can be obtained by minimizing the quadratic distance measure as

$$\hat{\boldsymbol{\theta}}_{II} = \underset{\boldsymbol{\theta} \in \Theta}{\operatorname{argmin}} (\hat{\boldsymbol{\theta}} - b_{NT}(\boldsymbol{\theta}))' \mathbf{W} (\hat{\boldsymbol{\theta}} - b_{NT}(\boldsymbol{\theta})), \quad (\text{A.16})$$

where \mathbf{W} is a positive definite weighting matrix. However, due to dimensionality, the above algorithm would be computationally demanding. Hence, we follow a simpler algorithm proposed by Gouriéroux et al. (2000), who observe that the objective function is (approximately) linear, and further assume that it holds in the multidimensional case, see also Phillips (2012) for the stationary case. Thus, we consider the following numerical optimization:

$$\boldsymbol{\theta}_{(n+1)} = \boldsymbol{\theta}_{(n)} + \lambda [\hat{\boldsymbol{\theta}} - \hat{b}_{NT}(\boldsymbol{\theta}_{(n)})], \text{ for } n = 0, 1, 2, \dots \quad (\text{A.17})$$

Starting from the (biased) initial estimate, $\hat{\boldsymbol{\theta}}$, we iterate the algorithm until the following convergence criterion is satisfied:

$$\|b_{NT}(\boldsymbol{\theta}) - \hat{\boldsymbol{\theta}}\| \leq c, \quad (\text{A.18})$$

where c is a small positive number.

To implement the above II procedure, we make three choices regarding the number of simulated paths, the estimation criterion, and the distribution of the data used for simulation. The first and the second choices are trivial. Gouriéroux et al. (2010) show that the II estimator can correct the bias and reduce the standard error effectively and simultaneously using a small

number of simulated paths (e.g., 10). In this study, we set the number of simulated paths to 50. Further, we also choose the estimation criterion in a flexible manner, depending on the baseline estimator used (e.g., MLE and GMM). The third choice is a more important one. It is crucial to calibrate the sampling distribution so that it can be close to the true DGP, although the latter is generally unknown. Gouriéroux et al. (2010) suggest using the auxiliary model to calibrate the simulated path of the data, in which case the number of parameters is the same in both models. We also follow this approach here and summarize our estimation algorithm in four steps below.

Step 1: We estimate the residuals, $\tilde{v}_{it} = \tilde{y}_{it} - \hat{\gamma}_{(0)}\tilde{y}_{i,t-1} - \hat{\beta}'_{(0)}\tilde{\mathbf{x}}_{it}$, where $\hat{\theta}_{(0)} = \left(\hat{\gamma}_{(0)}, \hat{\beta}_{(0)}\right)'$ is the FE estimate based on the original data. Next using the residuals, we estimate the variance of \tilde{v}_{it} , i.e., $\hat{\sigma}_v^2$.

Step 2: We generate the h th sample path of y_{it} using \tilde{v}_{it}^h drawn from a normal distribution with a mean of zero and a variance of $\hat{\sigma}_v^2$:

$$\tilde{y}_{it}^h = \hat{\gamma}_{(0)}\tilde{y}_{i,t-1}^h + \hat{\beta}'_{(0)}\tilde{\mathbf{x}}_{it} + \tilde{v}_{it}^h, \quad h = 1, \dots, H, \quad (\text{A.19})$$

where we use $\tilde{y}_{i1}^h = \hat{\beta}'_{(0)}\tilde{\mathbf{x}}_{i1} + \tilde{v}_{i1}$ and take $\tilde{\mathbf{x}}_{it}$ as given. Next, we obtain the FE estimate, namely $\hat{\theta}^h(\hat{\theta}_{(0)})$, using \tilde{y}_{it}^h and $\tilde{\mathbf{x}}_{it}$.

