BANK OF CANADA
BANQUE DU CANADA

Working Paper/Document de travail
2014-55

Persistent Leverage in Portfolio Sorts: An Artifact of Measurement Error?

by Michael Mueller

www.bank-banque-canada.ca

Bank of Canada Working Paper 2014-55

December 2014

Persistent Leverage in Portfolio Sorts: An Artifact of Measurement Error?

by

Michael Mueller

Financial Markets Department
Bank of Canada
Ottawa, Ontario, Canada K1A 0G9
mmueller@bankofcanada.ca

ISSN 1701-9397

Acknowledgements

I am grateful to my thesis supervisor, Murray Carlson, and Ron Giammarino, Adlai Fisher, Russell Lundholm, Jason Chen, Rob Heinkel, Kai Li, Henry Siu, Kin Lo, Nathalie Moyen, Michael Faulkender, Halit Gonenc, Danielle Xu, Atanas Mihov, Joshua Slive, Teodora Paligorova, Jonathan Witmer, and Scott Hendry. I would also like to thank seminar participants at the University of British Columbia, Florida International University, San Francisco State University, the Bank of Canada, and the NFA, EFMA, FMA and SFA conferences for their insights and helpful discussion.

Abstract

Studies such as Lemmon, Roberts and Zender (2008) demonstrate how stable firms' capital structures are over time, and raise the question of whether new theories of capital structure are needed to explain these phenomena. In this paper, I show that trade-off theory-based empirical proxies that are observed with error offer an alternative explanation for the persistence in portfolio-leverage levels. Measurement error noise equal to 80% of the cross-sectional variation in the market to book ratio, coupled with slight mismeasurement of other factors, matches simulated data moments to empirical moments. This suggests that unobserved investment opportunities play an important role in explaining leverage ratios.

JEL classification: G32, C18
Bank classification: Financial markets; Econometric and statistical methods

Résumé

Certains auteurs comme Lemmon, Roberts et Zender (2008) démontrent la stabilité à long terme de la structure financière des sociétés et se demandent s'il est nécessaire de formuler de nouvelles théories sur la structure du capital pour expliquer un tel phénomène. Dans la présente étude, l'auteur montre que les indicateurs empiriques fondés sur la théorie de l'arbitrage et observés de façon erronée offrent une autre explication de la persistance de l'effet de levier financier dans les portefeuilles. Le recours à un bruit d'erreur de mesure équivalant à 80 % de la variation transversale du ratio cours / valeur comptable, conjugué à la mesure légèrement inexacte d'autres facteurs, se traduit par une concordance entre les moments simulés et les moments empiriques. Ce résultat laisse à penser que les possibilités d'investissement non observées contribuent fortement à expliquer les ratios de levier financier.

Classification JEL : G32, C18
Classification de la Banque : Marchés financiers; Méthodes économétriques et statistiques

1 Introduction and Background

Financial leverage, the ratio of a firm's debt to its assets, is generally accepted to be a slow-moving variable.[1] A recent addition to the literature on leverage persistence is by Lemmon, Roberts, and Zender (2008), who document that when firms are sorted into portfolios based on their leverage, the average leverage levels of these portfolios do not converge to the unconditional mean, even after 20 years. Furthermore, the authors establish that this phenomenon persists even after controlling for factors that are believed to drive leverage: when firms are sorted into portfolios on the basis of residuals from a regression of leverage on a set of determinants motivated by the trade-off theory of capital structure, and then tracked for 20 years post-portfolio formation, the mean leverage levels of these portfolios still exhibit long-term persistence. Lemmon, Roberts, and Zender (2008) show that firm fixed effects, rather than any of the commonly used time-varying explanatory variables, explains most of this persistence. Since firm fixed effects are unsatisfactory from a theoretical point of view, the authors suggest that new theories of capital structure may be needed to explain the persistence of leverage ratios.

In this paper, I seek an explanation for this fixed-effect-like feature of leverage data in the context of portfolio sorts. In a framework motivated by the trade-off theory of capital structure, I focus on measurement error in empirical proxies of the true underlying economic variables as a possible cause of this phenomenon. I demonstrate that if leverage is governed by a persistent explanatory variable that is mismeasured, using the mismeasured explanatory variable in a regression creates an artificial persistence in residual-sorted portfolios in the following manner: conditional on an observed residual, future expectations of leverage are no longer equal to the unconditional mean. Instead, a large positive residual will forecast above-average future leverage. This occurs because the estimated residual is correlated with the true unobservable explanatory variable, which in turn predicts leverage.

It is reasonable to assume that the factors used in capital structure analysis contain measurement error. The market-to-book ratio is often used as a proxy for investment opportunities: under the trade-off theory, the larger this ratio, the less debt a firm should be willing to take on to take advantage of the interest tax shield, *ceteris paribus*. The real option to invest is riskier than assets in place, which would thus warrant a capital structure with less debt. As an example of how the market-to-book ratio could be inaccurate, consider Facebook between 2012 and 2014: over a span of less than two years, Facebook's market-to-book ratio almost quadrupled. Yet, it is hard to imagine

[1] For example, when modelled as an AR(1) process, the estimated autocorrelation coefficient for the leverage ratio tends to be around 0.9.

that the same is true of Facebook's actual, unobserved investment opportunities, since neither the company nor its business environment have changed that drastically over the two-year time frame.

While measurement error in economic variables is intuitively reasonable, a natural question is: how much measurement error is consistent with the documented persistence in portfolio sorts? I find that if we assume that a single factor drives leverage (we can think of this factor as a composite of many trade-off theory-based explanatory variables), then the measurement-error variance of this "composite" variable needs to be 42% larger than its cross-sectional variance to reproduce the Lemmon, Roberts, and Zender (2008) portfolio leverage levels. While this seems large, even much smaller levels of measurement error produce a remarkable level of persistence in residual-based portfolio sorts. For instance, if the ratio of measurement error to state noise in the explanatory variable is as low as 25%, the residual-based portfolios still exhibit a sizable amount of persistence. Therefore, measurement error is likely an important contributor to the persistence in residual-sorted leverage portfolios.

In a further analysis, I examine measurement error in the explanatory variables used in studies such as Lemmon, Roberts, and Zender (2008), Rajan and Zingales (1995) and Frank and Goyal (2009). Using a simulated method-of-moments approach, I match simulated data moments to empirical moments. The moments include portfolio leverage levels, summary statistics on leverage and explanatory factors, and typical estimated leverage regression parameters. I find that in this expanded calibration, low quantities of measurement error in profitability, tangibility, and industry leverage, coupled with a measurement-error variance equal to about 80% of the cross-sectional variation in the market-to-book ratio, result in a good match between simulated and empirical moments. This finding is consistent with studies such as Erickson and Whited (2006), who document that a large amount of the variability in the market-to-book ratio can be attributed to measurement error, and not to true Tobin's q. My finding suggests that unobserved investment opportunities play an important role in explaining leverage ratios. The puzzle of leverage persistence in a portfolio sort context may therefore be resolved by the proper measurement of existing economic variables, rather than by devising new theories of capital structure.

Related Literature

Twenty-six years after the seminal work of Modigliani and Miller (1958), Myers (1984) remarked that "we do not know how firms choose the debt, equity or hybrid securities they issue." Much effort has gone into a better understanding of what drives corporate capital structures, yet the question of whether firms do have a target capital structure towards which they actively adjust their debt/equity mix is open. Titman and Wessels (1988) find several variables that help predict a firm's capital structure, yet the variables do not correspond to any one theory. Fischer, Heinkel, and Zechner (1989) propose a dynamic capital structure model, where firms adjust towards an optimum, but are hampered by adjustment costs. Hovakimian, Opler, and Titman (2001) provide evidence that firms behave in a fashion consistent with a trade-off model, a finding that is echoed in Leary and Roberts (2005). Roberts (2002) shows that firms appear to adjust towards firm-specific time-varying targets, that adjustment speeds vary considerably across industries, and that accounting for measurement error increases the speed of adjustment. In their extensive survey, Graham and Harvey (2001) find some, though not particularly strong, support for the trade-off theory. Baker and Wurgler (2002), on the other hand, suggest that firms' issuance behaviour is driven by attempts to time the market, while Welch (2004) shows that firms appear to do nothing to counteract mechanistic stock return effects on market leverage. In Hennessy and Whited (2005), there is no leverage target towards which firms adjust, in spite of their optimizing behaviour. Chang and Dasgupta (2009) argue that the evidence for the trade-off theory is not as strong as it may seem, since random financing generates data that are similar to what actually is observed. Overall, the evidence for firms adjusting towards an optimal capital structure is mixed.

Lemmon, Roberts, and Zender (2008) contribute to this strand of literature. By analyzing leverage portfolios, they find that "high (low) levered firms tend to remain as such for over two decades." The authors sort firms into quartile portfolios on the basis of firm leverage, and track the portfolios' average leverage levels for 20 years. They find a large initial dispersion between the leverage portfolios. Over the years, the portfolios do converge to some extent (most of the convergence happens early on), but significant differences remain, even after 20 years. Controlling for known determinants of capital structure in a regression, and sorting on the regression residuals instead of on actual leverage has little effect on the results: the time series of the residual-based portfolios are very similar to those of the actual leverage-based portfolios. The authors thus characterize leverage ratios as containing a permanent, time-invariant leverage target and a transitory mean-reverting component, which may be attributed to active capital structure management. Furthermore, Lemmon, Roberts, and Zender (2008) show through a variance decomposition that a firm fixed effect has more explanatory power than any existing time-varying determinants, and thus is the factor that best explains the cross-section of capital structure. The authors also show that this effect pre-dates firms' IPOs, which suggests that "... changes in the distribution of control, the information environment, and the access to capital markets... do little to alter the relative costs and benefits that determine firms' preferred leverage ratios."

DeAngelo, DeAngelo, and Whited (2011) offer a potential explanation of leverage persistence. In their model, firms can finance investment by using retained earnings (cash), by issuing debt or by issuing equity. Carrying cash forces the firm to incur agency costs proportional to the cash balance. Issuing debt is costless, but credit rationing caps a firm's debt capacity. Equity issuance, on the other hand, is costly. Generally, firms will avoid carrying a cash balance owing to the associated agency costs. Instead, when installing new capital, they will free up debt capacity so as to avoid a costly equity issuance. This "transitory" debt, coupled with various frictions in the model and cross-sectional dispersion in profitability shocks, leads to actual leverage-based sorts that resemble those in Lemmon, Roberts, and Zender (2008).

Another recent effort to explain leverage persistence is by Menichini (2010). His model, which includes agency costs and endogenous investment, leverage, and dividend payouts generates portfolio sorts on both actual and unexpected leverage that exhibit long-term persistence. This occurs largely because in his model there is no single long-term mean towards which firms revert.

On the other hand, DeAngelo and Roll (2011) question the stability of capital structures altogether, and argue that it is the exception and not the rule. They find that many firms that have been listed for 20 or more years have leverage levels in at least three different quartiles.

The goal of this study is to reconcile some of the recent empirical findings regarding the persistence of capital structure. Lemmon, Roberts, and Zender (2008) provide evidence for leverage persistence via their portfolio sorts, and find that a firm fixed effect best explains the data. Both Roberts (2002) and Flannery and Rangan (2006) suggest that measurement error is partly responsible for the sluggish convergence of leverage ratios towards their mean, as measured by the adjustment speed parameter in a partial-adjustment framework. Flannery and Rangan (2006) also show that including firm fixed effects increases the estimated convergence speeds. My paper expands on this literature: I explore the channel through which measurement error can add to the persistence of leverage and load on a fixed effect in a portfolio sort setting. To extract the amount of measurement error necessary to reproduce the stylized facts, I use a model in which persistence stems from slow-moving, mismeasured leverage determinants rather than from a firm fixed effect.

2 Persistence in Portfolio Sorts and Possible Explanations

The data sample consists of firms listed in the annual Compustat database between 1965 and 2003. Financial institutions and firms with missing asset or debt values are excluded. Leverage is constrained to lie in the closed unit interval. Definitions of all variables are given in Appendix A. Explanatory variables are winsorized at the 1^{st} and 99^{th} percentiles. Table 1 presents summary statistics that are similar to those in Lemmon, Roberts, and Zender (2008). The table also shows a prominent feature of the data – namely, the existence of zero-leverage firms, whose proportion is sizable (see e.g., Strebulaev and Yang (2013)).

[Table 1 about here.]

The procedure for the portfolio sorts follows Lemmon, Roberts, and Zender (2008) and generates similar results: Starting in 1965, and then every year thereafter, I sort firms into four quartile portfolios on the basis of their book leverage level. I then compute the mean leverage of each portfolio for the next 20 years, keeping its composition constant (save for potential exits from the sample).[2] This results in a total of 38 portfolio time series, each with a length of 20 years or less. The portfolios are then averaged cross-sectionally in event time. Panel A of Figure 1 shows the long-term persistence in raw leverage that this procedure produces.

[Figure 1 about here.]

Panel A of Figure 1 is virtually identical to Panel A of Figure 1 in Lemmon, Roberts, and Zender (2008). There is wide cross-sectional dispersion among portfolios in the initial sorting period. This dispersion is followed by an initially quick convergence toward the overall mean, which starts to taper off noticeably as we move further away from the portfolio's formation year. After 20 years, there is still a 16-percentage-point difference between the highest and lowest leverage portfolios.

Since the pattern in Figure 1 could be the result of cross-sectional variation in the underlying determinants of firm leverage, Lemmon, Roberts, and Zender (2008) regress leverage on lagged firm size, profitability, tangibility, market-to-book equity and Fama-French 38-industry dummies.[3] In a variant of the original portfolio sorts, firms are

[2]Note that starting in 1983, the available time series for each portfolio will decrease by one year each year.

[3]The regressions are estimated every year, which allows for time-varying coefficient estimates. Size, profitability and an industry dummy are used, e.g., in Titman and Wessels (1988), while tangibility and market-to-book equity are used, e.g., in Rajan and Zingales (1995).

now sorted into portfolios based on the estimated regression residuals (the "unexpected leverage") instead of on actual leverage.

