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

Michael J. Mueller

Bank of Canada [1]

January 14, 2013

---

[1] Bank of Canada, 234 Wellington St., Ottawa ON K1A 0G9. Email: mmueller@bankofcanada.ca. Phone: +1-613-782-7675. Fax: +1-613-782-7136. I am very grateful to my thesis supervisor Murray Carlson, Ron Giammarino, Adlai Fisher, Russell Lundholm, Jason Chen, Rob Heinkel, Kai Li, Henry Siu, Kin Lo, Nathalie Moyen, Michael Faulkender, Alberto Romero, David Newton, Thomas Ruf, Zhongzhi Song, seminar participants at UBC, Florida International University, San Francisco State University, the Bank of Canada, and the NFA Conference, as well as an anonymous referee for their insights and helpful discussion.

Abstract

Firm leverage has been documented to be a slow-moving, persistent variable, even after controlling for leverage determinants. I show that if a firm's leverage dynamics are driven by a persistent explanatory variable that is measured with error, the mismeasured explanatory variable creates leverage persistence in a Lemmon et al. portfolio sort framework. In regression residual-sorted portfolios, a large positive residual will forecast above average future leverage. If a single factor drives leverage (we can think of this factor as a composite of many tradeoff theory-based explanatory variables), then the measurement error variance of this single "composite" variable needs to be 42% larger than its cross-sectional variance to reproduce the stylized facts of portfolio leverage persistence. Even small levels of measurement error produce a remarkable level of persistence in residual-based portfolio sorts. Furthermore, 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 suggests that unobserved investment opportunities may play an important role in explaining leverage ratios.

EFM Classification: 140

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[2]. Adding to the literature on leverage persistence, Lemmon et al. (2008) 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 show 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 various determinants motivated by the tradeoff 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 and slow convergence over time. In addition, Lemmon et al. (2008) find that a firm fixed effect is the variable that explains most of the initial cross-sectional variation and subsequent slow convergence. In this paper, I seek an explanation for this apparent firm fixed effect in the context of the Lemmon et al. (2008) portfolio sorts. My findings suggest that the apparent slow convergence of leverage towards an unconditional mean after controlling for determinants of leverage need not necessarily be taken as evidence against firms having a time-varying leverage target.

I focus on measurement error in empirical proxies of the true underlying economic variables as a culprit of the phenomenon. If leverage is governed by a persistent explanatory variable that is measured with error, using the mismeasured explanatory variable in a regression is problematic. It creates 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 is because the estimated residual is correlated with the true unobservable explanatory variable, which in turn predicts leverage.

It is reasonable to assume that the proxies used in capital structure analysis are tainted with measurement error. Consider asset tangibility, for instance. Under the tradeoff theory of capital structure, the more tangible a firm's collateral is, the more debt it will take on to take advantage of the interest tax shield, ceteris paribus. However, using a proxy like property, plant and equipment normalized by assets is an imperfect measure of true tangibility. For instance, let's compare a construction company and an airplane manufacturer with similar amounts of tangible assets, as defined by the afore-

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

mentioned ratio. In bankruptcy the construction company's cranes and excavators are easily transferable to another construction company, while the tooling used to assemble a Boeing 777 would be of questionable use in the assembly of an Airbus A-330, and thus less valuable. The empirical proxy, here given by the tangibility measure, will not always properly reflect the true underlying economic determinant.

I also obtain an estimate of the amount of measurement error consistent with the stylized facts. I find that if we assume that a single factor drives leverage (we can think of this factor as a composite of many tradeoff theory-based explanatory variables), then the measurement error variance of this single "composite" variable needs to be 42% larger than its cross-sectional variance to reproduce the Lemmon et al. (2008) findings. While this seems large, even 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 nonetheless exhibit a sizeable amount of persistence. Therefore, measurement error is likely to be an important contributor to the persistence in residual-sorted leverage portfolios, even if one takes the view that it is not the sole cause.

Finally, I examine measurement error in several explanatory variables consistent with studies such as Lemmon et al. (2008), Rajan and Zingales (1995) and Frank and Goyal (2009). By matching simulated data moments to empirical moments, I find that low quantities of measurement error in profitability, tangibility, and industry leverage, coupled with a measurement variance equal to about 80% of the cross-sectional variation in the market to book ratio, reproduce the stylized facts. This finding is consistent with other 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 true Tobin's q. My finding suggests that unobserved investment opportunities may play an important role in explaining leverage ratios.

Related Literature

Myers (1984) remarked that "we do not know how firms choose the debt, equity or hybrid securities they issue." Since then, much effort has gone into better understanding corporate capital structure, yet the question of whether firms do have a target capital structure towards which they 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 et al. (1989) propose a dynamic capital structure model, where firms adjust towards an optimum, but are hampered by adjustment costs. Hovakimian et al. (2001) provide evidence that firms behave in a fashion consistent with a tradeoff model, a finding that is echoed in Leary and Roberts (2005). Roberts (2001) 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 tradeoff theory. Baker and Wurgler (2002), on the other hand, suggest that firms' issuing behavior 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 tradeoff theory is not as strong as it may seem, as random financing generates data similar to what actually is observed. Overall, the evidence for firms adjusting towards an optimal capital structure is mixed, and Lemmon et al. (2008) also cast doubt on rebalancing behavior since they find firms' capital structures to be remarkably persistent over time.

By analyzing leverage portfolios, Lemmon et al. (2008) find that "high (low) levered firms tend to remain as such for over two decades", which seems to run counter to a world where firms actively rebalance their capital structures towards targets. 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 leaves their results largely unchanged. Instead of sorting on actual leverage, they now sort on the residual from a regression of leverage on lagged explanatory variables; the evolution of the average leverage of the resulting portfolios is similar to sorting on actual leverage directly. Persistent differences between the portfolios remain even 20 years after formation.

DeAngelo et al. (2011) offer a potential resolution of the Lemmon et al. (2008) results. In their model firms can finance investment either out of 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 there is credit rationing in place, which caps a firm's debt capacity. Equity issuance, on the other hand, is costly. Generally, firms will avoid carrying a cash balance due to the associated agency costs. Instead, they will free up debt capacity, so as to avoid a costly equity issuance when installing new capital. This "transitory" debt, coupled with various frictions in the model and cross-sectional dispersion in profitability shocks leads to leverage sorts that resemble those of Lemmon et al. (2008). However, the sorts are for raw leverage only: correctly controlling for the cross-sectional dispersion may reduce persistence in their model.