Step 3: We repeat Step 2 H times with $H = 50$, and calculate the binding function of $\hat{\theta}^h(\hat{\theta}_{(0)})$ as follows:

$$\hat{b}_{NT}(\hat{\theta}_{(0)}) = H^{-1} \sum_{h=1}^H \hat{\theta}^h(\hat{\theta}_{(0)}). \quad (\text{A.20})$$

Step 4: We define the difference between $\hat{\theta}_{(0)}$ and $\hat{b}_{NT}(\hat{\theta}_{(0)})$ as $\mathbf{d}_{(0)} = \|\hat{b}_{NT}(\hat{\theta}_{(0)}) - \hat{\theta}_{(0)}\|$. If $\mathbf{d}_{(0)} = 0$, then $\hat{\theta}_{(0)}$ will be an unbiased estimator of θ . Otherwise, we update

$$\hat{\theta}_{(n+1)} = \hat{\theta}_{(n)} + \lambda [\hat{\theta}_{(n)} - \hat{b}_{NT}(\hat{\theta}_{(n)})], \quad \text{for } n = 0, 1, 2, \dots \quad (\text{A.21})$$

We iterate the algorithm outlined in **Steps 1 - 4** until the convergence criterion in (A.18) is satisfied. We initially set $\lambda = 1$. For some simulated paths, however, the estimator tends to be close to the frontier of the set, $b_{NT}(\Theta)$ such that the algorithm fails to converge. In such a case we switch to $\lambda = 0.2$, as recommended by Gouriéroux et al. (2000).

3. Derivations of the formulas used in the DGP in the simulation experiments

3.1. Derivation of $Var(z_i)$ and $Cov(x_{it}, z_i)$

First, the covariance between x_{it} and z_i can be expressed as

$$Cov(x_{it}, z_i) = Cov(x_{it}, \bar{x}_i) - Cov(x_{it}, \bar{x}).$$

Using $\bar{x}_i = T^{-1} \sum_{s=1}^T x_{is}$, it is easily seen that

$$Cov(x_{it}, \bar{x}_i) = T^{-1} \left(\sum_{s=1}^T \rho^{|s-t|} \right) \sigma_x^2.$$

Next, using $\bar{x} = N^{-1} \sum_{i=1}^N \bar{x}_i$, and noticing that $Cov(x_{it}, \bar{x}_j) = 0$ for $i \neq j$, we have:

$$Cov(x_{it}, \bar{x}) = N^{-1} Cov(x_{it}, \bar{x}_i) = (NT)^{-1} \left(\sum_{s=1}^T \rho^{|s-t|} \right) \sigma_x^2.$$

Combining these results, we obtain:

$$Cov(x_{it}, z_i) = \left(\frac{N-1}{NT} \right) \left(\sum_{s=1}^T \rho^{|s-t|} \right) \sigma_x^2. \quad (\text{A.22})$$

Next, employing

$$Cov(\bar{x}_i, \bar{x}) = Cov\left(\bar{x}_i, \frac{1}{N} \sum_{i=1}^N \bar{x}_i\right) = \frac{1}{N} Var(\bar{x}_i),$$

since $Cov(\bar{x}_i, \bar{x}_j) = 0$ for $i \neq j$, and

$$Var(\bar{x}) = Var\left(\frac{1}{N} \sum_{i=1}^N \bar{x}_i\right) = \frac{1}{N^2} Var\left(\sum_{i=1}^N \bar{x}_i\right) = \frac{1}{N} Var(\bar{x}_i),$$

the variance of z_i is given by

$$Var(z_i) = Var(\bar{x}_i) + Var(\bar{x}) - 2Cov(\bar{x}_i, \bar{x}) = Var(\bar{x}_i) + \frac{1}{N} Var(\bar{x}_i) - \frac{2}{N} Var(\bar{x}_i).$$