Similar to Lemmon, Roberts, and Zender (2008), I also perform sorts based on "unexpected leverage," but instead of industry dummies, I use mean industry leverage (identified by Frank and Goyal (2009) as an important determinant of leverage). Panel B of Figure 1 depicts the residual-based sorts and reproduces the findings of Lemmon, Roberts, and Zender (2008) virtually identically (see their Panel A, Figure 2). The cross-sectional portfolio dispersion in the formation year is slightly reduced, when compared with sorting on actual leverage. However, large differences between the portfolios remain over the entire 20 years. This is inconsistent with a well-specified regression, where convergence of the portfolio-leverage averages towards the overall mean should speed up, since the residuals would not contain any information about firms' future leverage levels.

The persistent differences between leverage portfolios cast doubt on theories of capital structure that have the firm adjust towards some kind of optimal mix of debt and equity. However, before turning to new theories, it is useful to know to what extent existing theories are able to accommodate leverage persistence in a portfolio-sort setting. There are several possible channels that can give rise to the persistence of residual-based leverage portfolios. All are manifestations of the same underlying fact: the regression residuals must contain information about future levels of leverage. The first channel is that empirical specifications of leverage regressions are plagued by an omitted variable problem. In its simplest form, leverage is largely determined by a time-invariant firm fixed effect.[4] Since the omitted fixed effect is constant over time, sorting on the regression residual would lead to leverage persistence in the portfolios.

Another possibility is that regressions omit one or more time-varying persistent variables that determine leverage. As with a fixed effect, the regression residuals are no longer just noise, but contain information. An example of this strand of literature is the recent paper by DeAngelo, DeAngelo, and Whited (2011), who model firms as incurring transitory debt obligations that represent deliberate, but temporary, deviations from a target capital structure. When they carry out portfolio sorts based on their simulated firms' leverage, they find persistent portfolio-leverage levels. While DeAngelo, DeAngelo, and Whited (2011) do not sort on regression residuals, some of the persistence would likely remain in residual-based sorts: without accounting for the level of transitory debt, an omitted variable would be imbedded in the residual.

[4]A firm fixed effect can be thought of as every firm having its own intercept in the regression.

A third possibility is that the regressions are misrepresentations of the underlying economic mechanism. For instance, if firms face adjustment costs as in Fischer, Heinkel, and Zechner (1989) or Hennessy and Whited (2005), there would be no target leverage as implied by the regression. Instead, the firm would alter its capital structure only after leverage drifts outside of a certain range.

Finally, it is possible that our economic models are correct, but the empirical proxies for the benefits and costs of debt are inaccurate. If the explanatory variables were mismeasured, the regression residual and leverage itself would be correlated, and sorting on the residual would resemble sorting on leverage.

Generally, distinguishing conclusively between these alternatives is difficult. Including firm fixed effects in leverage regressions explains much of the cross-sectional variation among firms. It does not completely eliminate the interesting portfolio patterns in residual-based sorts, as shown in Panel C of Figure 1. While a fixed effect reduces initial dispersion, there is still no convergence, since the average leverage level of the low-leverage portfolio is now substantially higher than that of the high-leverage portfolio after 20 years. In essence, including a fixed effect de-means the portfolio-leverage time series, but the patterns, albeit shifted, still remain.

In the sections to follow, I show that the stylized facts of persistence in leverage sorts obtain if leverage is a function of one or more mismeasured explanatory variables that exhibit a certain degree of persistence. My results cannot prove that measurement error is the sole source of the phenomena I study. However, measurement error in explanatory variables is intuitively sensible and consistent with the data, which studies like Flannery and Rangan (2006), Roberts (2002), and Erickson and Whited (2006) confirm.

2.1 Persistence in Leverage Portfolios Due to Measurement Error

In this section, I formally examine persistence in leverage sorts, starting with the case where we can observe all variables perfectly. If leverage is determined by one or more persistent explanatory variables, then sorting firms into portfolios based on observed leverage will produce persistent portfolios. The persistence of the portfolio-leverage time series reflects the persistence of the determinants. When firms are sorted into portfolios based on residuals from a regression that controls for the persistent determinants of leverage, differences between portfolio-leverage levels also arise naturally during the formation period. This happens because the residual is, by construction, correlated with the dependent variable, leverage. However, if the regression residuals are uncorrelated over time, then any initial difference between the portfolios should completely vanish in the subsequent period. Finally, I show how persistent differences in residual-based portfolios arise when explanatory variables are measured with error. In my setup, I do not assume measurement error to be persistent or firm-specific. The only persistent variable is the true, but unobserved regressor.

2.1.1 Base Case: A Correctly Specified Model

I begin with a world where leverage is a function of a single persistent explanatory variable, which is perfectly measured. There exists a panel of firms, where i indexes a firm and t indexes time. The dependent variable of interest – leverage – is denoted by lev_{it}. Its true relationship to the explanatory variable x_{it} (e.g., size, profitability, or the book-to-market ratio) is given by:

$$lev_{it} = \beta x_{it} + u_{it} \tag{1}$$

where $u_{it} \sim N(0, \sigma^2_{u_{it}})$ is an error term and βx_{it} can be thought of as firm i's leverage target, towards which the firm fully adjusts every time period. The firm's actual leverage lev_{it} equals its target, plus a random deviation u_{it}. This deviation, which Lemmon, Roberts, and Zender (2008) refer to as unexpected leverage, represents an exogenous shock that occurs after adjustment to the target has taken place. For instance, a change in the market value of the firm's equity would cause actual leverage to deviate from the target. The leverage determinant x_{it} follows an AR(1) process of the form:

$$x_{it} = \phi x_{it-1} + \epsilon_{it} \tag{2}$$

where $\phi > 0$ and $\epsilon_{it} \sim N(0, \sigma^2_{\epsilon_{it}})$. For simplicity, I assume that the explanatory variable and, hence, leverage, have a mean of 0.[5] Under this specification, leverage directly inherits the dynamics of the explanatory variable. Tomorrow's expected leverage, conditional on today's observed leverage, is governed by the magnitude of the autocorrelation coefficient of the AR(1) process, since $\mathbb{E}(lev_{it}|lev_{it-1}) = \phi lev_{it-1}$.

If the value of ϕ is large,[6] then if we form portfolios by sorting on leverage and track their evolution over time, the high-leverage portfolios decline only slowly towards

[5]This assumption is not crucial and will be relaxed in Section 3 We could easily add a mean to the explanatory variable x_{it} without affecting any of the conclusions. Alternatively, since it is possible in my setup for leverage to be negative, it is perhaps most natural to think of the lev_{it} as logit leverage $ln\left(\frac{lev}{1-lev}\right)$. An inverse logit transformation would map leverage from the real line back to the unit interval.

[6]The assumption of a slow-moving explanatory variable is reasonable, since the empirical factors and the underlying capital structure determinants they proxy for are both persistent. Empirically, the persistence of the tangibility ratio is $\phi = 0.95$, for example.

the unconditional mean, while the leverage of low-leverage portfolios increases equally slowly towards the mean. The persistent difference between a high-leverage portfolio and a low-leverage portfolio reflects the persistence in the explanatory variable. Figure 2 illustrates this via a simulation. Leverage is a function of a persistent explanatory variable x_{it}, whose autocorrelation coefficient is $\phi = 0.85$. The persistence in the explanatory variable is clearly reflected in the slow convergence of the leverage portfolios in Panel A: the high-leverage and low-leverage portfolios have not converged to the unconditional mean of 0 after 20 time periods.

[Figure 2 about here.]

If instead of sorting on actual leverage, we sort on unexpected leverage, i.e., on the residuals obtained from a regression of lev_{it} on x_{it}, convergence happens immediately after the sorting period. Since there is no information in the regression residual about future values of the regressor and, hence, leverage, next period's average portfolio leverage drops to its unconditional mean of zero right away, irrespective of the magnitude of the residual that we condition on. To see this analytically, combine (1) with (2) to obtain the following sample regression equation:

$$lev_{it+1} = \beta\phi x_{it} + \beta\epsilon_{it+1} + u_{it+1} \tag{3}$$

The expectation of next period's leverage lev_{it+1}, conditional on this period's estimated regression residual $\hat{u}_{it}$, obtained by running regression (1), is:

$$\begin{aligned} \mathbb{E}[lev_{it+1}|\hat{u}_{it}] &= \beta\phi\mathbb{E}[x_{it}|\hat{u}_{it}] + \beta\mathbb{E}[\epsilon_{it+1}|\hat{u}_{it}] + \mathbb{E}[u_{it+1}|\hat{u}_{it}] \\ &= 0 \end{aligned} \tag{4}$$

since all three expectations on the RHS are equal to zero. $\mathbb{E}[x_{it}|\hat{u}_{it}] = \mathbb{E}[x_{it}] = 0$ follows from the orthogonality of the residuals to the regressor. The second expectation vanishes owing to the independence of ϵ_{it+1} and $\hat{u}_{it}$, while the last expectation equals zero because of the temporal independence of the regression residuals. Thus, under a correctly specified model of leverage, conditioning on the estimated residuals does not produce persistent differences between leverage portfolios.

Panel B of Figure 2 shows the results for sorting on unexpected leverage (the estimated regression residual) instead of on leverage itself. Since the regression is well-specified, today's residual contains no information about tomorrow's leverage, and both portfolios converge to the unconditional mean after one time period.

2.1.2 Case II: Persistence as a Consequence of Measurement Error

Under a correctly specified model of leverage, conditioning on the estimated residuals does not produce persistent differences between leverage portfolios. This is no longer true if we measure a slowly moving explanatory variable with error. To understand the transmission mechanism, I first assume that the regressor x_{it} is not directly observable, but a mismeasured regressor x^*_{it} is:

$$x^*_{it} = x_{it} + \eta_{it} \tag{5}$$

where $\eta_{it} \sim N(0, \sigma^2_{\eta_{it}})$ denotes measurement error. The leverage equation is still given by (1), and the AR(1) process of the explanatory variable is given by (2). The error terms of all three equations (i.e., u_{it}, ϵ_{it} and η_{it}) are independent. If we run a regression of lev_{it} on x^*_{it}, the sample regression equation (* indicates a coefficient or variable that is affected by measurement error, and $\hat{}$ denotes a regression estimate) is:

$$lev_{it} = \hat{\beta}^* x^*_{it} + \hat{u}^*_{it} \tag{6}$$

The estimated slope coefficient of regression equation (6) is biased towards 0 (see Appendix B.1):

$$\hat{\beta}^* = \frac{cov(x^*_{it}, lev_{it})}{\sigma^2_{x^*_{it}}} = \beta \frac{\sigma^2_{x_{it}}}{\sigma^2_{x_{it}} + \sigma^2_{\eta_{it}}} \leq \beta \tag{7}$$

The estimated regression residuals $\hat{u}^*_{it}$ are now biased as well. If we use the residuals $\hat{u}^*_{it}$ to form portfolios at time t and track the leverage of these portfolios over time, next period's expected portfolio leverage is no longer equal to zero (or to the unconditional mean, more generally):

Proposition 1. *Suppose that leverage is determined by* $lev_{it} = \beta x_{it} + u_{it}$*, where* $x_{it} = \phi x_{it-1} + \epsilon_{it}$*, and all noise terms are normally distributed. If we regress leverage on the mismeasured observable variable* $x^*_{it} = x_{it} + \eta_{it}$*, then expected leverage next period, conditional on this period's estimated regression residual* $\hat{u}^*_{it}$*, is a function of the estimated residual:*

$$\mathbb{E}(lev_{it+1} | \hat{u}^*_{it}) = \phi \underbrace{\left[1 + \frac{\sigma^2_{u_{it}}}{\beta^2}\left(\frac{1}{\sigma^2_{\eta_{it}}} + \frac{1}{\sigma^2_{x_{it}}}\right)\right]^{-1}}_{=c\geq 0} \hat{u}^*_{it} \tag{8}$$

Proof. See Appendix B.2. □

Equation (8) shows that next period's expected leverage is directly linked to this period's regression residual via the coefficient c. Its sign is positive, which implies that the expected leverage, conditional on a positive residual, will overstate the true expected leverage (and understate true expected leverage for a negative residual). This creates an artificial leverage dispersion when we track leverage portfolios. The rate at which the dispersion disappears is directly governed by the coefficient ϕ, the persistence in the underlying latent explanatory variable.

The link between the regression residual and expected leverage is that the mismeasured residual now contains information about the magnitude of the true explanatory variable x_{it}, which in turn determines leverage. Recall the general expression for expected portfolio leverage, conditional on sorting on the regression residual:

$$\mathbb{E}[lev_{it+1}|\hat{u}_{it}^*] = \beta\phi\mathbb{E}[x_{it}|\hat{u}_{it}^*] + \beta\mathbb{E}[\epsilon_{it+1}|\hat{u}_{it}^*] + \mathbb{E}[u_{it+1}|\hat{u}_{it}^*] \tag{9}$$

As under the no-measurement-error scenario, the second and third expectations on the RHS are still equal to zero. The first expectation on the RHS, however, is no longer equal to 0:

Lemma 1. *In the setup described in Proposition 1, the expectation of the regressor, conditional on the estimated regression residual $\hat{u}_{it}^*$, is:*

$$\begin{aligned}
\mathbb{E}(x_{it}|\hat{u}_{it}^*) &= \mathbb{E}(x_{it}) + \frac{Cov(x_{it}, \hat{u}_{it}^*)}{Var(\hat{u}_{it}^*)}\left[\hat{u}_{it}^* - \mathbb{E}(\hat{u}_{it}^*)\right] \\
&= \underbrace{\left[\beta + \frac{\sigma_{u_{it}}^2}{\beta}\left(\frac{1}{\sigma_{\eta_{it}}^2} + \frac{1}{\sigma_{x_{it}}^2}\right)\right]^{-1}}_{=b} \hat{u}_{it}^* \qquad (10)
\end{aligned}$$

Proof. See Appendix B.2, beginning with (43). □

This expression relates the expectation of x_{it}, conditional on an estimated residual $\hat{u}_{it}^*$, to the true parameters of the underlying processes, which are captured in the coefficient b. Importantly, knowing a particular value of $\hat{u}_{it}^*$ tells us something about the value of the true x_{it}. This is because the estimated residual is *not* orthogonal to the true regressor, i.e., $\mathbb{E}(x_{it}|\hat{u}_{it}^*) \neq \mathbb{E}(x_{it})$, unlike in the setup without measurement error. With a latent explanatory variable, if the true relationship between lev_{it} and x_{it} is positive (i.e. $\beta > 0$), then a larger residual $\hat{u}_{it}^*$ predicts a true x_{it} that is above its unconditional mean. Conversely, if $\beta < 0$, then a larger residual $\hat{u}_{it}^*$ predicts a true x_{it} that is below its unconditional mean.