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 contain long-term persistence as in Lemmon et al. (2008). This obtains 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 capital structure stability altogether, and argue that it is the exception, and not the rule. They find that many firms which 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. Both Roberts (2001) and Flannery and Rangan (2006) suggest that measurement error may be partly responsible for the sluggish convergence in leverage ratios towards their mean, as measured by the adjustment speed parameter in a partial adjustment framework. The latter authors also show that including firm fixed effects speeds up estimated convergence speeds. Lemmon et al. (2008) also provide evidence for leverage persistence via their portfolio sorts, and suggest that a fixed effect may be responsible. My paper expands on this literature in the following way: I make precise the channel in which measurement error can add to the persistence of leverage and load on a fixed effect in the Lemmon et al. (2008) portfolio sort setting, and extract the amount of measurement necessary to reproduce the stylized facts.

2 Replication of the Lemmon et al. (2008) Findings

I start with a data sample that is comparable to that of Lemmon et al. (2008). The sample consists of firms listed in the annual Compustat database between 1965 and 2003. Financials 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} percentile. Table 1 presents summary statistics, which are similar to those in Lemmon et al. (2008). The table also displays a prominent feature of the data, namely the existence of zero-leverage firms, whose proportion is, in fact, sizeable (see e.g. Strebulaev and Yang (2012)).

[Table 1 about here.]

Next, I carry out the Lemmon et al. (2008) portfolio sorts. The procedure is as follows: Starting in 1965, and then every year thereafter, I sort firms into 4 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). Note that starting in 1983, the portfolios' time series length will decrease by one year every year. This results in a total of 38 portfolio time series of length 20 years, or less. The portfolios are then averaged cross-sectionally in event time, where the 'event' is the initial sort. Panel A of Figure 1 shows the results of this procedure.

[Figure 1 about here.]

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

Since the pattern uncovered in Figure 1 could be the consequence of cross-sectional variation in underlying determinants of firm leverage, Lemmon et al. (2008) regress leverage on lagged firm size, profitability, tangibility, market-to-book equity, and Fama-French 38-industry dummies[3]. The regressions are estimated every year, which allows for time-varying coefficient estimates. Then Lemmon et al. (2008) repeat the portfolio sorts with a modification: firms are now sorted into portfolios based on the estimated regression residuals (the "unexpected leverage") instead of on actual leverage.

I follow the Lemmon et al. (2008) methodology with a slight modification: 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 Lemmon et al. (2008)'s findings virtually identically (see their Panel A, Figure 2). The cross-sectional portfolio dispersion in the formation year is still large, albeit slightly reduced as compared to sorting on actual leverage. In addition, the portfolio dispersion remains persistent, and significant differences between the portfolios remain over the entire 20 years. In contrast, under a well-specified regression, 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

[3]Size, profitability and an industry dummy are used in e.g. Titman and Wessels (1988), while tangibility and market-to-book equity are used e.g. in Rajan and Zingales (1995).

levels. In a further variance decomposition of ANCOVA models, Lemmon et al. (2008) point out that a firm fixed effect is the largest component of the explained sum of squares, and largely subsumes other determinants of leverage.

Lemmon et al. (2008)'s findings are striking, and more importantly, 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 towards new theories, it is useful to know to what extent existing theories are able to accommodate the Lemmon et al. (2008) leverage portfolio graphs.

3 Possible Explanations

There are several possible channels that could give rise to the persistence of residual-based leverage portfolios. However, they are all manifestations of the same underlying cause: 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, it is possible that leverage is largely determined by a time-invariant firm fixed effect, as suggested by Lemmon et al. (2008). A firm fixed effect can be thought of as every firm having its own intercept in the regression. Since the omitted intercept 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 important information. An example of this strand of literature is the recent paper by DeAngelo et al. (2011), who model firms as incurring transitory debt obligations that represent deliberate, but temporary, deviations from a target capital structure. Carrying out portfolio sorts on their simulated firms also results in persistent leverage portfolio. While they do not sort on the basis of regression residuals, some of the persistence would likely remain: the failure to properly account for the level of transitory debt would leave an omitted variable imbedded in the residual.

A third possibility is that the regressions are misspecified from an economic viewpoint. For instance, if the firms face adjustment costs as in Fischer et al. (1989) or Hennessy and Whited (2005), there would be no target leverage as implied by the regression. Instead, the firm may choose to not alter its capital structure while leverage is within a certain range.

Finally, it is possible that our economic models are correct, but our empirical proxies for the benefits and costs of debt are inaccurate. Having mismeasured explanatory variables would again create a correlation between the regression residual and leverage itself, and sorting on the residual would resemble sorting on leverage.

Generally, distinguishing conclusively between these alternatives is difficult. Including firm fixed effects in the regression explains much of the cross-sectional variation between firms, because each firm is now allowed its own intercept. However, it does not 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 still is no convergence, as the average leverage level of the low leverage portfolio now is substantially higher than that of the high leverage portfolio after 20 years. In essence, including a fixed effect demeans the portfolio leverage time series, but the patterns, albeit shifted, still remain.

In the sections to follow, I show that the stylized facts obtain if leverage is a function of one or more mismeasured explanatory variables that exhibit a certain degree of persistence. I concede that my results cannot conclusively prove that measurement error is, in fact, the culprit of the phenomena I study. However, I will argue that measurement error in explanatory variables is intuitively sensible and consistent with the data, which studies like Flannery and Rangan (2006), Roberts (2001), and Erickson and Whited (2006) confirm.

3.1 Time Series Persistence in Leverage Portfolios as a Result of Measurement Error

Differences between portfolio leverage levels naturally arise during the formation period when firms are sorted into portfolios based on regression residuals. This is 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.

In this section, I formally examine persistence in leverage sorts, starting with the case where we can observe all variables perfectly. Here, sorting on the regression residual will produce initial dispersion that immediately vanishes in the periods after portfolio formation. I then show how persistence arises when the explanatory variable is 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.