Therefore,

$$Var(z_i) = \left(\frac{N-1}{N} \right) Var(\bar{x}_i). \quad (\text{A.23})$$

The variance of $\bar{x}_i = T^{-1} \sum_{t=1}^T x_{it}$ is then given by

$$\text{Var}(\bar{x}_i) = T^{-2} \text{Var} \left(\sum_{t=1}^T x_{it} \right) = T^{-2} \sum_{t=1}^T \sum_{s=1}^T \text{Cov}(x_{it}, x_{is}),$$

where

$$\begin{aligned} T^{-2} \sum_{t=1}^T \sum_{s=1}^T \text{Cov}(x_{it}, x_{is}) &= T^{-2} \left\{ 2 \left[\sum_{t=0}^{T-1} (T-t) \rho^t \right] - T \right\} \text{Var}(x_{it}) \\ &= \left[\frac{T^{-1} (1 - \rho^2) - 2T^{-2} (\rho - \rho^{T+1})}{(1 - \rho)^2} \right] \text{Var}(x_{it}). \end{aligned}$$

After some tedious, but otherwise straightforward, algebra, we finally obtain:

$$\text{Var}(\bar{x}_i) = \sigma_{\xi}^2 \left[\frac{T^{-1} (1 - \rho^2) - 2T^{-2} (\rho - \rho^{T+1})}{(1 - \rho)^2 (1 - \rho^2)} \right]. \quad (\text{A.24})$$

3.2. Derivation of φ_{it} and ψ_{it}

We rewrite the DGP as follows:

$$y_{it} = \frac{\beta}{1 - \gamma L} x_{it} + \frac{1}{1 - \gamma L} v_{it} + \frac{1}{1 - \gamma} \eta_i, \quad (\text{A.25})$$

where L is a lag operator. Replacing x_{it} and v_{it} in (A.25) with

$$x_{it} = \frac{1}{1 - \rho L} \xi_{it}, \quad v_{it} = \frac{1}{1 - \phi L} \varepsilon_{it},$$

we then obtain:

$$y_{it} = \beta \varphi_{it} + \psi_{it} + \frac{\eta_i}{1 - \gamma}, \quad (\text{A.26})$$

where $\varphi_{it} = \frac{1}{(1 - \gamma L)(1 - \rho L)} \xi_{it}$ and $\psi_{it} = \frac{1}{(1 - \gamma L)(1 - \phi L)} \varepsilon_{it}$ can be expressed as two mutually independent stationary AR(2) processes:

$$\varphi_{it} = (\gamma + \rho) \varphi_{i,t-1} - \gamma \rho \varphi_{i,t-2} + \xi_{it}, \quad (\text{A.27})$$

$$\psi_{it} = (\gamma + \phi) \psi_{i,t-1} - \gamma \phi \psi_{i,t-2} + \varepsilon_{it}. \quad (\text{A.28})$$

3.3. Derivation of the variance of the signal, σ_s^2

The variance of $s_{it} = w_{it} - v_{it}$ is defined as

$$\begin{aligned} \text{Var}(w_{it} - v_{it}) &= \text{Var}(w_{it}) + \text{Var}(v_{it}) - 2\text{Cov}(w_{it}, v_{it}) \\ &= \beta^2 \text{Var}(\varphi_{it}) + \text{Var}(\psi_{it}) + \text{Var}(v_{it}) - 2\text{Cov}(\psi_{it}, v_{it}), \end{aligned}$$

where

$$\text{Var}(v_{it}) = \frac{\sigma_\varepsilon^2}{1 - \phi^2},$$

$$\text{Cov}(\psi_{it}, v_{it}) = \text{Var}(v_{it}) + \gamma \text{Cov}(v_{it}, v_{i,t-1}) = \text{Var}(v_{it}) + \gamma\phi \text{Var}(v_{it}) = (1 + \gamma\phi) \text{Var}(v_{it}).$$

Thus,

$$\begin{aligned} \sigma_s^2 &= \beta^2 \text{Var}(\varphi_{it}) + \text{Var}(\psi_{it}) - (1 + 2\gamma\phi) \text{Var}(v_{it}) \\ &= \beta^2 \sigma_\xi^2 \left[1 + \frac{(\gamma + \rho)^2 (\gamma\rho - 1)}{1 + \gamma\rho} - (\gamma\rho)^2 \right]^{-1} \\ &\quad + \sigma_v^2 (1 - \phi^2) \left[1 + \frac{(\gamma + \phi)^2 (\gamma\phi - 1)}{1 + \gamma\phi} - (\gamma\phi)^2 \right]^{-1} - (1 + 2\gamma\phi) \sigma_v^2 \\ &= \beta^2 \sigma_\xi^2 A + \sigma_v^2 B, \end{aligned} \tag{A.29}$$