The more mismeasured the regressor is, the more persistent are the residual-based portfolio leverage levels. In Figure 3, I illustrate the effect of measurement error when the value of the AR(1) coefficient of the regressor is $\phi = 0.85$. The figure plots the relationship between portfolio-leverage dispersion and the magnitude of measurement error, when firms are sorted based on regression residuals. I include two lines for reference: the solid line shows the sort based on leverage itself, while the dotted line shows a residual-based sort without measurement error. In the latter case, the portfolios collapse to the unconditional mean immediately after the sorting period, as discussed previously. The dashed lines show sorts for two levels of the measurement error: $\sigma_\eta \in \{0.5, 1\}$. The ratio of measurement noise to state noise in the regressor is thus also $\sigma_\eta/\sigma_\epsilon \in \{0.5, 1\}$. The larger the quantity of measurement error is, the more the residual-based sorts start to resemble leverage-based sorts. At the higher level of measurement error, the dispersion in portfolio leverage is about 50% of the dispersion when sorting is done on leverage itself.

[Figure 3 about here.]

Lemma 2. *In the setup described in Proposition 1, the estimated regression residual* $\hat{u}^*_{it}$ *will exhibit persistence:*

$$\mathbb{E}(\hat{u}^*_{it+1}|\hat{u}^*_{it}) = \phi(\beta - \hat{\beta}^*)\mathbb{E}(x_{it}|\hat{u}^*_{it}) = \left[\phi(\beta - \hat{\beta}^*)b\right] \hat{u}^*_{it} \tag{11}$$

where b *is as defined in Lemma 1.*

Proof. See Appendix B.3. □

Lemma 2 shows that when a persistent regressor is mismeasured, the estimated regression residual itself will exhibit persistence. This persistence can be linked to the explanatory power of a firm fixed effect: increasing persistence (via a higher value of ϕ) and a larger attenuation bias in the cross-sectional β coefficient will increase the explanatory power of a firm fixed effect. This occurs because the estimated firm fixed effect loads on the persistent error term.

3 Extracting Measurement Error from Explanatory Variables: Calibration

I now turn to the important question of how much measurement error is needed to reproduce both the wide initial dispersion between the residual-based portfolios, and the slow post-sort convergence evident in the data. To assess whether a reasonably calibrated model with measurement error in the explanatory variables can satisfactorily explain the data, I employ two different approaches: in the first approach, described in Sections 3.1 to 3.2, I use the portfolio sorts based on actual leverage as the starting point. If leverage is determined cross-sectionally by $lev_{it} = \beta x_{it} + u_{it}$, then in every time period, a firm's leverage is equal to target leverage βx_{it} plus an error term u_{it}. Therefore, the leverage time series of the portfolios based on *actual* leverage display the same dynamics as the true leverage target, and thus can be used to infer the target's law of motion. Furthermore, if the regressions underpinning the *residual*-based sorts correctly identified this true target, the leverage levels of the residual-sorted portfolios should converge to the unconditional mean immediately. Since they do not, I use the portfolio-leverage time series for the residual-based sorts to establish how mismeasured the leverage target (and thus its determinant) needs to be in order to be consistent with the residual-based sorts.

The second approach to quantifying the amount of measurement error needed to reproduce persistent leverage in portfolio sorts is outlined in Section 3.3. There, I examine four actual explanatory variables consistent with the Lemmon, Roberts, and Zender (2008) study, and determine how mismeasured each of these needs to be in order to be consistent with the portfolio sorts, as well as with other observed data moments.

3.1 Estimating Target-Leverage Dynamics

The first approach consists of a two-step method: Step 1 parameterizes the law of motion for a firm's target leverage. Step 2 (see Section 3.2) uses this law of motion in conjunction with residual-based portfolio sorts to arrive at an estimate of measurement error.

To start, consider again the setup from Section 2.1, where leverage lev_{it} is a function of a slow-moving factor x_{it}. This factor evolves according to an AR(1) process, but the true realizations of the process are latent. The observed values x^*_{it} contain *iid* measurement-error terms η_{it}:

$$
\begin{aligned}
lev_{it} &= \beta x_{it} + u_{it} && (12)\\
x_{it} &= \phi_0 + \phi_1 x_{it-1} + \epsilon_{it} && (13)\\
x^*_{it} &= x_{it} + \eta_{it} && (14)
\end{aligned}
$$

where $u_{it} \sim N(0, \sigma^2_u)$, $\epsilon_{it} \sim N(0, \sigma^2_\epsilon)$, and $\eta_{it} \sim N(0, \sigma^2_\eta)$. An intercept ϕ_0 is included in the AR(1) process for the leverage determinant to allow for a positive leverage mean, as in the data (see Figure 1 for the mean leverage time series of the four leverage portfolios).

In the cross-sectional specification in equation (12), actual leverage lev_{it} can be viewed as the sum of two components: a leverage target $\widehat{lev_{it}} \equiv \beta x_{it} = \mathbb{E}[lev_{it}|x_{it}]$, which the firm adjusts to every period, and a random deviation from the target u_{it}. Implicit in this representation is the assumption that there are no adjustment costs that would cause the firm to deviate systematically from its target for multiple periods. While the target in (12) is determined by only a single variable, using the target-leverage representation does allow the flexibility of viewing the target as a function of potentially many explanatory variables, so the above setup of only one explanatory factor does not result in a loss of generality. Substituting target leverage in (12) and (13) above gives the following system:

$$
\begin{aligned}
lev_{it} &= \widehat{lev}_{it} + u_{it} && (15)\\
\widehat{lev}_{it} &= \varphi_0 + \varphi_1 \widehat{lev}_{it-1} + \varepsilon_{it} && (16)\\
x^*_{it} &= x_{it} + \eta_{it} && (17)
\end{aligned}
$$

The law of motion for target leverage $\widehat{lev}_{it}$ in (16) is the same as that for the original factor[7] in (13), scaled by the constant β.

To estimate the parameter values in (15) to (17), I proceed as follows: in the first step, I use the portfolio-leverage time series (sorted on *actual* leverage) to parameterize (15) and (16). After estimating target-leverage dynamics, I then determine how mismeasured (by virtue of mismeasuring the underlying factors) the target needs to be in the cross-sectional regressions for the patterns in the *residual-based* leverage sorts to obtain. Without loss of generality, I simplify the analysis by examining two rather than four portfolios: a "high-leverage" portfolio and a "low-leverage" portfolio. This does not affect the results, since initial convergence and long-term persistence are still evident with only two portfolios.

Using the actual leverage-based portfolios, I estimate four parameters in equations (15) and (16): the cross-sectional error variance σ_u^2 in the leverage equation, and the intercept φ_0, slope coefficient φ_1 and error variance σ_ε^2 in the equation for the AR(1) process governing target leverage. I simulate equations (15) and (16) above for both realized leverage and the target, and then find parameter values that minimize the sum of the squared differences between actual portfolio leverage and simulated portfolio leverage, i.e.,

$$\min_{\Phi} \sum_i \sum_t \left(PFlev_{it}^{sim} - PFlev_{it}^{act}\right)^2 \tag{18}$$

where the parameter vector $\Phi = \{\sigma_u^2, \varphi_0, \varphi_1, \sigma_\varepsilon^2\}$, and $PFlev_{it}$ denotes the leverage of portfolio i (i indexes high and low leverage) at time t. The parameter estimates are as follows:

	φ_0	φ_1	σ_ε	σ_u
Estimate	0.021	0.930	0.066	0.080
Std. Error	(0.012)	(0.009)	(0.003)	(0.010)

[7]If more than one explanatory factor is included, the target dynamics can be thought of as a linear combination of AR(1) processes, which would result in an ARMA representation for the target (see e.g., Granger and Newbold (1977)).

The estimated coefficients are of reasonable magnitude, in line with what a pooled regression would yield. In addition, simulating the actual leverage-based portfolio sorts using the parameter values above provides a good fit to the real data, as shown in Figure 4.

[Figure 4 about here.]

3.2 Estimating Measurement Error by Extracting the Mismeasured Target

Having estimated the true leverage target $\widehat{lev_{it}} = \beta x_{it}$ consistent with portfolios sorted on actual leverage, I now extract a *mismeasured* leverage target $\widehat{lev^*_{it}}$ consistent with the *residual-based* portfolios. The mismeasured target allows us to compute the mismeasured residuals u^*_{it}, which form the basis of the residual-based portfolio sorts because

$$u^*_{it} = lev_{it} - \widehat{lev^*_{it}} \tag{19}$$

Using a noisy determinant in the regression implies that the target leverage (i.e., the regression's predicted leverage value) is also mismeasured. As mentioned before, I avoid explicitly modelling an explanatory variable x_{it}, or x^*_{it} in its mismeasured form, but focus on the target instead. This approach highlights the intuition behind the results.[8] It is possible to recover the mismeasured target leverage $\widehat{lev^*_{it}}$, because we can express it as a function of the true target:

Proposition 2. *Suppose that leverage dynamics are given by equations (12) through (14). Using a mismeasured explanatory variable x^*_{it} in the cross-sectional leverage regression will cause target leverage $\widehat{lev^*_{it}} = \mathbb{E}(lev^*_{it}|x^*_{it})$ (the fitted regression value) to be mismeasured as well. If the mismeasured target is expressed in terms of the true target $\widehat{lev_{it}}$ by the following regression:*

$$\widehat{lev^*_{it}} = \alpha_0 + \alpha_1 \widehat{lev_{it}} + e_{it} \tag{20}$$

then the parameters in (20) are functions of known data moments and the noise-to-

[8] The target-leverage substitution also improves identification, since no cross-sectional β-coefficient needs to be estimated.

signal ratio a of the mismeasured x^*_{it}:

$$\alpha_0 = (1-\alpha_1)\mathbb{E}(\widehat{lev_{it}}) \tag{21}$$

$$\alpha_1 = \frac{1}{1+a} \tag{22}$$

$$\sigma_e^2 = Var(\widehat{lev_{it}})\frac{a}{(1+a)^2} \tag{23}$$

$$a = \frac{\sigma_\eta^2}{\sigma_x^2} \tag{24}$$

Proof. See Appendix C. □

Proposition 2 states that knowledge of the true target dynamics, via the methodology in Section 3.1, permits an explicit solution for the mismeasured target leverage in (20). The unknown parameters α_0, α_1 and σ_e^2 are functions of known data moments and a given ratio of measurement noise to cross-sectional variation ${}^{\sigma_\eta^2}/_{\sigma_x^2} = a$. This ratio can thus be used to indirectly quantify the amount of measurement error in equation (14), and also allows for an explicit solution for the parameters in Proposition 2.

An observation about Proposition 2 is warranted: if measurement error is present ($\sigma_\eta^2 \geq 0$), then the variance of the mismeasured target is less than the variance of the true target, i.e., $Var(\widehat{lev^*_{it}}) \leq Var(\widehat{lev_{it}})$ (see (64) in Appendix C for a proof). The larger the amount of measurement error in the underlying leverage determinant, the smaller the variation in estimated target leverage will be. In a univariate regression, the attenuation in the estimated slope coefficient rises with the amount of measurement error in the explanatory variable. This naturally results in a larger estimated intercept, which makes intuitive sense: the "best" predicted value of the dependent variable in the presence of an increasingly noisy explanatory variable simply approaches the dependent variable's unconditional mean, which does not vary over time.

This reasoning translates directly to the relationship between the estimated mismeasured target leverage and the true target, as given by (20). If the target were perfectly measured, then $\alpha_0 = 0$, $\alpha_1 = 1$ and $\sigma_e^2 = 0$. As the measurement noise in the observed explanatory variable increases, the mismeasured target will become more stable relative to the true target: $\alpha_0 > 0$ and $\alpha_1 < 1$, while σ_e^2 will increase at first and then decrease again. In the limit, with the signal-to-noise ratio of x^*_{it} approaching 0, the mismeasured target is constant with $\alpha_0 = \mathbb{E}(lev)$, $\alpha_1 = 0$, and $\sigma_e^2 = 0$.

While the mismeasured target will always be less variable than the true target, it still equals the true target, on average: $\mathbb{E}(\widehat{lev^*}) = \mathbb{E}(\widehat{lev})$ (see (56) in Appendix C for a proof). Intuitively, this is due to regression mechanics: the mean predicted value will equal the dependent variable's unconditional mean, regardless of whether there is measurement error in the explanatory variable. Naturally, this only holds in an unconditional sense; if a given true x_{it} is above its unconditional mean, the mismeasured target will underestimate the true target and vice versa. Figure 5 illustrates this point for various levels of the noise-to-signal ratio a. As a increases, the effect becomes more visible. For instance, in Panel D with $a = 1.25$, when true target leverage is below its unconditional mean of 0.27, the mismeasured target tends to be larger than the true target, i.e., it is closer to the unconditional mean of the leverage variable.

[Figure 5 about here.]