3.1.1 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 it 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 et al. (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 $\epsilon_t \sim N(0, \sigma^2_{\epsilon_{it}})$. In the above, I am implicitly assuming that the explanatory variable, and hence leverage, have a mean of 0[4]. 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[5], then if we form portfolios by sorting on leverage and track their evolution over time, the high leverage portfolios decline only slowly towards the unconditional mean, while the leverage of low leverage portfolios increase 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.]

[4] This assumption is not crucial. We could easily add a mean to the explanatory variable x_{it} without affecting any of the conclusions. Furthermore, 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.

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

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 due 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.

3.1.2 Case II: Persistence as a Consequence of Measurement Error

In Section 3.1.1 I explain how, 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 $\eta_{it} \sim N(0, \sigma^2_{\eta_{it}})$ is measurement error, and 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 $\widehat{}$ 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 no longer unbiased (see Appendix B.1):

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

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:*

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

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 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. *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 the expectation of the regressor, conditional on the estimated regression residual* $\hat{u}^*_{it}$ *is:*

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

Proof. See Appendix B.2, beginning with (44). □

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 compactly 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 before. The dashed lines show sorts for two levels of 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 do 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. *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 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. □

Finally, Lemma 2 shows that when a persistent regressor is mismeasured, the estimated regression residual itself will exhibit persistence. 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 is because the estimated firm fixed effect will load on the persistent error term.

4 Extracting Measurement Error From Explanatory Variables: A Calibration

In Section 3.1.2, I show both analytically and in simulations how measurement error in a slow-moving explanatory variable can lead to leverage persistence in residual-based portfolio sorts. While that section thus shows the theoretical channel through which measurement error can produce persistence in portfolio sorts, it does not answer the important question of how much measurement error is needed to reproduce wide initial dispersion between the residual-based portfolios, followed by slow convergence.

The objective of this section is to assess whether a reasonably calibrated model with measurement error in explanatory variables can satisfactorily explain the data. I do this with two different approaches: in the first approach, described in Sections 4.1 - 4.2, I use the Lemmon et al. (2008) actual leverage portfolio sorts 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 portfolio leverage time series from the *actual* leverage-based sorts 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 needs to be in order to be consistent with the residual-based sorts.

The second approach of quantifying the amount of measurement error needed to reproduce the Lemmon et al. (2008) portfolio sorts is outlined in Section 4.3. There, I examine four actual explanatory variables consistent with the Lemmon et al. (2008) study, namely profitability, tangibility (which measures 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. I determine how mismeasured each of these needs to be in order to be consistent not only with the portfolio sorts, but also with other observed data moments.

4.1 Estimating Target Leverage Dynamics

Under the first approach, I begin by parameterizing the law of motion for a firm's target leverage. In Section 4.2 I use this law of motion in conjunction with residual-based portfolio sorts to back out an estimate of measurement error. To start, consider again the setup from Section 3.1, 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 η_{it}:

$$lev_{it} = \beta x_{it} + u_{it} \tag{12}$$

$$x_{it} = \phi_0 + \phi_1 x_{it-1} + \epsilon_{it} \tag{13}$$

$$x^*_{it} = x_{it} + \eta_{it} \tag{14}$$

where $u_{it} \sim N(0, \sigma_u^2)$, $\epsilon_{it} \sim N(0, \sigma_\epsilon^2)$, and $\eta_{it} \sim N(0, \sigma_\eta^2)$. An intercept ϕ_0 is included in the AR(1) process for the leverage determinant to allow for a non-zero mean. This is necessary because actual leverage is bounded between 0 and 1; the leverage portfolios have a positive mean, as seen in Figure 1, for instance.

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 just 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 in (13), scaled by the constant β. If more than one explanatory factor were included, the target dynamics can be thought of as a linear combination of AR(1) processes, which would result in an ARMA representation[6] for the target.

To estimate parameter values in (15) - (17), I proceed as follows: in the first step, I use the Lemmon et al. (2008) 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 only two portfolios, a "high leverage" portfolio and a "low leverage" portfolio, as opposed to the 4 portfolios in their original study. This does not affect the results; the main Lemmon et al. (2008) conclusions, namely initial convergence and long-term persistence, are still evident with only 2 portfolios.

Using the actual leverage-based portfolios, I estimate four parameters in equations (15) and (16): the cross-sectional error variance σ_u^2, and for the AR(1) process governing target leverage the intercept φ_0, slope coefficient φ_1 and error variance σ_ε^2. I simulate the system (15) - (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.

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

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)

[6]see e.g. Granger and Newbold (1977)

The estimated coefficients are of reasonable magnitudes, roughly in line with what a pooled regression would yield. In addition, simulating the Lemmon et al. (2008) *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.]

4.2 Estimating Measurement Error by Extracting the Mismeasured Target

Section 4.1 retrieves the dynamics of a leverage target $\widehat{lev_{it}} = \beta x_{it}$ by calibrating an AR(1) process for the true target to the Lemmon et al. (2008) portfolios sorted on actual leverage. The objective now is to extract a *mismeasured* leverage target $\widehat{lev^*_{it}}$ consistent with the *residual-based* portfolios. The mismeasured target will allow us to compute the mismeasured residuals, which form the basis of the residual-based portfolio sorts because

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

Equation (19) decomposes a leverage observation into a mismeasured target and a mismeasured residual. 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 modeling an explanatory variable x_{it}, or x^*_{it} in its mismeasured form, but focus on the target instead. 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 $\widehat{lev^*_{it}}$ (the fitted regression value) to be mismeasured as well. This mismeasured target can be expressed in terms of the true target $\widehat{lev_{it}}$ by the following regression:*

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

where

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

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

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

$$a = \frac{\sigma^2_\eta}{\sigma^2_x} \tag{24}$$

Proof. See Appendix C. □

Proposition 2 states that knowledge of the true target dynamics, via the methodology in Section 4.1, permits an explicit solution for the mismeasured target leverage in (20). The unknown parameters α_0, α_1 and σ_e^2 are functions only 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. Furthermore, as long as measurement error is present, the variance of the mismeasured target $Var(\widehat{lev}_{it}^*)$ is always less than the variance of the true target $Var(\widehat{lev}_{it})$ (see (65) in Appendix C for a proof). In fact, the larger the amount of measurement error in the underlying leverage determinant, the *smaller* will be the variation in estimated target leverage. This phenomenon warrants a short explanation. As we increase the amount of measurement error on the right hand side in a univariate regression, we increase attenuation in the slope coefficient. This naturally results in a larger estimated intercept. For instance, consider an extreme example where the signal-to-noise ratio of the explanatory variable goes to zero, i.e. the observed x_{it}^* is (almost) completely white noise. In this case, the estimated intercept will approach the unconditional mean of the dependent variable. This makes intuitive sense: the 'best' predicted value of the dependent variable in the presence of an (almost) useless explanatory variable should just be the dependent variable's unconditional mean.