where

$$A = \left[1 + \frac{(\gamma + \rho)^2 (\gamma\rho - 1)}{1 + \gamma\rho} - (\gamma\rho)^2 \right]^{-1},$$

$$B = (1 - \phi^2) \left[1 + \frac{(\gamma + \phi)^2 (\gamma\phi - 1)}{1 + \gamma\phi} - (\gamma\phi)^2 \right]^{-1} - (1 + 2\gamma\phi).$$

Finally, the variances of ξ_{it} and ε_{it} can be obtained as follows:

$$\sigma_\xi^2 = \sigma_v^2 \left(\frac{\vartheta - B}{\beta^2 A} \right) \text{ and } \sigma_\varepsilon^2 = (1 - \phi^2) \sigma_v^2. \tag{A.30}$$

3.4. Derivation of the initial conditions

Following Kiviet (1995), we can easily obtain the initial conditions for x_{i0} , φ_{i0} , and ψ_{i0} as follows: First,

$$x_{i0} = \xi_{i0} (1 - \rho^2)^{-1/2}. \quad (\text{A.31})$$

Next,

$$\varphi_{i0} = \xi_{i0} \left[\frac{\text{Var}(\varphi_{it})}{\sigma_\xi^2} \right]^{1/2}, \quad (\text{A.32})$$

$$\varphi_{i1} = \varphi_{i0} \text{Corr}(\varphi_{it}, \varphi_{i,t-1}) + \xi_{i1} \left[\frac{\text{Var}(\varphi_{it})}{\sigma_\xi^2} \right]^{1/2} \left[1 - \text{Corr}(\varphi_{it}, \varphi_{i,t-1})^2 \right]^{1/2}, \quad (\text{A.33})$$

where

$$\text{Var}(\varphi_{it}) = \sigma_\xi^2 [1 - (\gamma + \rho) \text{Corr}(\varphi_{it}, \varphi_{i,t-1}) + \gamma\rho \text{Corr}(\varphi_{it}, \varphi_{i,t-2})]^{-1},$$

$$\text{Corr}(\varphi_{it}, \varphi_{i,t-1}) = \frac{\gamma + \rho}{1 + \gamma\rho},$$

$$\text{Corr}(\varphi_{it}, \varphi_{i,t-2}) = (\gamma + \rho) \text{Corr}(\varphi_{it}, \varphi_{i,t-1}) - \gamma\rho.$$

Finally,

$$\psi_{i0} = \varepsilon_{i0} \left[\frac{\text{Var}(\psi_{it})}{\sigma_\varepsilon^2} \right]^{1/2}, \quad (\text{A.34})$$

$$\psi_{i1} = \psi_{i0} \text{Corr}(\psi_{it}, \psi_{i,t-1}) + \varepsilon_{i1} \left[\frac{\text{Var}(\psi_{it})}{\sigma_\varepsilon^2} \right]^{1/2} \left[1 - \text{Corr}(\psi_{it}, \psi_{i,t-1})^2 \right]^{1/2}, \quad (\text{A.35})$$

where

$$\text{Var}(\psi_{it}) = \sigma_\varepsilon^2 [1 - (\gamma + \phi) \text{Corr}(\psi_{it}, \psi_{i,t-1}) + \gamma\phi \text{Corr}(\psi_{it}, \psi_{i,t-2})]^{-1},$$

$$\text{Corr}(\psi_{it}, \psi_{i,t-1}) = \frac{\gamma + \phi}{1 + \gamma\phi},$$

$$\text{Corr}(\psi_{it}, \psi_{i,t-2}) = (\gamma + \phi) \text{Corr}(\psi_{it}, \psi_{i,t-1}) - \gamma\phi.$$