I next recover the implied ratio a of measurement noise to cross-sectional variation in the explanatory variable that minimizes the sum of the squared differences between the simulated and actual portfolio-leverage levels, sorted on mismeasured residuals:

$$\min_a \sum_i \sum_t \left(PFlev_{it}^{sim} - PFlev_{it}^{act}\right)^2 \tag{25}$$

$PFlev_{it}$ denotes the average leverage of portfolio i, where i indexes high and low leverage at time t. To do the portfolio sorts embedded in the above minimization, I first compute the mismeasured target $\widehat{lev_{it}^*}$ from the true target $\widehat{lev_{it}}$ via (20), and then solve equation (19) for the residual u_{it}^*. Figure 6 shows the results of this minimization.

[Figure 6 about here.]

The simulated residual-based portfolio sorts most closely match the empirical ones with a noise-to-signal ratio of $a = 1.42$ (std. error = 0.12), i.e., the variance of the measurement error needs to be 42% larger than the cross-sectional variation of the true but unobserved explanatory variable x. Clearly, this amount of measurement error seems large, but the sole factor x in this setup serves as a stand-in for all the determinants of leverage. In a multivariate world, high levels of measurement error in one variable can counterbalance low levels of measurement error in another.

Another consideration is that in the previous calibration, each portfolio received an equal weighting. Weighting some observations more heavily than others would also reduce the value of a. Furthermore, while $a = 1.42$ results in the best fit, even small quantities of measurement error relative to the variance of the explanatory variable produce a surprising amount of persistence in the residual-based sorts. Figure 7 shows portfolio sorts at different levels of a. For instance, Panel D shows that a noise-to-signal ratio as low as 0.75 still produces a good fit to the actual portfolios in year 5 and beyond. While large quantities of measurement error are needed to completely reproduce the stylized facts, much more moderate levels still produce a notable amount of leverage persistence in the sorts. This reinforces the view that measurement error is a contributing factor to persistence in residual-based portfolio sorts.

[Figure 7 about here.]

3.3 Multi-Variable Calibration with *iid* Measurement Error

I now turn to whether measurement error in explanatory variables similar to those used by Lemmon, Roberts, and Zender (2008) is able to reproduce the leverage time series of both the actual- and residual-based portfolio sorts. I focus on profitability; tangibility (as a measure of how tangible a firm's collateral assets are); the market-to-book ratio (as a proxy for investment opportunities); and industry leverage (as a measure of industry-specific leverage targets, in lieu of an industry fixed effect). Firm size is excluded, since it is not stationary and thus would not conform to my setup of modelling the explanatory variables as AR(1) processes.

The estimation by simulated method of moments proceeds in a similar fashion to that in Section 3.2. In particular, the economy consists of simulated firms whose leverage dynamics are governed by the following system of equations:

$$
\begin{aligned}
lev_{it} &= \boldsymbol{\beta}'(1 \;\; \boldsymbol{x_{it}})' + u_{it} && (26)\\
&= \begin{pmatrix}\beta_0 & \beta_{Prof} & \beta_{Tang} & \beta_{MB} & \beta_{IndLev}\end{pmatrix}\begin{pmatrix}1\\ Prof_{it}\\ Tang_{it}\\ MB_{it}\\ IndLev_{it}\end{pmatrix} + u_{it} && (27)\\
\boldsymbol{x_{it}} &= \boldsymbol{\phi_0} + \boldsymbol{\phi_1}\boldsymbol{x_{it-1}} + \boldsymbol{\epsilon_{it}} && (28)\\
&= \begin{pmatrix}\phi_0^{Prof}\\ \phi_0^{Tang}\\ \phi_0^{MB}\\ \phi_0^{IndLev}\end{pmatrix} + \begin{pmatrix}\phi_1^{Prof} & 0 & 0 & 0\\ 0 & \phi_1^{Tang} & 0 & 0\\ 0 & 0 & \phi_1^{MB} & 0\\ 0 & 0 & 0 & \phi_1^{IndLev}\end{pmatrix}\boldsymbol{x_{it-1}} + \begin{pmatrix}\epsilon_{it}^{Prof}\\ \epsilon_{it}^{Tang}\\ \epsilon_{it}^{MB}\\ \epsilon_{it}^{IndLev}\end{pmatrix} && (29)\\
\boldsymbol{x^*_{it}} &= \boldsymbol{x_{it}} + \boldsymbol{\eta_{it}} && (30)
\end{aligned}
$$

The errors are all normally distributed with $u_{it} \sim N(0, \sigma_u^2)$, $\boldsymbol{\epsilon_{it}} \sim N(0, \boldsymbol{\Sigma_\epsilon})$, and $\boldsymbol{\eta_{it}} \sim N(0, \boldsymbol{\Sigma_\eta})$.[9] Leverage is determined in the cross-section by an intercept and the

[9]Vectors and matrices are denoted by bold letters.

four explanatory factors, which are all modelled as AR(1) processes. Firms differ in terms of the realization of a particular variable, but the coefficients in the model are the same for all firms. The true explanatory variable vector $\boldsymbol{x_{it}}$ is latent; the observable $\boldsymbol{x^*_{it}}$ is measured with error η_{it}: The explanatory variables are imperfect proxies for the true economic fundamentals driving leverage.

The covariance matrix of the innovations of the AR(1) processes $\boldsymbol{\Sigma_\epsilon}$ is diagonal, as is the covariance matrix of the measurement-error terms $\boldsymbol{\Sigma_\eta}$:

$$\boldsymbol{\Sigma_\epsilon} = \begin{pmatrix} \sigma^2_{\epsilon_{Prof}} & 0 & 0 & 0 \\ 0 & \sigma^2_{\epsilon_{Tang}} & 0 & 0 \\ 0 & 0 & \sigma^2_{\epsilon_{MB}} & 0 \\ 0 & 0 & 0 & \sigma^2_{\epsilon_{IndLev}} \end{pmatrix} \tag{31}$$

$$\boldsymbol{\Sigma_\eta} = \begin{pmatrix} \sigma^2_{\eta_{Prof}} & 0 & 0 & 0 \\ 0 & \sigma^2_{\eta_{Tang}} & 0 & 0 \\ 0 & 0 & \sigma^2_{\eta_{MB}} & 0 \\ 0 & 0 & 0 & \sigma^2_{\eta_{IndLev}} \end{pmatrix} \tag{32}$$

There are a total of 22 unknown parameters in this formulation: the intercepts, slopes, and error variances of the AR(1) process (12 parameters), the cross-sectional betas and the error variance σ^2_u (6 parameters), and the measurement-error variances (4 parameters).

To reduce the number of free parameters in the model, the unconditional means of the noisy explanatory variables are inferred directly from the data. This is possible since mismeasured and latent explanatory variables have the same mean: $\boldsymbol{\mu_{x^*}} = \mathbb{E}(\boldsymbol{x^*_{it}}) = \mathbb{E}(\boldsymbol{x_{it}} + \boldsymbol{\eta_{it}}) = \mathbb{E}(\boldsymbol{x_{it}}) = \boldsymbol{\mu_x}$. This allows me to express the intercepts of the latent AR(1) processes as functions of the empirical means of the respective variables and estimates of $\boldsymbol{\phi_1}$, which is a free parameter matrix:

$$\boldsymbol{\phi_0} = (\boldsymbol{I_4} - \boldsymbol{\phi_1})\,\boldsymbol{\mu_x} \tag{33}$$

$$= (\boldsymbol{I_4} - \boldsymbol{\phi_1})\,\boldsymbol{\mu_{x^*}} \tag{34}$$

$\boldsymbol{I_4}$ denotes a 4×4 identity matrix. In fact, several other parameters could be inferred directly from the variance of leverage $Var(lev)$, and from the variance matrix of the noisy explanatory variables $\boldsymbol{\Sigma_{x^*}}$.[10] However, forcing the constraints that the model imposes on these parameters to hold exactly is too restrictive and results in a poor fit. Instead, the variances are added as moment conditions, which results in the simulated values being close to the data values without the need to match them exactly.

[10]We could relate the variance matrix for the AR(1) innovations $\boldsymbol{\Sigma_\epsilon}$ to the variance of the noisy

3.3.1 Identification

To reduce the dimensionality of the parameter space, I calculate the intercepts for the autoregressive processes directly from the data via (34). This pares down the free structural parameters to a total of 18: the matrix $\boldsymbol{\phi_1}$, which contains the slope coefficients for the explanatory variables, the innovation standard deviation matrix $\boldsymbol{\Sigma_\epsilon}$, and the measurement-error variance matrix $\boldsymbol{\Sigma_\eta}$. Furthermore, the parameter vector $\boldsymbol{\beta}$, which governs the cross-sectional relationship between leverage and its determinants, along with the standard deviation of the cross-sectional residual σ_u, has to be estimated.

The structural parameters underlying the latent processes are obtained by matching simulated sample moments to data moments. Broadly speaking, the data moments consist of sample statistics for leverage and the explanatory variables, panel and time-series regression parameters, and the portfolio-leverage levels of the Lemmon, Roberts, and Zender (2008) portfolio sorts. Since I assume that the actual data on explanatory variables are contaminated by measurement error, all data moments involving explanatory variables are mismeasured as well. In particular, I use the following moments:

i. The intercepts ϕ_0^* and slope coefficients ϕ_1^* for each explanatory variable (i.e., profitability, tangibility, market-to-book and industry leverage), which are obtained by regressing each observed mismeasured explanatory variable on its lagged value (8 moments):

$$x_{it}^* = \phi_0^* + \phi_1^* x_{it-1}^* + \epsilon_{it}^* \tag{37}$$

ii. The variance of each mismeasured explanatory variable σ_x^{2*}, and the variance of leverage σ_{lev}^2 (5 moments).

iii. The cross-sectional coefficients $\boldsymbol{\beta^*}$ from a regression of leverage on the noisy determinants (5 moments):

$$\begin{aligned} lev_{it} &= \boldsymbol{\beta^{*\prime}}(1 \;\; \boldsymbol{x_{it}^{*\prime}}) + u_{it}^* \qquad \text{where} && (38)\\ \boldsymbol{\beta^{*\prime}} &= (\beta_0^* \;\; \beta_{Prof}^* \;\; \beta_{Tang}^* \;\; \beta_{MB}^* \;\; \beta_{IndLev}^*) && (39)\end{aligned}$$

and $\boldsymbol{x_{it}^*}$ is the vector of mismeasured explanatory variables.

---

factors $\boldsymbol{\Sigma_{x^*}}$, and the variance of the regression residual σ_u^2 to the variance of leverage $Var(lev)$ via:

$$\begin{aligned} \boldsymbol{\Sigma_\epsilon} &= \left(\boldsymbol{I_4} - \boldsymbol{\phi_1}'\boldsymbol{\phi_1}\right) \; \boldsymbol{\Sigma_x} \\ &= \left(\boldsymbol{I_4} - \boldsymbol{\phi_1}'\boldsymbol{\phi_1}\right)\left(\boldsymbol{\Sigma_{x^*}} - \boldsymbol{\Sigma_\eta}\right) && (35)\\ \sigma_u^2 &= Var(lev) - \boldsymbol{\beta\Sigma_x\beta}' && (36)\end{aligned}$$

iv. The time series of portfolio-leverage levels after sorting on both actual and unexpected leverage (80 moments in total). A time series consists of 20 portfolio-leverage levels for each "high-leverage" and "low-leverage" portfolio.

For both actual and simulated data, the moments are collected in vectors $\boldsymbol{m^{act}}$ and $\boldsymbol{m^{sim}}$, respectively. The structural parameters collected in the vector $\Phi = (\boldsymbol{\phi_1} \;\; \boldsymbol{\Sigma_\epsilon} \;\; \boldsymbol{\Sigma_\eta} \;\; \boldsymbol{\beta} \;\; \sigma_u)$ are found by minimizing the sum of the squared differences between actual moments and simulated moments:

$$\min_{\Phi} \; (\boldsymbol{m^{act}} - \boldsymbol{m^{sim}})'(\boldsymbol{m^{act}} - \boldsymbol{m^{sim}}) \tag{40}$$

This minimization makes the simulated moments as close to their actual counterparts as possible by picking the "best" structural parameter values.

3.3.2 Results

The estimated structural parameters of this procedure, along with their standard errors,[11] are listed in Table 2. Table 3 presents a comparison of empirical data moments, their simulated counterparts based on mismeasured variables, and moments that are based on the estimated true latent parameters. Table 4 gives two estimates of the ratio of measurement noise to state noise for each simulated explanatory variable. The first estimate is the ratio of the measurement-error variance to the variance of the latent underlying variable, while the second estimate is the ratio of the measurement-error variance to the variance of the observed variable, which thus includes the measurement-error variance in the denominator. Finally, Figure 8 shows the portfolio sorts on actual and residual leverage, which are obtained with the estimated parameter values.

[Table 2 about here.]

[Table 3 about here.]

[Table 4 about here.]

[Figure 8 about here.]

For both the tangibility and industry-leverage ratios, the calibrated values of the latent processes are very close to the empirical data values. As measured by the AR(1)

[11]The standard errors are bootstrapped: First, all empirical moments are recalculated for subsets of the Compustat universe. I then estimate structural parameters for each of the subsamples. The standard errors are given by the standard deviations of the estimated structural parameters.

parameter and shown in Table 3, the estimated persistence for tangibility is 0.936 (empirical data value of 0.952), while it is 0.891 for industry leverage (empirical data value of 0.908). The estimated magnitude of the measurement-error standard deviation σ_η is small in both instances, and well below the standard deviation of the innovation σ_ϵ in the respective AR(1) process (see Table 2). This results in a ratio of measurement-error variance to latent-variable variance σ_η^2/σ_x^2 of 0.021 for tangibility and 0.018 for industry leverage (see Table 4, column (1)). Similar values for the measurement-error ratio are obtained if the variance of the observed explanatory variable is used instead. Consistent with the small magnitude of the estimated measurement error terms, the structural β-coefficients for both variables are close to their empirical counterparts (see Table 3).