This reasoning translates directly to the relationship between 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 (57) in Appendix C for a proof). Intuitively, this is due to regression mechanics: the mean predicted value will equal the dependent variable' 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 4 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 squared differences between the simulated and actual portfolios, 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 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 one needs to keep in mind that the sole factor x in my setup serves as a stand-in for all the determinants of leverage. For instance, in a multivariate world, high levels of measurement error in one variable counterbalances low levels of measurement error in another.

Another consideration is that in the previous calibration each portfolio received an equal weighting, which resulted in $a = 1.42$. Weighting some observations more heavily than others may also reduce the value of a. Furthermore, $a = 1.42$ results in the *best* fit, but it is instructive to assess the impact that lower noise-to-signal ratios a have on the residual-based portfolio sorts. Therefore, I simulate the portfolios for various levels of a. The results are shown in Figure 7.

[Figure 7 about here.]

The simulations show that 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. For instance, Panel 4 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. This proves that while large quantities of measurement error are needed to reproduce the stylized facts exactly, much more moderate levels still result in a good fit. Panel 2 depicts portfolios where $a = 0.25$; even after 20 time periods, a persistent difference remains between the high and low leverage portfolios. Furthermore, the difference in year 20 between the simulated portfolios is almost as large as that of the actual portfolios. This reinforces the view that measurement error is likely a contributing factor to persistence in residual-based portfolio sorts.

4.3 Multi-Variable Calibration with *iid* Measurement Error

In Section 4.2, I obtain an estimate of the noise-to-signal ratio that best reproduces the residual-based portfolio sorts in an implied single variable framework. I next investigate whether measurement error in explanatory variables similar to those used by Lemmon et al. (2008) is able to reproduce the leverage time series of both the actual- and residual-based portfolio sorts. I focus on profitability, tangibility, market-to-book, and industry leverage (in lieu of an industry fixed effect) as explanatory variables. Firm size is excluded, since it is not stationary and thus would not conform to my setup of modeling the explanatory variables as AR(1) processes.

The estimation procedure, a form of the simulated method of moments framework, proceeds in a similar fashion to that in Section 4.2. In particular, the economy consists of simulated firms whose leverage dynamics are governed by the following system of equations:

$$lev_{it} = \boldsymbol{\beta}'(1 \quad \boldsymbol{x_{it}})' + u_{it} \tag{26}$$

$$= \begin{pmatrix} \beta_0 & \beta_{Prof} & \beta_{Tang} & \beta_{MB} & \beta_{IndLev} \end{pmatrix} \begin{pmatrix} 1 \\ Prof_{it} \\ Tang_{it} \\ MB_{it} \\ IndLev_{it} \end{pmatrix} + u_{it} \tag{27}$$

$$\boldsymbol{x_{it}} = \boldsymbol{\phi_0} + \boldsymbol{\phi_1} \boldsymbol{x_{it-1}} + \boldsymbol{\epsilon_{it}} \tag{28}$$

$$= \begin{pmatrix} \phi_0^{Prof} \\ \phi_0^{Tang} \\ \phi_0^{MB} \\ \phi_0^{IndLev} \end{pmatrix} + \begin{pmatrix} \phi_1^{Prof} & 0 & 0 & 0 \\ 0 & \phi_1^{Tang} & 0 & 0 \\ 0 & 0 & \phi_1^{MB} & 0 \\ 0 & 0 & 0 & \phi_1^{IndLev} \end{pmatrix} \boldsymbol{x_{it-1}} + \begin{pmatrix} \epsilon_{it}^{Prof} \\ \epsilon_{it}^{Tang} \\ \epsilon_{it}^{MB} \\ \epsilon_{it}^{IndLev} \end{pmatrix} \tag{29}$$

$$\boldsymbol{x_{it}^*} = \boldsymbol{x_{it}} + \boldsymbol{\eta_{it}} \tag{30}$$

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 $\boldsymbol{\eta_{it}} \sim N(0, \boldsymbol{\Sigma_\eta})$. Leverage is determined in the cross-section by an intercept and the four explanatory factors, which are all modeled 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. This reflects the fact that 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_u^2 (6 parameters), and the measurement error variances (4 parameters).

To reduce the number of free parameters in the model, I infer the unconditional means of the noisy explanatory variables 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:

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

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

I use $\boldsymbol{I_4}$ to denote a 4×4 identity matrix. In fact, several other parameters could be inferred directly from the data[7], namely the variance of leverage, $Var(lev)$, and the variance matrix of the noisy explanatory variables $\boldsymbol{\Sigma_{x^*}}$. However, forcing the constraints

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

that the model imposes on the variances 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.

4.3.1 Identification

To reduce the dimensionality of the parameter space, I compute 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 $\boldsymbol{\Sigma_\eta}$ need to be estimated. 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, the parameter estimates of the mismeasured AR(1) processes driving the explanatory variables, the regression parameters from a panel regression of leverage on its noisy determinants, and the portfolio leverage levels of the Lemmon et al. (2008) portfolio sorts. Since I assume that the actual data on explanatory variables is contaminated by measurement error, all data moments involving explanatory variables are mismeasured as well. In particular, I use the following moments:

1. 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

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

$$\boldsymbol{\Sigma_\epsilon} = \left(\boldsymbol{I_4} - \boldsymbol{\phi_1}'\boldsymbol{\phi_1}\right) \boldsymbol{\Sigma_x} \quad (35)$$

$$= \left(\boldsymbol{I_4} - \boldsymbol{\phi_1}'\boldsymbol{\phi_1}\right)\left(\boldsymbol{\Sigma_{x^*}} - \boldsymbol{\Sigma_\eta}\right) \quad (36)$$

$$\sigma_u^2 = Var(lev) - \boldsymbol{\beta}\boldsymbol{\Sigma_x}\boldsymbol{\beta}' \quad (37)$$

The first equation is a rewritten expression for the variance of a vector of AR(1) processes. Solving for the variances of the error terms ϵ requires the slope coefficients and the variances of the latent explanatory variables, which I express as the difference between the variances of the observed mismeasured variables and the variances of the measurement error terms. The last expression computes the variance of (27) to solve for the variance of the residual.