4. Simulation in the presence of random financing choices

Chang and Dasgupta (2009) examine the dynamics of corporate capital structure using simulated data, where firms are assumed to follow a random type of financing to offset their financing deficit, k_{it} : firms issue debt and equity with probabilities p and $1 - p$, respectively. We now examine the performance of each estimator in an additional simulation in which the leverage processes are generated on the basis of this type of random financing. As in Chang and Dasgupta (2009), with random financing the leverage ratio evolves as follows:

$$y_{i,t+1} = \begin{cases} \frac{k_{it}}{1+k_{it}} + \left(\frac{1}{1+k_{it}}\right)y_{it} & \text{with probability } p \\ \left(\frac{1}{1+k_{it}}\right)y_{it} & \text{with probability } 1 - p \end{cases}, \quad i = 1, \dots, N; t = 0, \dots, T - 1. \quad (\text{A.36})$$

We then extend Chang and Dasgupta's (2009) simulation design in two important ways. First, to be consistent with the main simulations in Section 3 of our paper, we incorporate the firm fixed effect and an idiosyncratic error into the DGP:

$$y_{i,t+1} = \begin{cases} \frac{k_{it}}{1+k_{it}} + \left(\frac{1}{1+k_{it}}\right)y_{it} + \eta_i + v_{i,t+1} & \text{with probability } p \\ \left(\frac{1}{1+k_{it}}\right)y_{it} + \eta_i + v_{i,t+1} & \text{with probability } 1 - p \end{cases}, \quad i = 1, \dots, N; t = 0, \dots, T - 1. \quad (\text{A.37})$$

Second, while Chang and Dasgupta (2009) fix k_{it} , we relax this strong restriction and allow for heterogeneity in k_{it} across firms via several experiments: we draw k_{it} from the uniform distribution across firms and over time.

For simplicity, in (A.37), we let $p = 0.5$. That is, the firm's financial decision is determined by a coin toss. In our simulations, we estimate the following AR(1) model:

$$y_{i,t+1} = \gamma y_{it} + \eta_i + v_{i,t+1}, \quad i = 1, \dots, N; t = 0, \dots, T - 1, \quad (\text{A.38})$$

where we set $(N, T) = (500, 10)$ and use 1,000 replications. We further set $y_{i0} = 0$ but discard the first 10 years of observations to reduce the effect of the initial values. We draw $v_{i,t+1}$ from the standard normal distribution, and η_i from the uniform distribution, $U(-a, a)$, for $a > 0$. To control for the loading factor, μ , we set $a = \sqrt{3}\mu$. Then we let $\mu = 1$ and so $a = \sqrt{3}$. That is, we

draw η_i from $U(-\sqrt{3}, \sqrt{3})$. Note that, from (A.37), the AR(1) coefficient is $\gamma = 1/(1 + k_{it})$.¹

Table A.1 below presents the results of three simulation experiments. First, as in Chang and Dasgupta (2009), we simply fix k_{it} at 0.15 for all firm-year observations (the choice of $k_{it} = 0.15$ corresponds to the empirical mean of the financing deficit ratio). Then, the true AR(1) coefficient is $1/(1 + 0.15) \approx 0.870$. The results contained in the first row suggest that AH-IV produces the most accurate estimate of the AR(1) coefficient, γ , although LD-GMM, LDP-GMM, LSDVC, BC, II, and FD-GMM also perform reasonably well. However, DPF, SYS-GMM, and POLS generate a moderate to large amount of upward bias, while FE produces a substantial amount of downward bias.

Second, we allow k_{it} to vary across firms and over time. To do so, we draw k_{it} from $U(0, 0.3)$. The AR(1) coefficient is thus equal to $\mathbf{E}[1/(1 + k_{it})] \approx 0.875$.² We find that the results, reported in the second row, are very similar to those documented for the first experiment.

In the third and final experiment, we allow for both issuing and retiring/repurchasing decisions. That is, the firm issues debt (or retires it if $k_{it} < 0$) with probability p , and issues equity (or repurchases it if $k_{it} < 0$) with probability $1 - p$, where $p = 0.5$. Specifically, we generate values from $k_{it} \sim U(-0.2, 0.2)$ and find that in this case $\mathbf{E}[1/(1 + k_{it})] \approx 1.014$. The results in the third row show that AH-IV, FD-, LD-, LDP-GMM, LSDVC, and BC have negligible bias, while DPF, SYS-GMM, and POLS also perform reasonably well.