Table 3 shows that latent profitability ($\phi_1 = 0.832$, see "Struc. Value" column) is more persistent than observed profitability ($\phi_1^* = 0.775$). The depressed observed ϕ_1^* coefficient is caused by measurement error in observed profitability with an estimated standard deviation of $\sigma_\eta = 0.105$ (see Table 2), which also induces a slight downward bias in the cross-sectional β^*. Relative to tangibility and industry leverage, the measurement error ratios for profitability have increased to 0.090 and 0.083, respectively (see Table 4). These values are still low; for example, the latter implies that only 8.3% of the variation in observed profitability is due to measurement error.

The most interesting result obtains for the market-to-book ratio. The latent AR(1) process has an estimated value of $\phi_1 = 0.931$, while the empirical process has a value of $\phi_1^* = 0.534$ (see Table 3). The simulated ϕ_1^* value, obtained by regressing the simulated mismeasured market-to-book ratio on its lagged value, is 0.530, which is very close to the empirical estimate. The discrepancy between latent and observed ϕ_1 is caused by a measurement-error standard deviation that is large when compared with that for the other variables. Its value is $\sigma_\eta = 1.476$, which exceeds the standard deviation of the innovation term in the AR(1) process $\sigma_\epsilon = 0.603$, as shown in Table 2. The resulting measurement-error ratio is $\sigma_\eta^2/\sigma_x^2 = 0.802$, which drops to $\sigma_\eta^2/\sigma_{x^*}^2 = 0.445$ if we use the variance of the observed market-to-book ratio in the denominator (see Table 4). This latter value implies that 44.5% of the observed variation in the market-to-book ratio is driven by noise. While this seems large, the market-to-book ratio as a proxy for investment opportunities can, ex ante, be expected to be noisy. Erickson and Whited (2006) state that "all observable measures or estimates of the true incentive to invest [...] are likely to contain measurement error." Using a classical errors-in-variables model with the investment-to-capital ratio on the LHS and average q on the RHS, Erickson and Whited (2006) report that approximately 59% of the variation in book-value-based measures of Tobin's q is driven by noise, and only 41% is driven by variation in the true unobservable q. This is consistent with my results, where 55% of the variation in the market-to-book ratio is due to variation in true q.

My estimates of the structural parameters produce a variance in the observed market-to-book ratio of 4.895, which is equal to its empirical counterpart. Thus, the results are not driven by an unnaturally high total variance in the market-to-book ratio. In the simulated cross-section, the true latent β-coefficient for the market-to-book ratio is -0.105, which is larger than the empirical value of -0.006 (Table 3). The simulated mismeasured observed value for β_{MB} is -0.058. My results suggest that a market-to-book ratio, which is a poor proxy for true investment opportunities, plays an important role in the persistence of the residual-based portfolio sorts. Since an option to invest is riskier than the investment itself, firms with a high true q would optimally choose to carry lower amounts of leverage. However, this effect is obscured in the data owing to the high amount of measurement error inherent in the market-to-book ratio.

Overall, the estimation produces sensible parameter values, and the simulated moments closely resemble their empirical data counterparts, as a comparison of the "Data Value" and "Sim. Value" columns in Table 3 reveals. Finally, Figure 8 shows the results of the portfolio sorts. Using the estimated values of the structural parameters in Table 2 produces a close fit between empirical and simulated portfolio leverage time series, regardless of whether the sort is done on actual or residual leverage. While the simulated residual-based portfolios exhibit less dispersion than their empirical counterparts in years two to five, they track the empirical time series closely in the other time periods. This shows that low levels of measurement error in profitability, size and industry leverage, coupled with a larger, yet realistic, amount of measurement error inherent in using book-value-based proxies of Tobin's q offers a potential explanation of the documented persistence in leverage portfolios.[12]

4 Conclusion

Persistence in residual-based leverage portfolios is a well-documented fact. While this persistence can result from the omission of either a firm fixed effect or time-varying variables, I show that it also arises when slow-moving explanatory variables in a leverage regression are measured with error. Sorting firms into portfolios based on these regression residuals will exhibit similar portfolio-leverage persistence as sorting firms into portfolios based on actual leverage.

[12]In unreported results, model fit improves by allowing for a slight autocorrelation in the measurement-error terms themselves.

Being able to predict future leverage with the regression residuals implies that target leverage is mismeasured. I find that if the leverage target is modelled as being determined by a single composite factor of a number of possible trade-off theory variables, then the measurement-error variance of this latent factor needs to be 142% of its cross-sectional variance to reproduce the stylized empirical facts. This number is large, but is nonetheless a useful measure, since it can be interpreted as an aggregate estimate of how mismeasured the explanatory variables would need to be. However, even much lower amounts of measurement error still produce remarkably persistent residual-based portfolio sorts. Therefore, even if measurement error alone is not sufficient to fully account for the persistence of leverage in the setting of regression-residual-based portfolio sorts, it is nonetheless likely to be an important contributor.

I also examine measurement error in several important explanatory variables, namely the firm's profitability, the tangibility of its assets, the market-to-book ratio, and industry leverage. I find that low quantities of measurement error in profitability, tangibility and industry leverage, coupled with a measurement-error variance equal to about 80% of the cross-sectional variation in the market-to-book ratio, produce a good fit of simulated sample data moments to empirical moments. This level of measurement error in the market-to-book variable, which proxies for Tobin's q, is consistent with other studies such as Erickson and Whited (2006), and suggests that unobserved investment opportunities play an important role in explaining leverage ratios, and, hence, in the persistence of the residual-based portfolio sorts.

The focus of this paper is on capital structure. However, portfolio sorts are also a popular tool to evaluate the returns from trading strategies, and to test asset pricing models. Measurement quality is an important consideration for the risk factors in these models, so my work also has implications for the asset pricing applications of portfolio sorts.

Appendices

A Variable Definitions

Data are taken from the annual Compustat database between 1965 and 2003. Financial firms and companies with missing asset or debt values are excluded from the sample. Leverage is constrained to lie in the closed unit interval. Size, profitability, tangibility and the market-to-book ratio are winsorized at the 1^{st} and 99^{th} percentiles. The construction of each variable is as follows:

$$
\begin{aligned}
\text{Leverage} &= \frac{\text{Short-Term Debt [34] + Long-Term Debt [9]}}{\text{Book Assets [6]}} \\
\text{Total Debt} &= \text{Short-Term Debt + Long-Term Debt} \\
\text{Size} &= ln(\text{Book Assets [6]}) \\
\text{Profitability} &= \frac{\text{Operating Income before Depreciation [13]}}{\text{Book Assets [6]}} \\
\text{Tangibility} &= \frac{\text{PPE [13]}}{\text{Book Assets [6]}} \\
\text{Market Equity} &= \text{Share Price [199] * Shares Outstanding [54]} \\
\text{Market-to-Book} &= \frac{\text{Market Equity + Total Debt + Pref. Stock Liq. Value [10] - Def. Taxes [35]}}{\text{Book Assets [6]}}
\end{aligned}
$$

B Derivations

B.1 Attenuation Bias

Intermediate Steps:

$$\begin{aligned}
\hat{\beta}^* &= \frac{cov(x^*_{it}, lev_{it})}{\sigma^2_{x^*_{it}}} = \frac{\mathbb{E}\left[(x_{it}+\eta_{it})(\beta x_{it}+u_{it})\right] - \mathbb{E}(x_{it}+\eta_{it})\mathbb{E}(\beta x_{it}+u_{it})}{\mathbb{E}[(x_{it}+\eta_{it})^2] - \mathbb{E}(x_{it}+\eta_{it})^2} \\
&= \frac{\beta\mathbb{E}(x^2_{it}) - \beta\mathbb{E}(x_{it})^2}{\mathbb{E}(x^2_{it}) + \mathbb{E}\eta^2_{it} - \mathbb{E}(x_{it})^2} \\
&= \beta\frac{\sigma^2_{x_{it}}}{\sigma^2_{x_{it}} + \sigma^2_{\eta_{it}}} \qquad (41)
\end{aligned}$$

B.2 Conditional Expectation of Leverage Under Measurement Error

Expected portfolio leverage, conditional on sorting on the mismeasured regression residual, is:

$$\mathbb{E}[lev_{it+1}|\hat{u}^*_{it}] = \beta\phi\mathbb{E}[x_{it}|\hat{u}^*_{it}] + \beta\mathbb{E}[\epsilon_{it+1}|\hat{u}^*_{it}] + \mathbb{E}[u_{it+1}|\hat{u}^*_{it}] \qquad (42)$$

The second and third expectations on the RHS are equal to zero. $\mathbb{E}[\epsilon_{it+1}|\hat{u}^*_{it}] = 0$ since next period's innovation in the explanatory variable is independent of this year's estimated residual. Similarly, next period's true residual in the leverage regression is independent of this period's estimated residual, so $\mathbb{E}[u_{it+1}|\hat{u}^*_{it}] = 0$. The first expectation on the RHS is not equal to 0, however. The residual $\hat{u}^*_{it}$ contains information about the true x_{it}, so $\mathbb{E}(x_{it}|\hat{u}^*_{it}) \neq \mathbb{E}(x_{it})$. To see this, assume that x_{it} and $\hat{u}^*_{it}$ are normally distributed random variables. Start with a scalar version of the conditional expectation of multivariate normal random variables: [13]

$$\mathbb{E}(x_{it}|\hat{u}^*_{it}) = \mathbb{E}(x_{it}) + \frac{Cov(x_{it}, \hat{u}^*_{it})}{Var(\hat{u}^*_{it})}\left[\hat{u}^*_{it} - \mathbb{E}(\hat{u}^*_{it})\right] \qquad (43)$$

Now express $\hat{u}^*_{it}$ as $\hat{u}^*_{it} = lev_{it} - \hat{\beta}^* x^*_{it} = \beta x_{it} + u_{it} - \hat{\beta}^*(x_{it}+\eta_{it}) = (\beta - \hat{\beta}^*)x_{it} - \hat{\beta}^*\eta_{it} + u_{it}$

[13] Let $x_1 \dots x_N$ be multivariate normal, and collect $(x_1 \dots x_m)'$ in a vector x_a, and $(x_{m+1} \dots x_N)'$ in a vector x_b $(1 \leq m \leq N-1)$. Then stack the vectors and let $\mathbf{x} = \begin{pmatrix} x_a \\ x_b \end{pmatrix}$ with mean $\begin{pmatrix} \mu_a \\ \mu_b \end{pmatrix}$ and covariance matrix $\mathbf{\Sigma} = \begin{pmatrix} \Sigma_a & \Sigma_{ab} \\ \Sigma_{ba} & \Sigma_b \end{pmatrix}$. Then $\mathbb{E}(x_a|x_b) = \mu_a + \Sigma_{ab}\Sigma_b^{-1}(x_b - \mu_b)$, where $\Sigma_{ab}\Sigma_b^{-1}$ can be interpreted as the coefficients of a regression of x_a on x_b (see, for example, Greene (2003)).

and substitute:

$$\begin{aligned}
\mathbb{E}(x_{it}|\hat{u}^*_{it}) &= \mathbb{E}(x_{it}) + \frac{\mathbb{E}\left[(\beta-\hat{\beta}^*)x^2_{it} - \hat{\beta}^*\eta_{it}x_{it} + u_{it}x_{it}\right] - \mathbb{E}(x_{it})\mathbb{E}[(\beta-\hat{\beta}^*)x_{it}]}{\mathbb{E}\left[(\beta-\hat{\beta}^*)^2x^2_{it} + (\hat{\beta}^*)^2\eta^2_{it} + u^2_{it}\right] - \left[(\beta-\hat{\beta}^*)\mathbb{E}(x_{it})\right]^2}\hat{u}^*_{it} \\
&= \mathbb{E}(x_{it}) + \frac{(\beta-\hat{\beta}^*)\left[\mathbb{E}(x^2_{it}) - \mathbb{E}(x_{it})^2\right]}{(\beta-\hat{\beta}^*)^2[\mathbb{E}(x^2_{it}) - \mathbb{E}(x_{it})^2] + (\hat{\beta}^*)^2\mathbb{E}(\eta^2_{it}) + \mathbb{E}(u^2_{it})}\hat{u}^*_{it} \\
&= \mathbb{E}(x_{it}) + \frac{(\beta-\hat{\beta}^*)\sigma^2_{x_{it}}}{(\beta-\hat{\beta}^*)^2\sigma^2_{x_{it}} + (\hat{\beta}^*)^2\sigma^2_{\eta_{it}} + \sigma^2_{u_{it}}}\hat{u}^*_{it} \qquad (44)
\end{aligned}$$

Expanding the quadratic in the denominator and substituting $\hat{\beta} = \beta\frac{\sigma^2_{x_{it}}}{\sigma^2_{x_{it}}+\sigma^2_{\eta_{it}}}$ gives

$$\begin{aligned}
\mathbb{E}(x_{it}|\hat{u}^*_{it}) &= \mathbb{E}(x_{it}) + \frac{(\beta-\hat{\beta}^*)\sigma^2_{x_{it}}}{(\beta^2 - 2\beta\hat{\beta}^* + (\hat{\beta}^*)^2)\sigma^2_{x_{it}} + (\hat{\beta}^*)^2\sigma^2_{\eta_{it}} + \sigma^2_{u_{it}}}\hat{u}^*_{it} \\
&= \mathbb{E}(x_{it}) + \frac{(\beta-\hat{\beta}^*)\sigma^2_{x_{it}}}{\beta(\beta-2\hat{\beta}^*)\sigma^2_{x_{it}} + \hat{\beta}^*\beta\frac{\sigma^2_{x_{it}}}{\sigma^2_{x_{it}}+\sigma^2_{\eta_{it}}}(\sigma^2_{x_{it}}+\sigma^2_{\eta_{it}}) + \sigma^2_{u_{it}}}\hat{u}^*_{it} \\
&= \mathbb{E}(x_{it}) + \frac{(\beta-\hat{\beta}^*)\sigma^2_{x_{it}}}{\beta(\beta-\hat{\beta}^*)\sigma^2_{x_{it}} + \sigma^2_{u_{it}}}\hat{u}^*_{it} \\
&= \mathbb{E}(x_{it}) + b\cdot\hat{u}^*_{it}; \qquad b = \left(\beta + \frac{\sigma^2_{u_{it}}(\sigma^2_{x_{it}}+\sigma^2_{\eta_{it}})}{\beta\sigma^2_{x_{it}}\sigma^2_{\eta_{it}}}\right)^{-1} \qquad (45)
\end{aligned}$$