(8 moments):

$$x^*_{it} = \phi^*_0 + \phi^*_1 x^*_{it-1} + \epsilon^*_{it} \tag{38}$$

2. The variance of each mismeasured explanatory variable σ^{2*}_x, and the variance of leverage σ^2_{lev} (5 moments).

3. The cross-sectional coefficients β^* from a regression of leverage on the noisy determinants (5 moments):

$$lev_{it} = \boldsymbol{\beta^{*\prime}}(1 \quad \boldsymbol{x^{*\prime}_{it}}) + u^*_{it} \qquad \text{where} \tag{39}$$

$$\boldsymbol{\beta^{*\prime}} = (\beta^*_0 \quad \beta^*_{Prof} \quad \beta^*_{Tang} \quad \beta^*_{MB} \quad \beta^*_{IndLev}) \tag{40}$$

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

4. The time series of portfolio leverage levels obtained 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 $\Phi = (\boldsymbol{\phi_1} \quad \boldsymbol{\Sigma_\epsilon} \quad \boldsymbol{\Sigma_\eta} \quad \boldsymbol{\beta} \quad \sigma_u)$ are found by minimizing the sum of squared differences between actual moments and simulated moments:

$$\min_{\Phi} \, (\boldsymbol{m^{act}} - \boldsymbol{m^{sim}})'(\boldsymbol{m^{act}} - \boldsymbol{m^{sim}}) \tag{41}$$

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

4.3.2 Results

The estimated structural parameters of this procedure, along with their standard errors[8], are listed in Table 2. Table 3 presents a comparison of empirical data moments, their simulated counterpart based on mismeasured variables, and moments that are based on the estimated true latent parameters. For each explanatory variable, I present intercept and slope coefficient of the AR(1) process and its variance. For the cross-sectional relationship between leverage and its determinants, I present the β-coefficients for each variable (including an intercept), as well as the leverage variance. Table 4 gives two estimates of the ratio of measurement noise to state noise for each

[8]The standard errors are obtained by bootstrapping: 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 then given by the standard deviations of the estimated structural parameters.

simulated explanatory variable. The first estimate is the ratio of measurement error variance to variance of the latent underlying variable, while the second estimate is the ratio of measurement error variance to 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 ratio, the calibrated values of the latent processes are very close to the empirical data values. As measured by the AR(1) 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)). Very similar values for the measurement error ratio are obtained if the variance of the observed explanatory variable is used instead. Consistent with the low quantity of estimated measurement error, the structural β-coefficients for both variables are close to their empirical counterparts (see Table 3).

Table 3 reveals that latent profitability ($\phi_1 = 0.832$) 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 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 0.09 and 0.083, respectively (see Table 4). These value are still low; for example, the latter value 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). Note that the simulated ϕ_1^* value, obtained by regressing simulated mismeasured market-to-book 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 compared to that in the other variables. Its value is $\sigma_\eta = 1.476$, which exceeds the standard deviation of the innovation term in the AR(1) process, whose value is $\sigma_\epsilon = 0.603$, 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." Their paper notes that this is because accounting information inaccurately reflects both the market value of debt and the replacement value of assets, and because strong assumptions are needed for Tobin's q to accurately reflect a firm's incentive to invest. Using a classical errors-in-variables model with the investment-to-capital ratio on the left-hand side and average q on the right-hand side, 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 model, 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 exactly 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, this means that the true latent β-coefficient for the market-to-book ratio is -0.105, which is much larger than the empirical value of -0.006 (Table 3). The simulated mismeasured observed value for β_{MB} is -0.058, which is larger than the data value. My results nonetheless 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 due 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 2-5, 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, is able to produce close matches to the Lemmon et al. (2008) portfolio sorts, and thus offers a potential explanation of their findings[9].

5 Conclusion

Persistence in residual-based leverage portfolios is a well-documented fact. While this persistence may be a consequence of a firm fixed effect or omitted time-varying variables, I show that it can also arise when slow-moving explanatory variables in a leverage regression are measured with error. Sorting firms into portfolios based on these regression residuals will resemble sorting firms into portfolios based on actual leverage.

Being able to predict future leverage with the regression residuals implies that target leverage is mismeasured. I find that if we view the target as being determined by a single composite factor of a number of possible tradeoff theory variables, then the measurement error variance of this latent factor needs to be 42% larger than its cross-sectional variance. This number is large, but nonetheless a useful measure, as one can interpret it as an aggregate estimate of how mismeasured the explanatory variables would need to be. In addition, a measurement error with a magnitude of 75% of the state noise variance of the latent variable still produces persistent residual-based portfolio sorts. Therefore, even if one takes the view that measurement error alone is not sufficient to fully account for the persistence in residual-sorted leverage portfolios, it nonetheless is likely to be an important contributor, since sizeable persistence in the residual-based portfolios arises even at low ratios of measurement error to state noise in the explanatory variable.

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 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. Furthermore, the level of measurement error in the market-to-book variable, which proxies for Tobin's q, is consistent with

[9] In unreported results, I determine that model fit can be improved by allowing for a slight autocorrelation in the measurement error terms themselves.

other studies such as Erickson and Whited (2006), and suggests that unobserved investment opportunities may play an important role in explaining leverage ratios, and 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 returns of trading strategies, and to test asset pricing models. Measurement quality is an important consideration for the risk factors in these models, so my work may have implications for the asset pricing applications of portfolio sorts as well.