In sum, our simulation experiments show that AH-IV is the most favorable method, while LD-GMM, LDP-GMM, LSDVC, and BC are also acceptable. In contrast, DPF generally has

¹Unlike EF, we show that in the random financing simulation, the SOA should not be equal to zero. This is because, by construction, the true SOA is $\delta = k_{it}/(1 + k_{it})$, meaning the SOA would only be equal to zero in a hypothetical scenario where the firm did not have any financing deficit, i.e., $k_{it} = 0$, and leverage was a non-stationary process, i.e., $\gamma = 1$.

²It is approximated using the Taylor expansion up to the fourth moment, as follows:

$$\begin{aligned} \mathbf{E}\left[\frac{1}{1 + k_{it}}\right] &\approx \frac{1}{1 + \mathbf{E}[k_{it}]} - \frac{\mathbf{E}[k_{it} - \mathbf{E}[k_{it}]]}{(1 + \mathbf{E}[k_{it}])^2} + \frac{\mathbf{E}[(k_{it} - \mathbf{E}[k_{it}])^2]}{(1 + \mathbf{E}[k_{it}])^3} - \frac{\mathbf{E}[(k_{it} - \mathbf{E}[k_{it}])^3]}{(1 + \mathbf{E}[k_{it}])^4} + \frac{\mathbf{E}[(k_{it} - \mathbf{E}[k_{it}])^4]}{(1 + \mathbf{E}[k_{it}])^5} \\ &= \frac{1}{1 + \frac{(u+l)}{2}} + \frac{\frac{1}{12}(u-l)^2}{\left\{1 + \frac{(u+l)}{2}\right\}^3} + \frac{9}{5} \left[\frac{\frac{1}{144}(u-l)^4}{\left\{1 + \frac{(u+l)}{2}\right\}^5} \right], \end{aligned}$$

where $k_{it} \sim U(l, u)$.

non-negligible (upward) bias, and is always outperformed by these estimators. Unlike in Elsas and Florysiak (2014), our simulation results suggest that no estimator can help researchers to distinguish between real target adjustment behavior and mechanical mean reversion caused by random financing.

Additional references

Goncalves, S., and L. Kilian. 2004. Bootstrapping autoregressions with conditional heteroskedasticity of unknown form. *Journal of Econometrics* 123: 89-120.

MacKinnon, J.G. 2002. Bootstrap inference in econometrics. *Canadian Journal of Economics* 35: 615-645.

Phillips, P.C.B. 2012. Folklore theorems, implicit maps and indirect inference. *Econometrica* 80: 425-454.

Table A.1: Performance of alternative estimators in random-financing leverage processes

Simulation	$\mathbf{E}\left(\frac{1}{1+k_{it}}\right)$	POLS	FE	AH-IV	FD-GMM	SYS-GMM	LD-GMM	LDP-GMM	LSDVC	BC	DPF	II
$k_{it} = 0.15$	0.870	0.137	-0.239	0.005	-0.054	0.106	0.012	0.026	-0.020	-0.048	0.074	0.043
$k_{it} \sim U(0, 0.3)$	0.875	0.131	-0.243	0.005	-0.051	0.099	0.017	0.027	-0.023	-0.048	0.064	0.042
$k_{it} \sim U(-0.2, 0.2)$	1.014	0.050	-0.116	0.000	0.001	0.048	-0.001	0.000	0.006	0.004	0.034	0.089

Notes: This table reports the performance of each estimator when the leverage processes (DGPs) are generated according to random financing (A.37). $\mathbf{E}[1/(1+k_{it})]$ is the true AR(1) coefficient. k_{it} is the financing deficit divided by total assets. The performance of each estimator is evaluated by the bias of the AR(1) coefficient estimate ($\hat{\gamma}$).