In my setup, $\mathbb{E}[x_{it}] = 0$, so the expectation of x_{it} conditional on the regression residual $\hat{u}^*_{it}$ is

$$\mathbb{E}(x_{it}|\hat{u}^*_{it}) = b\cdot\hat{u}^*_{it} = \left[\beta + \frac{\sigma^2_{u_{it}}}{\beta}\left(\frac{1}{\sigma^2_{\eta_{it}}} + \frac{1}{\sigma^2_{x_{it}}}\right)\right]^{-1}\hat{u}^*_{it} \qquad (46)$$

Finally, substitute (46) into (9) to obtain an expression for the conditional expectation for next period's leverage:

$$\mathbb{E}(lev_{it+1}|\hat{u}^*_{it}) = \beta\phi b\cdot\hat{u}^*_{it} = \phi\ \underbrace{\left[1 + \frac{\sigma^2_{u_{it}}}{\beta^2}\left(\frac{1}{\sigma^2_{\eta_{it}}} + \frac{1}{\sigma^2_{x_{it}}}\right)\right]^{-1}}_{=c\geq 0}\hat{u}^*_{it} \qquad (47)$$

B.3 Residual Persistence

As before, express the regression residual as

$$\begin{aligned}
\hat{u}^*_{it+1} &= (\beta-\hat{\beta}^*)x_{it+1} - \hat{\beta}^*\eta_{it+1} + u_{it+1} \\
&= (\beta-\hat{\beta}^*)(\phi x_{it} + \epsilon_{it}) - \hat{\beta}^*\eta_{it+1} + u_{it+1} \qquad (48)
\end{aligned}$$

Then

$$\begin{aligned}\mathbb{E}(\hat{u}^*_{it+1}|\hat{u}^*_{it}) &= (\beta-\hat{\beta}^*)\left[\phi\mathbb{E}(x_{it}|\hat{u}^*_{it})+\mathbb{E}(\epsilon_{it}|\hat{u}^*_{it})\right]-\hat{\beta}^*\mathbb{E}(\eta_{it+1}|\hat{u}^*_{it})+\mathbb{E}(u_{it+1}|\hat{u}^*_{it}) \\ &= (\beta-\hat{\beta}^*)\phi\mathbb{E}(x_{it}|\hat{u}^*_{it}) = \left[\phi(\beta-\hat{\beta}^*)b\right]\hat{u}^*_{it} \qquad (49)\end{aligned}$$

C Implied Target Leverage Derivations

Derivation of α_0

Begin with the relationship between the mismeasured target and the true target:

$$\widehat{lev^*} = \alpha_0 + \alpha_1\widehat{lev} + e \qquad (50)$$

Taking expectations:

$$\mathbb{E}(\widehat{lev^*}) = \alpha_0 + \alpha_1\mathbb{E}(\widehat{lev}) + 0 \qquad (51)$$

The mismeasured target and the true target are equal, on average, i.e., $\mathbb{E}(\widehat{lev^*}) = \mathbb{E}(\widehat{lev})$. To see this, start with the regression specification where the explanatory variable x^* is measured with error (* denotes that a variable or parameter is affected by measurement error):

$$lev = \beta_0^* + \beta_1^* x^* + \epsilon^* \qquad (52)$$

Taking expectations:

$$\begin{aligned}\mathbb{E}(lev) &= \beta_0^* + \beta_1\frac{\sigma_x^2}{\sigma_x^2+\sigma_\eta^2}\mathbb{E}(x+\eta)+\mathbb{E}(\epsilon^*) \\ &= \beta_0^* + \frac{\sigma_x^2}{\sigma_x^2+\sigma_\eta^2}\mathbb{E}(lev) \qquad (53)\end{aligned}$$

Therefore,

$$\beta_0^* = \mathbb{E}(lev)\left(1-\frac{\sigma_x^2}{\sigma_x^2+\sigma_\eta^2}\right) \qquad (54)$$

Mismeasured target leverage is given by

$$\widehat{lev^*} = \beta_0^* + \beta_1^* x^* \qquad (55)$$

Substituting for β_0^* and β_1^* shows that the mismeasured target equals the true target (and, hence, actual leverage), on average:

$$\begin{aligned}\mathbb{E}(\widehat{lev^*}) &= \mathbb{E}(lev)\left(1-\frac{\sigma_x^2}{\sigma_x^2+\sigma_\eta^2}\right)+\frac{\sigma_x^2}{\sigma_x^2+\sigma_\eta^2}\beta_1\mathbb{E}(x+\eta) \\ &= \mathbb{E}(lev)\left(1-\frac{\sigma_x^2}{\sigma_x^2+\sigma_\eta^2}\right)+\frac{\sigma_x^2}{\sigma_x^2+\sigma_\eta^2}\mathbb{E}(lev) \\ &= \mathbb{E}(lev) \qquad (56)\end{aligned}$$

Substituting (56) into (51) then yields an expression for α_0:

$$\begin{aligned} \mathbb{E}(\widehat{lev^*}) = \mathbb{E}(\widehat{lev}) &= \alpha_0 + \alpha_1 \mathbb{E}(\widehat{lev}) + 0 \\ \alpha_0 &= (1-\alpha_1)\mathbb{E}(\widehat{lev}) \end{aligned} \tag{57}$$

Derivation of α_1

Since (50) above is a regression equation:

$$\alpha_1 = \frac{cov(\widehat{lev}, \widehat{lev^*})}{Var(\widehat{lev})} \tag{58}$$

Expanding the numerator:

$$\begin{aligned} cov(\widehat{lev}, \widehat{lev^*}) &= Cov\left(\beta_1 x, \beta_0^* + \beta_1^*(x+\eta)\right) \\ &= Cov\left(\beta_1 x, \mathbb{E}(lev)\left(1 - \frac{\sigma_x^2}{\sigma_x^2+\sigma_\eta^2}\right) + \beta_1 \frac{\sigma_x^2}{\sigma_x^2+\sigma_\eta^2}(x+\eta)\right) \\ &= Cov\left(\beta_1 x, \beta_1 \frac{\sigma_x^2}{\sigma_x^2+\sigma_\eta^2} x\right) + \underbrace{Cov\left(\beta_1 x, \beta_1 \frac{\sigma_x^2}{\sigma_x^2+\sigma_\eta^2}\eta\right)}_{=0} \\ &= \beta_1^2 \frac{\sigma_x^2}{\sigma_x^2+\sigma_\eta^2} Var(x) \end{aligned} \tag{59}$$

Substituting:

$$\begin{aligned} \alpha_1 &= \frac{cov(\widehat{lev}, \widehat{lev^*})}{Var(\widehat{lev})} = \frac{\beta_1^2 \frac{\sigma_x^2}{\sigma_x^2+\sigma_\eta^2} Var(x)}{\beta_1^2 Var(x)} \\ &= \frac{\sigma_x^2}{\sigma_x^2+\sigma_\eta^2} \end{aligned} \tag{60}$$

Finally, let $a = \frac{\sigma_\eta^2}{\sigma_x^2}$, and substitute into (60):

$$\alpha_1 = \frac{1}{1+a} \tag{61}$$

Derivation of σ_e^2

Start again with (50), and compute the variance:

$$Var(\widehat{lev^*}) = \alpha_1^2 Var(\widehat{lev}) + \sigma_e^2 \tag{62}$$

We can calculate $Var(\widehat{lev})$ from calibrating the true target to resemble the leverage-sorted portfolios. To compute $Var(\widehat{lev^*})$, start again with

$$\widehat{lev^*} = \beta_0^* + \beta_1^* x^* = \beta_0^* + \beta_1 \frac{\sigma_x^2}{\sigma_x^2+\sigma_\eta^2}(x+\eta) \tag{63}$$

Then compute the variance:

$$
\begin{aligned}
Var(\widehat{lev^*}) &= Var\left(\beta_1 \frac{\sigma_x^2}{\sigma_x^2+\sigma_\eta^2}(x+\eta)\right) \\
&= \left(\beta_1 \frac{\sigma_x^2}{\sigma_x^2+\sigma_\eta^2}\right)^2 (\sigma_x^2+\sigma_\eta^2) \\
&= \beta_1^2 \frac{(\sigma_x^2)^2}{\sigma_x^2+\sigma_\eta^2} \\
&= \underbrace{Var(\widehat{lev}) \frac{\sigma_x^2}{\sigma_x^2+\sigma_\eta^2}}_{\leq Var(\widehat{lev})} \qquad (64)
\end{aligned}
$$

Substitute (64) into (62), and solve for σ_e^2:

$$
Var(\widehat{lev}) \frac{\sigma_x^2}{\sigma_x^2+\sigma_\eta^2} = \alpha_1^2 Var(\widehat{lev}) + \sigma_e^2 \qquad (65)
$$

Again, express measurement error as a fraction of the variability of the true x: $\sigma_\eta^2 = a\sigma_x^2$. We can now solve for the implied variance of the residual e as a function of the amount of measurement error present:

$$
\begin{aligned}
\sigma_e^2 &= Var(\widehat{lev}) \left[\frac{\sigma_x^2}{\sigma_x^2+\sigma_\eta^2} - \left(\frac{\sigma_x^2}{\sigma_x^2+\sigma_\eta^2}\right)^2\right] \\
&= Var(\widehat{lev}) \left[\frac{1}{1+a} - \frac{1}{(1+a)^2}\right] \\
&= Var(\widehat{lev}) \frac{a}{(1+a)^2} \qquad (66)
\end{aligned}
$$

References

Baker, Malcolm, and Jeffrey Wurgler. 2002. "Market Timing and Capital Structure." *Journal of Finance* 55 (1): 1–32.

Chang, Xin, and Sudipto Dasgupta. 2009. "Target Behaviour and Financing: How Conclusive is the Evidence?" *Journal of Finance* 64 (4): 1767–1796.

DeAngelo, Harry, Linda DeAngelo, and Toni Whited. 2011. "Capital Structure Dynamics and Transitory Debt." *Journal of Financial Economics* 99 (2): 235–261.

DeAngelo, Harry, and Richard Roll. 2011. "How stable are corporate capital structures?" *Journal of Finance*, forthcoming.

Erickson, Tim, and Toni M. Whited. 2006. "On the accuracy of different measures of q." *Financial Management* 35 (3): 5–33.

Fischer, Edwin, Robert Heinkel, and Josef Zechner. 1989. "Dynamic Capital Structure Choice: Theory and Tests." *Journal of Finance* 44 (1): 19–40.

Flannery, Mark, and Kasturi Rangan. 2006. "Partial Adjustment toward Target Capital Structures." *Journal of Financial Economics* 79 (3): 469–506.

Frank, Murray, and Vidhan Goyal. 2009. "Capital Structure Decisions: Which Factors are Reliably Important?" *Financial Management* 38 (1): 1–37.

Graham, John, and Campbell Harvey. 2001. "The Theory and Practice of Corporate Finance: Evidence from the Field." *Journal of Financial Economics* 60 (2): 187–243.

Granger, Clive, and Paul Newbold. 1977. *Forecasting economic time series.* New York, NY: Academic Press.

Greene, William H. 2003. *Econometric Analysis.* Upper Saddle River, NJ: Pearson Education, Inc.

Hennessy, Christoper, and Toni M. Whited. 2005. "Debt Dynamics." *Journal of Finance* 60 (3): 1129–1165.

Hovakimian, Armen, Tim Opler, and Sheridan Titman. 2001. "The Debt-Equity Choice." *Journal of Financial and Quantitative Analysis* 36 (1): 1–24.

Leary, Mark, and Michael Roberts. 2005. "Do Firms Rebalance Their Capital Structure?" *Journal of Finance* 60 (6): 2575–2620.

Lemmon, Michael, Michael Roberts, and Jaime Zender. 2008. "Back to the Beginning: Persistence and the Cross-Section of Corporate Capital Structure." *Journal of Finance* 63 (4): 1575–1608.

Menichini, Amilcar. 2010. "Financial frictions and capital structure choice: A structural estimation." Unpublished working paper, The University of Arizona.

Modigliani, Franco, and Merton H. Miller. 1958. "The cost of capital, corporation finance and the theory of investment." *American Economic Review* 48 (3): 261–297.

Myers, Stewart. 1984. "The Capital Structure Puzzle." *Journal of Finance* 39 (3): 575–592.

Rajan, Raghuram, and Luigi Zingales. 1995. "What Do We Know about Capital Structure: Some Evidence from International Data." *Journal of Finance* 50 (5): 1421–1460.

Roberts, Michael. 2002. "The Dynamics of Capital Structure: An Empirical Analysis of a Partially Observable System." Unpublished working paper, Duke University.

Strebulaev, Ilya, and Baozhong Yang. 2013. "The Mystery of Zero-Leverage Firms." *Journal of Financial Economics* 109 (1): 1–23.

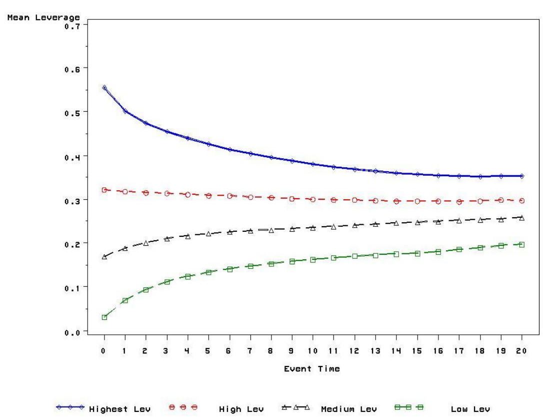
Titman, Sheridan, and Roberto Wessels. 1988. "The Determinants of Capital Structure Choice." *Journal of Finance* 43 (1): 1–19.