Appendices

A Variable Definitions

Data are taken from the annual Compustat database between 1965 and 2003. The variable definitions mirror those in Lemmon et al. (2008). Financials and firms 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} percentile. 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:

$$\begin{aligned}\hat{\beta}^* &= \frac{cov(x_{it}^*, lev_{it})}{\sigma^2_{x_{it}^*}} = \frac{\mathbb{E}\left[(x_{it}+\eta_{it})(\beta x_{it}+u_{it})\right] - \mathbb{E}(x_{it}+\eta_{it})\mathbb{E}(\beta x_{it}+u_{it})}{\mathbb{E}[(x_{it}+\eta_{it})^2] - \mathbb{E}(x_{it}+\eta_{it})^2} \\ &= \frac{\beta\mathbb{E}(x_{it}^2) - \beta\mathbb{E}(x_{it})^2}{\mathbb{E}(x_{it}^2) + \mathbb{E}\eta_{it}^2 - \mathbb{E}(x_{it})^2} \\ &= \beta\frac{\sigma^2_{x_{it}}}{\sigma^2_{x_{it}} + \sigma^2_{\eta_{it}}} \qquad (42)\end{aligned}$$

B.2 Conditional Expectation of Leverage Under Measurement Error

We want to compute expected portfolio leverage, conditional on sorting on the mismeasured regression residual:

$$\mathbb{E}[lev_{it+1}|\hat{u}_{it}^*] = \beta\phi\mathbb{E}[x_{it}|\hat{u}_{it}^*] + \beta\mathbb{E}[\epsilon_{it+1}|\hat{u}_{it}^*] + \mathbb{E}[u_{it+1}|\hat{u}_{it}^*] \qquad (43)$$

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, however, is not equal to 0. 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 [10]:

$$\mathbb{E}(x_{it}|\hat{u}_{it}^*) = \mathbb{E}(x_{it}) + \frac{Cov(x_{it}, \hat{u}_{it}^*)}{Var(\hat{u}_{it}^*)}\left[\hat{u}_{it}^* - \mathbb{E}(\hat{u}_{it}^*)\right] \qquad (44)$$

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

[10] 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 e.g. Greene (2003)).

and substitute:

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

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

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

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

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

Finally, substitute (47) into (9) to obtain an expression for the conditional expectation for next period's leverage:

$$
\mathbb{E}(lev_{it+1}|\hat{u}^*_{it}) = \beta\phi b\cdot\hat{u}^*_{it} = \phi\underbrace{\left[1 + \frac{\sigma^2_{u_{it}}}{\beta^2}\left(\frac{1}{\sigma^2_{\eta_{it}}} + \frac{1}{\sigma^2_{x_{it}}}\right)\right]^{-1}}_{=c\geq 0}\hat{u}^*_{it} \qquad (48)
$$

B.3 Residual Persistence

As before, express the regression residual as

$$
\begin{aligned}
\hat{u}^*_{it+1} &= (\beta-\hat{\beta}^*)x_{it+1} - \hat{\beta}^*\eta_{it+1} + u_{it+1} \\
&= (\beta-\hat{\beta}^*)(\phi x_{it} + \epsilon_{it}) - \hat{\beta}^*\eta_{it+1} + u_{it+1} \qquad (49)
\end{aligned}
$$

Then

$$\begin{aligned}\mathbb{E}(\hat{u}^*_{it+1}|\hat{u}^*_{it}) &= (\beta-\hat{\beta}^*)\left[\phi\mathbb{E}(x_{it}|\hat{u}^*_{it})+\mathbb{E}(\epsilon_{it}|\hat{u}^*_{it})\right]-\hat{\beta}^*\mathbb{E}(\eta_{it+1}|\hat{u}^*_{it})+\mathbb{E}(u_{it+1}|\hat{u}^*_{it}) \\ &= (\beta-\hat{\beta}^*)\phi\mathbb{E}(x_{it}|\hat{u}^*_{it}) = \left[\phi(\beta-\hat{\beta}^*)b\right]\hat{u}^*_{it} \qquad (50)\end{aligned}$$

C Implied Target Leverage Derivations

Derivation of α_0

Begin with the relationship between mismeasured target and true target:

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

Taking expectations:

$$\mathbb{E}(\widehat{lev^*}) = \alpha_0 + \alpha_1\mathbb{E}(\widehat{lev}) + 0 \qquad (52)$$

Mismeasured target and 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):

$$lev = \beta^*_0 + \beta^*_1 x^* + \epsilon^* \qquad (53)$$

Taking expectations:

$$\begin{aligned}\mathbb{E}(lev) &= \beta^*_0 + \beta_1\frac{\sigma^2_x}{\sigma^2_x+\sigma^2_\eta}\mathbb{E}(x+\eta)+\mathbb{E}(\epsilon^*) \\ &= \beta^*_0 + \frac{\sigma^2_x}{\sigma^2_x+\sigma^2_\eta}\mathbb{E}(lev) \qquad (54)\end{aligned}$$

Therefore,

$$\beta^*_0 = \mathbb{E}(lev)\left(1-\frac{\sigma^2_x}{\sigma^2_x+\sigma^2_\eta}\right) \qquad (55)$$

Mismeasured target leverage is given by

$$\widehat{lev^*} = \beta^*_0 + \beta^*_1 x^* \qquad (56)$$

Substituting for β^*_0 and β^*_1 shows that the mismeasured target equals true target (and hence actual leverage), on average:

$$\begin{aligned}\mathbb{E}(\widehat{lev^*}) &= \mathbb{E}(lev)\left(1-\frac{\sigma^2_x}{\sigma^2_x+\sigma^2_\eta}\right)+\frac{\sigma^2_x}{\sigma^2_x+\sigma^2_\eta}\beta_1\mathbb{E}(x+\eta) \\ &= \mathbb{E}(lev)\left(1-\frac{\sigma^2_x}{\sigma^2_x+\sigma^2_\eta}\right)+\frac{\sigma^2_x}{\sigma^2_x+\sigma^2_\eta}\mathbb{E}(lev) \\ &= \mathbb{E}(lev) \qquad (57)\end{aligned}$$

Substituting (57) into (52) then yields an expression for α_0:

$$\begin{aligned}\mathbb{E}(\widehat{lev^*}) = \mathbb{E}(\widehat{lev}) &= \alpha_0 + \alpha_1 \mathbb{E}(\widehat{lev}) + 0 \\ \alpha_0 &= (1-\alpha_1)\mathbb{E}(\widehat{lev}) \qquad (58)\end{aligned}$$