Welch, Ivo. 2004. "Capital Structure and Stock Returns." *Journal of Political Economy* 112 (1): 106–131.

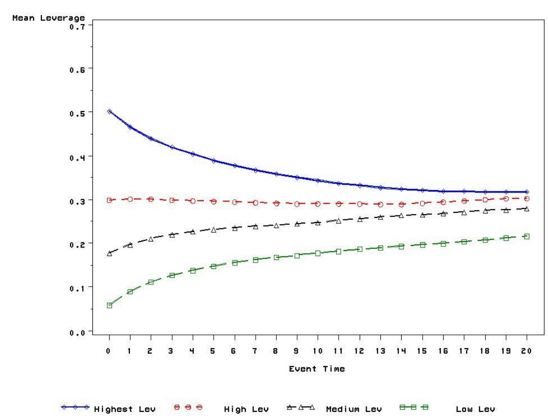
Figures

Figure 1: Average Leverage of Book Leverage Portfolios

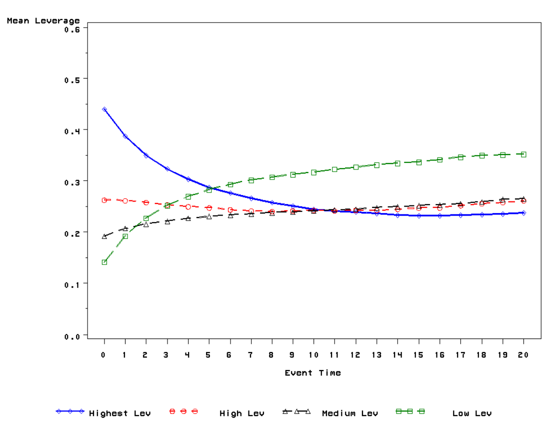
Using the 1965-2003 sample of nonfinancial Compustat firms, I sort firms into 4 portfolios. In Panel A, the sort is based on the firm's actual level of book leverage. In Panel B, the sort is based on residuals from a regression of book leverage on lagged size, market-to-book, profitability, tangibility and mean industry leverage. In Panel C, a firm fixed effect (FE) is added to the other explanatory variables. I then compute the mean leverage of each portfolio for the next 20 years, keeping its composition constant. I repeat this procedure for all years until the end of the sample period. The resulting 38 portfolio time series are then averaged in event time. Variables are defined in Appendix A.



Panel A: Sort on Actual Leverage



Panel B: Sort on Unexpected Leverage



Panel C: Sort on Unexpected Lev. with FE

Figure 2: Simulated Portfolio Convergence

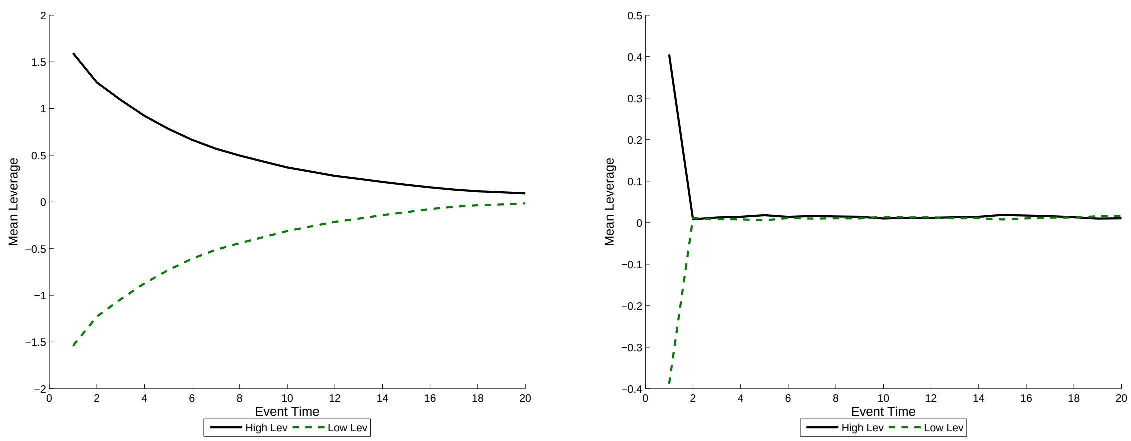
The two panels show the evolution of leverage portfolios where simulated firms are sorted into either a high- or a low-leverage portfolio. In Panel A, the sort is based on actual leverage at time 0, while in Panel B, it is based on unexpected leverage at time 0. Unexpected leverage is the residual obtained from a cross-sectional regression of leverage on its determinant, which is estimated each year. The firms are kept in their respective portfolios for 20 years. The sort is carried out every year for 40 years, giving rise to 40 time series, each being 20 years long. The time series are then averaged in event time within each portfolio, resulting in the graphs above. Individual firm time series are produced as follows: each period, leverage is determined as a function of an explanatory variable x:

$$lev_{it} = \beta x_{it} + u_{it} \tag{67}$$

where $\beta = 1$ and $u_{it} \sim N(0, 0.25)$. The leverage determinant x_{it} follows an AR(1) process:

$$x_{it} = \phi x_{it-1} + \epsilon_{it} \tag{68}$$

with $\phi = 0.85$ and $\epsilon_{it} \sim N(0, 1)$. The time series for x is simulated for 160 time periods, of which only the last 60 are retained to approximate a steady state. I simulate a cross-section of 5,000 firms.



Panel A: Sort on Actual Leverage

Panel B: Sort on Unexpected Leverage

Figure 3: Comparison of Portfolio-Leverage Dispersion as a Function of Measurement Error

The simulation setup is as before, e.g., in Figure 2, but only the mismeasured regressor x^*_{it} is available. I simulate the following system 5,000 times:

$$lev_{it} = \beta x_{it} + u_{it} \tag{69}$$

$$x_{it} = \phi x_{it-1} + \epsilon_{it} \tag{70}$$

$$x^*_{it} = x_{it} + \eta_{it} \tag{71}$$

where $\beta = 1$, $u_{it} \sim N(0, 0.25)$, $\phi = 0.85$, and $\epsilon_{it} \sim N(0, 1)$. The available regressor x^* is imperfectly measured. I perform the residual-based portfolio sorts as before, for three levels of measurement error: $\sigma_\eta \in \{0, 0.5, 1\}$. The ratio of measurement noise to state noise in the regressor is thus also $\sigma_\eta/\sigma_\epsilon \in \{0, 0.5, 1\}$. The *leverage*-based portfolio sort (solid line) is included for reference. The average portfolio-leverage levels are shown over an event horizon of 20 time periods.

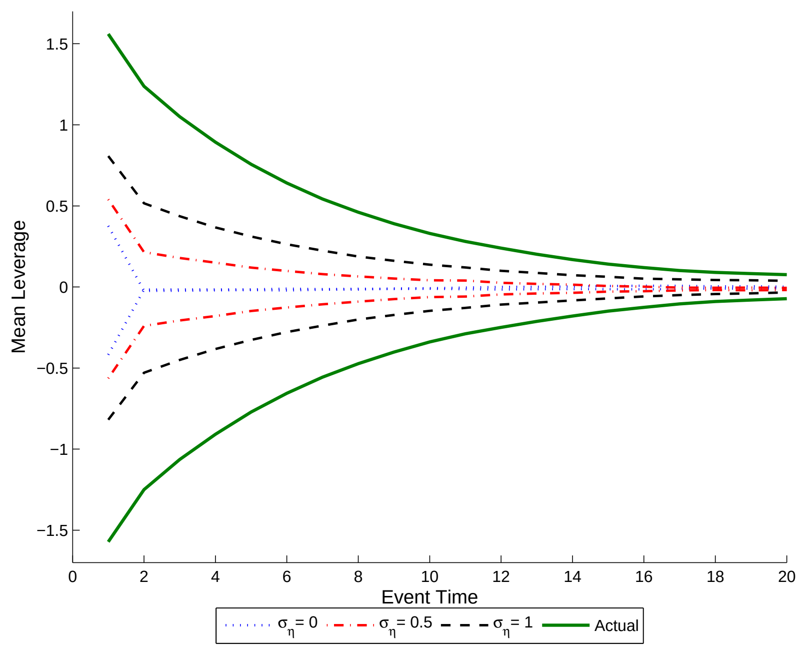


Figure 4: Data-Implied Target Leverage Dynamics

I model leverage lev as a function of its target $\widehat{lev}$, which in turn is an AR(1) process:

$$lev_t = \widehat{lev}_t + u_t \tag{72}$$
$$\widehat{lev}_t = \varphi_0 + \varphi_1 \widehat{lev}_{t-1} + \varepsilon_t \tag{73}$$

The red crosses correspond to actual portfolio-leverage levels. I simulate a panel of 1,000 firms, and choose parameter values for the system above such that the simulated data most closely resemble the actual data points by minimizing the sum of the squared deviations:

$$\min_{\Phi} \sum_i \sum_t \left(PFlev_{it}^{sim} - PFlev_{it}^{act}\right)^2 \tag{74}$$

where i indexes whether a data point belongs to a high- or low-leverage portfolio at time t, and the parameter vector $\Phi = \{\sigma_u^2, \varphi_0, \varphi_1, \sigma_\varepsilon^2\}$. The parameters are, respectively: the cross-sectional error variance in (72), as well as the intercept, slope and error variance for the AR(1) process governing target leverage in (73). The estimates are as follows:

	φ_0	φ_1	σ_ε	σ_u
Estimate	0.021	0.930	0.066	0.080
Std. Error	(0.012)	(0.009)	(0.003)	(0.010)

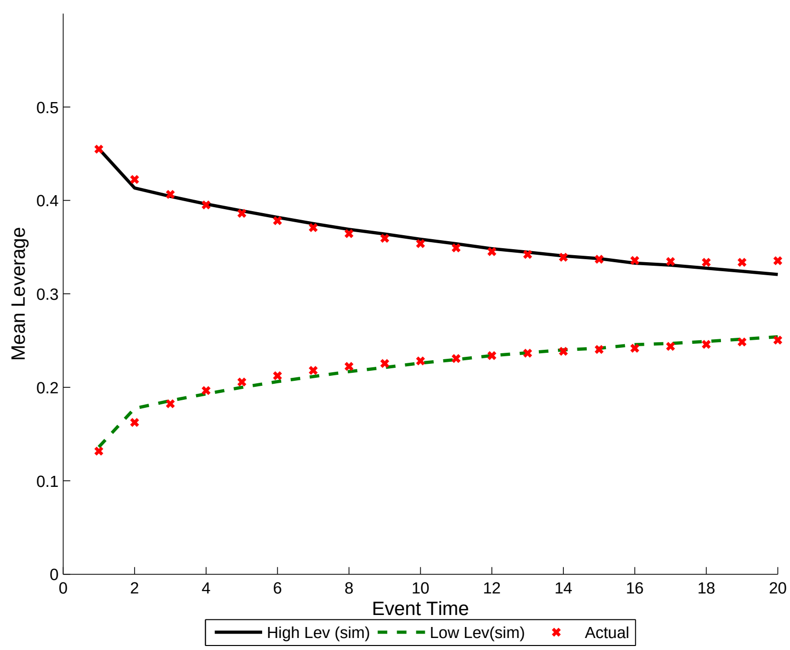
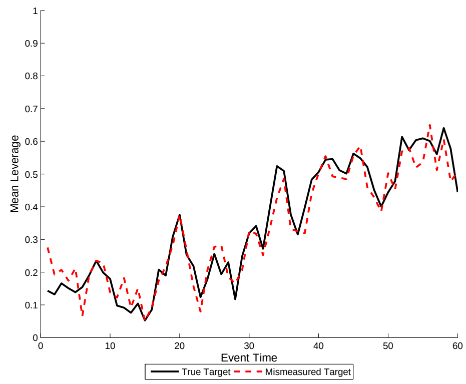


Figure 5: Sample Mismeasured Target-Leverage Paths

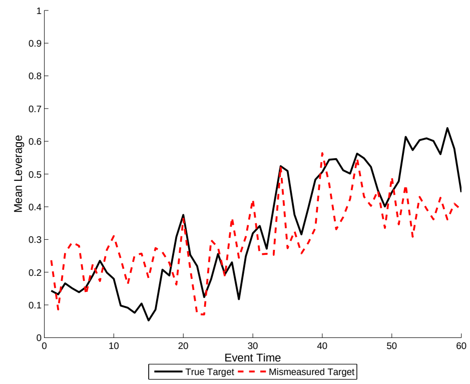
I first simulate a true target based on the parameters recovered via (18): $\Phi = \{\varphi_0 = 0.021,\ \varphi_1 = 0.93,\ \sigma_\varepsilon = 0.066,\ \sigma_u = 0.080\}$. The mismeasured target is then given by

$$\begin{aligned}
\widehat{lev^*} &= \alpha_0 + \alpha_1 \widehat{lev} + e && (75)\\
\alpha_0 &= (1-\alpha_1)\mathbb{E}(\widehat{lev}) && (76)\\
\alpha_1 &= \frac{1}{1+a} &&\\
\sigma_e^2 &= Var(\widehat{lev})\frac{a}{(1+a)^2} && (77)
\end{aligned}$$

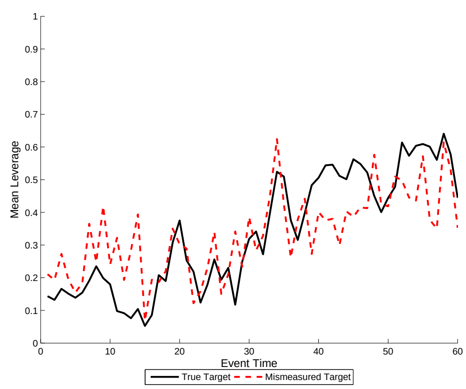
The four panels show sample leverage paths for different levels of the noise-to-signal ratio $a = \frac{\sigma_\eta^2}{\sigma_x^2}$. The true target is the same in all panels.



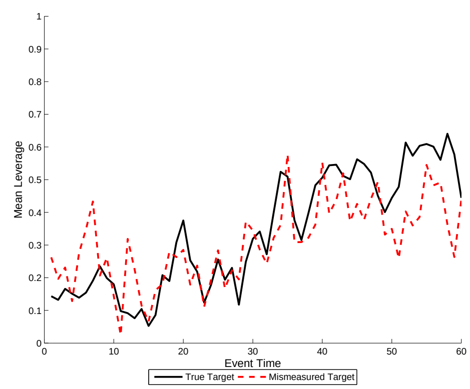
Panel A: $a = 0.1$



Panel B: $a = 0.5$



Panel C: $a = 0.75$



Panel D: $a = 1.25$

Figure 6: Implied Measurement Error from Residual-Based Sorts

The figure shows the implied ratio of measurement noise to variation in the true explanatory variable x. After parameterizing the dynamics of the true target, the mismeasured target $\widehat{lev^*}$ can be backed out via

$$\widehat{lev^*} = \alpha_0 + \alpha_1 \widehat{lev} + e \tag{78}$$

$$\alpha_0 = (1-\alpha_1)\mathbb{E}(\widehat{lev}) \tag{79}$$

$$\alpha_1 = \frac{1}{1+a}$$

$$\sigma_e^2 = Var(\widehat{lev})\frac{a}{(1+a)^2} \tag{80}$$

The red crosses correspond to actual portfolio-leverage levels. I simulate a panel of 1,000 firms, and choose the noise-to-signal ratio $a = \sigma_\eta^2/\sigma_x^2$ for the above system such that the simulated data most closely resemble the actual data points by minimizing the sum of squared deviations:

$$\min_a \sum_i \sum_t \left(PFlev_{it}^{sim} - PFlev_{it}^{act}\right)^2 \tag{81}$$

where i indexes whether a data point belongs to a high- or low-leverage portfolio at time t. The minimum of the objective function is reached at $a = 1.42$ (std. error $= 0.12$). The resulting fit is shown above.

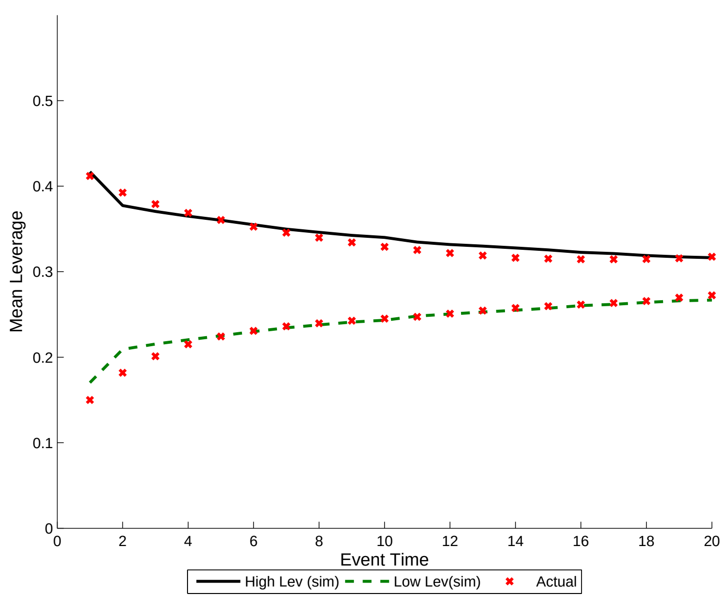
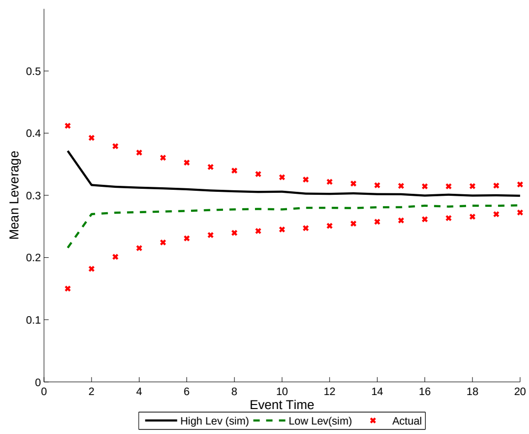
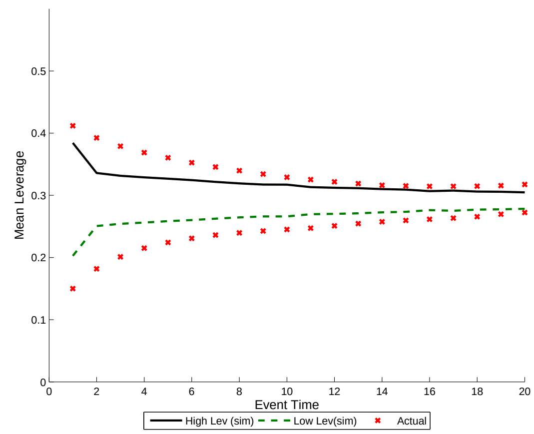


Figure 7: Residual-Sorted Leverage Portfolios at Different Implied Levels of Measurement Error

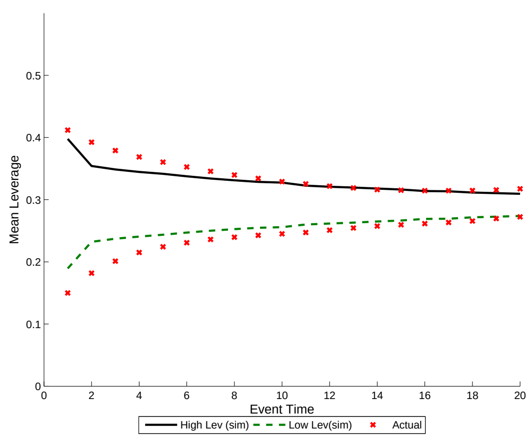
I simulate the set of equations in Figure 6 for different values of the noise-to-signal ratio $a = \sigma_\eta^2/\sigma_x^2$. Firms are sorted into portfolios based on residuals.



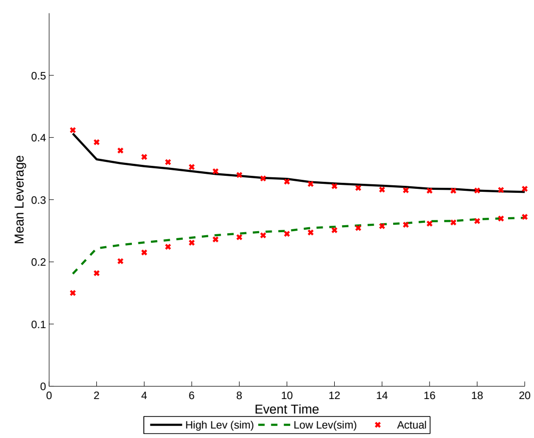
Panel A: $a = 0.1$



Panel B: $a = 0.25$



Panel C: $a = 0.5$

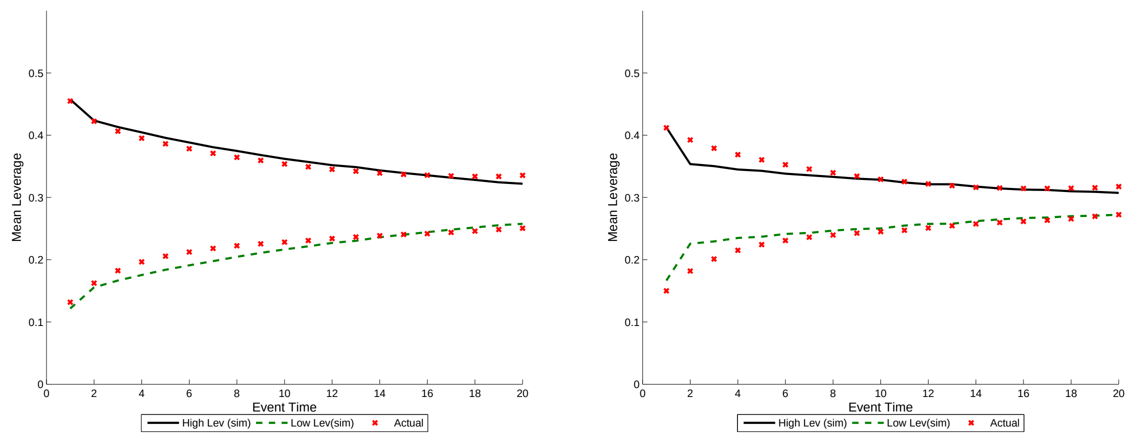


Panel D: $a = 0.75$

Figure 8: Average Leverage of Portfolios Sorted on Simulated "Actual" and "Unexpected" Leverage with *iid* Measurement Error

Panel A shows the evolution of the high- and low-leverage portfolios, when firms are sorted into portfolios based on simulated leverage. Firms are simulated using the parameters from Table 2, which are obtained by the calibration described in Section 3.3. Every period, simulated firms are sorted into either a high- or low-leverage portfolio, whose composition is held constant for 20 time periods. The figure shows the average leverage of the simulated portfolios in each year (solid and dashed lines). The simulated portfolios closely resemble the real data, depicted by the red crosses.

Panel B shows the results of doing the residual-based sort: leverage is regressed on mismeasured profitability, tangibility, market-to-book and industry leverage, and firms are then sorted into portfolios on the basis of the regression residual. The portfolio-leverage levels in years 5 and onward again closely resemble the real data, while the initial dispersion is lower than in the data.



Panel A:
Sort on Actual Leverage

Panel B:
Sort on Unexpected Leverage

Tables

Table 1: Summary Statistics

Summary statistics over the sample period 1965-2003 for non-financial firms on Compustat. Variable definitions are provided in Appendix A.

Variable	Mean	Minimum	Median	Maximum	Std Dev
lev	0.27	0.00	0.24	1.00	0.21
profit	0.05	-2.37	0.11	0.44	0.32
tang	0.34	0.00	0.28	0.93	0.25
MB	1.73	0.18	1.00	21.21	2.45
LnSize	4.18	-1.47	4.03	10.45	2.38

Table 2: Estimated Structural Parameters, with *iid* Measurement Error

This table lists the structural parameters governing the time-series and cross-sectional properties of the latent variables profitability, tangibility, market-to-book, and industry leverage in the four-variable calibration modelled via equations (26) through (32), as well as standard errors. The parameter values are found by minimizing the squared distance between simulated sample moments and actual data moments. The moments chosen are described in Section 3.3.1.

Variable	Parameter	Estimate	Std. Error
Profitability	ϕ_1	0.832	0.013
	σ_ϵ	0.194	0.007
	σ_η	0.105	0.006
Tangibility	ϕ_1	0.936	0.011
	σ_ϵ	0.090	0.011
	σ_η	0.038	0.009
Market-to-book	ϕ_1	0.931	0.011
	σ_ϵ	0.603	0.059
	σ_η	1.476	0.067
Industry leverage	ϕ_1	0.891	0.009
	σ_ϵ	0.039	0.010
	σ_η	0.012	0.004
Cross-sectional parameters	β_0	0.129	0.014
	β_{Prof}	-0.070	0.016
	β_{Tang}	0.115	0.011
	β_{MB}	-0.105	0.007
	β_{IndLev}	0.859	0.031
	σ_u	0.082	0.005

Table 3: Actual and Simulated Moments, with *iid* Measurement Error

This table lists actual data moments in the "Data Value" column, their simulated counterparts (excluding the portfolio-leverage levels) in the "Sim. Value" column, and the latent structural values in the "Struc. Value" column. The simulated moments are computed from simulated mismeasured variables using the estimated structural parameters from Table 2, and are described in Section 3.3.1. The latent structural values are obtained with the estimated structural parameter values from Table 2, and are again included here for ease of comparison.

Variable	Parameter	Data Value	Sim. Value	Struc. Value
Profitability	ϕ_0^*	0.009	0.009	0.007
	ϕ_1^*	0.775	0.764	0.832
	σ_x^{2*}	0.132	0.134	0.123
Tangibility	ϕ_0^*	0.017	0.031	0.023
	ϕ_1^*	0.952	0.916	0.936
	σ_x^{2*}	0.057	0.067	0.066
Market-to-book	ϕ_0^*	0.616	0.616	0.092
	ϕ_1^*	0.534	0.530	0.931
	σ_x^{2*}	4.895	4.895	2.717
Industry leverage	ϕ_0^*	0.028	0.036	0.033
	ϕ_1^*	0.908	0.879	0.891
	σ_x^{2*}	0.007	0.008	0.008
Cross-sectional	β_0^*	0.013	0.077	0.129
parameters	β_{Prof}^*	-0.066	-0.063	-0.070
	β_{Tang}^*	0.099	0.112	0.115
	β_{MB}^*	-0.006	-0.058	-0.105
	β_{IndLev}^*	0.835	0.834	0.859
Leverage	σ_{lev}^2	0.034	0.044	0.044

Table 4: Measurement-Error Ratio with *iid* Measurement Error

For each explanatory variable, column (1) shows estimates of the ratio of measurement noise σ_η^2 to variance in the latent explanatory variable σ_x^2, while column (2) shows the ratio of measurement noise σ_η^2 to total variance $\sigma_{x^*}^2$. The total variance is the variance of the mismeasured observed variable and thus includes the measurement-error variance. The values shown are computed with the structural parameter values in Table 2, which minimize the calibration's sum of squared errors.

	(1) σ_η^2/σ_x^2	(2) $\sigma_\eta^2/\sigma_{x^*}^2$
Profitability	0.090	0.083
Tangibility	0.021	0.021
Market-to-book	0.802	0.445
Industry leverage	0.018	0.018