Derivation of α_1

Since (51) above is a regression equation:

$$\alpha_1 = \frac{cov(\widehat{lev}, \widehat{lev^*})}{Var(\widehat{lev})} \qquad (59)$$

Expanding the numerator:

$$\begin{aligned}cov(\widehat{lev}, \widehat{lev^*}) &= Cov\left(\beta_1 x, \beta_0^* + \beta_1^*(x+\eta)\right) \\ &= Cov\left(\beta_1 x, \mathbb{E}(lev)\left(1 - \frac{\sigma_x^2}{\sigma_x^2+\sigma_\eta^2}\right) + \beta_1 \frac{\sigma_x^2}{\sigma_x^2+\sigma_\eta^2}(x+\eta)\right) \\ &= Cov\left(\beta_1 x, \beta_1 \frac{\sigma_x^2}{\sigma_x^2+\sigma_\eta^2} x\right) + \underbrace{Cov\left(\beta_1 x, \beta_1 \frac{\sigma_x^2}{\sigma_x^2+\sigma_\eta^2}\eta\right)}_{=0} \\ &= \beta_1^2 \frac{\sigma_x^2}{\sigma_x^2+\sigma_\eta^2} Var(x) \qquad (60)\end{aligned}$$

Substituting:

$$\begin{aligned}\alpha_1 &= \frac{cov(\widehat{lev}, \widehat{lev^*})}{Var(\widehat{lev})} = \frac{\beta_1^2 \frac{\sigma_x^2}{\sigma_x^2+\sigma_\eta^2} Var(x)}{\beta_1^2 Var(x)} \\ &= \frac{\sigma_x^2}{\sigma_x^2+\sigma_\eta^2} \qquad (61)\end{aligned}$$

Finally, let $a = \frac{\sigma_\eta^2}{\sigma_x^2}$, and substitute into (61):

$$\alpha_1 = \frac{1}{1+a} \qquad (62)$$

Derivation of σ_e^2

Start again with (51), and compute the variance:

$$Var(\widehat{lev^*}) = \alpha_1^2 Var(\widehat{lev}) + \sigma_e^2 \qquad (63)$$

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

$$\widehat{lev^*} = \beta_0^* + \beta_1^* x^* = \beta_0^* + \beta_1 \frac{\sigma_x^2}{\sigma_x^2+\sigma_\eta^2}(x+\eta) \qquad (64)$$

Then compute the variance:

$$
\begin{aligned}
Var(\widehat{lev^*}) &= Var\left(\beta_1 \frac{\sigma_x^2}{\sigma_x^2+\sigma_\eta^2}(x+\eta)\right) \\
&= \left(\beta_1 \frac{\sigma_x^2}{\sigma_x^2+\sigma_\eta^2}\right)^2 (\sigma_x^2+\sigma_\eta^2) \\
&= \beta_1^2 \frac{(\sigma_x^2)^2}{\sigma_x^2+\sigma_\eta^2} \\
&= \underbrace{Var(\widehat{lev}) \frac{\sigma_x^2}{\sigma_x^2+\sigma_\eta^2}}_{\leq Var(\widehat{lev})} \qquad (65)
\end{aligned}
$$

Substitute (65) into (63), and solve for σ_e^2:

$$
Var(\widehat{lev}) \frac{\sigma_x^2}{\sigma_x^2+\sigma_\eta^2} = \alpha_1^2 Var(\widehat{lev}) + \sigma_e^2 \qquad (66)
$$

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:

$$
\begin{aligned}
\sigma_e^2 &= Var(\widehat{lev}) \left[\frac{\sigma_x^2}{\sigma_x^2+\sigma_\eta^2} - \left(\frac{\sigma_x^2}{\sigma_x^2+\sigma_\eta^2}\right)^2 \right] \\
&= Var(\widehat{lev}) \left[\frac{1}{1+a} - \frac{1}{(1+a)^2} \right] \\
&= Var(\widehat{lev}) \frac{a}{(1+a)^2} \qquad (67)
\end{aligned}
$$

References

Baker, M. and J. Wurgler (2002). Market Timing and Capital Structure. *Journal of Finance 55*, 1–32.

Chang, X. and S. Dasgupta (2009). Target Behaviour and Financing: How Conclusive is the Evidence? *Journal of Finance*.

DeAngelo, H., L. DeAngelo, and T. Whited (2011). Capital structure dynamics and transitory debt. *Journal of Financial Economics 99*, 235–261.

DeAngelo, H. and R. Roll (2011). How stable are corporate capital structures? *Working Paper*.

Erickson, T. and T. M. Whited (2006). On the accuracy of different measures of q. *Financial Management*, 5–33.

Fischer, E., R. Heinkel, and J. Zechner (1989). Dynamic Capital Structure Choice: Theory and Tests. *Journal of Finance 44*, 19–40.

Flannery, M. and K. Rangan (2006). Partial Adjustment toward Target Capital Structures. *Journal of Financial Economics 79*, 469–506.

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

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

Granger, C. W. J. and P. Newbold (1977). Forecasting economic time series. *Academic Press*.

Greene, W. H. (2003). Econometric Analysis. *Prentice Hall*.

Hennessy, C. and T. M. Whited (2005). Debt dynamics. *Journal of Finance 60*, 1129–1165.

Hovakimian, A., T. Opler, and S. Titman (2001). The Debt-Equity Choice. *Journal of Financial and Quantitative Analysis 36*(1), 1–24.

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

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

Menichini, A. (2010). Financial frictions and capital structure choice: A structural estimation. *Working Paper*.

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

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

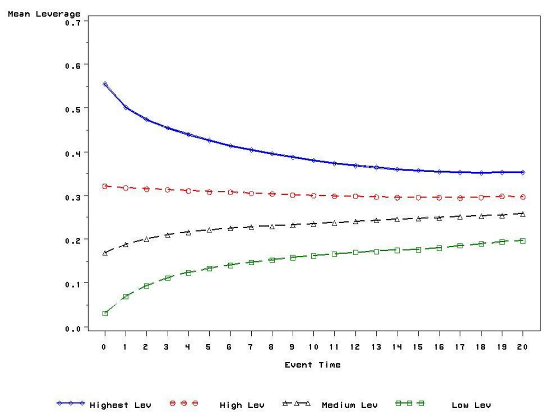
Roberts, M. (2001). The Dynamics of Capital Structure: An Empirical Analysis of a Partially Observable System. *Working Paper*.

Strebulaev, I. and B. Yang (2012). The Mystery of Zero-Leverage Firms. *NBER Working Paper*.

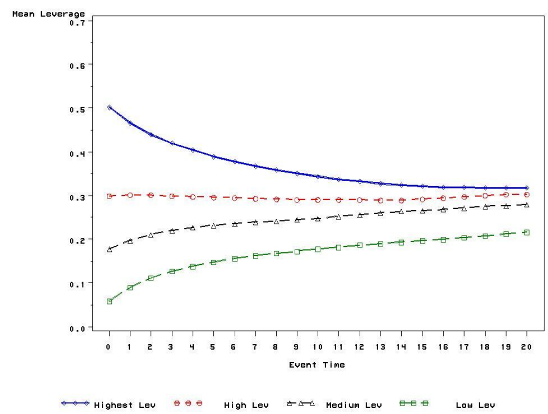
Titman, S. and R. Wessels (1988). The Determinants of Capital Structure Choice. *Journal of Finance 43*(1), 1–19.

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

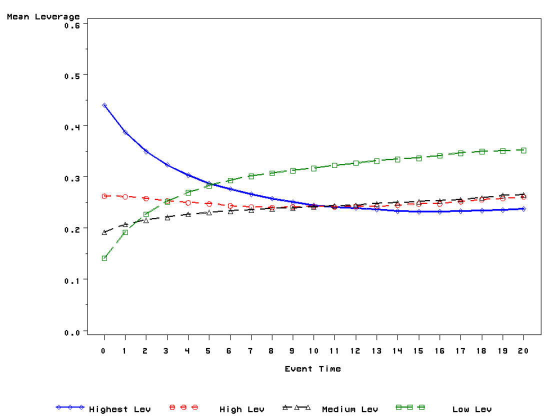
Figures



Panel A: Sort on Actual Leverage



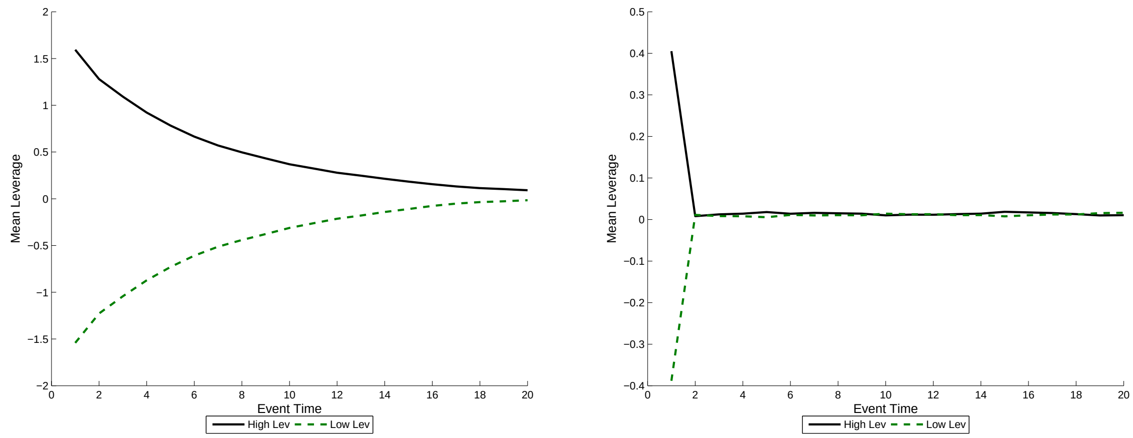
Panel B: Sort on Unexpected Leverage



Panel C: Sort on Unexpected Leverage with FE

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 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

Figure 2: 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:

$$lev_{it} = \beta x_{it} + u_{it} \tag{68}$$

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

$$x_{it} = \phi x_{it-1} + \epsilon_{it} \tag{69}$$

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.

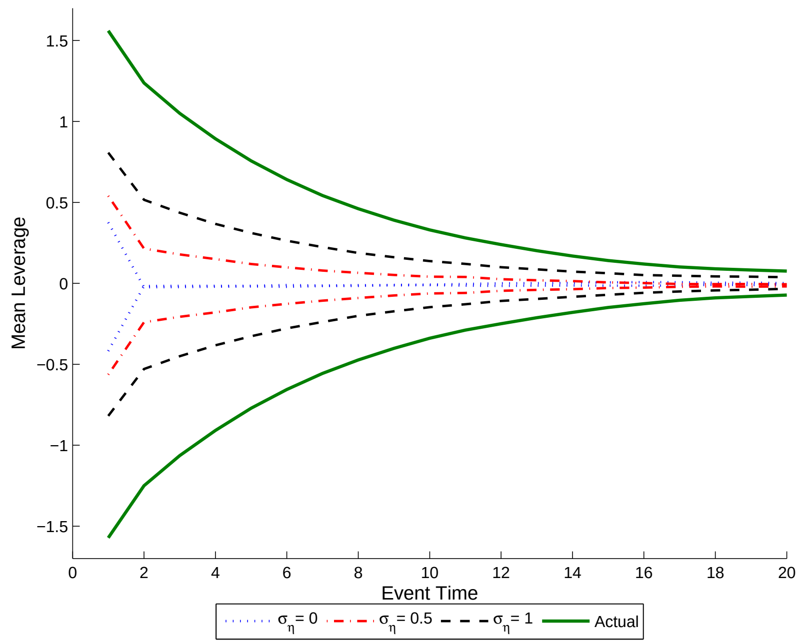


Figure 3: Comparison of Portfolio Leverage Dispersion as a Function of Measurement Error. The simulation setup is as before, e.g. in Figure 2, but only the mismeasured regressor x^*_{it} is available. I simulate the following system 5,000 times:

$$lev_{it} = \beta x_{it} + u_{it} \tag{70}$$

$$x_{it} = \phi x_{it-1} + \epsilon_{it} \tag{71}$$

$$x^*_{it} = x_{it} + \eta_{it} \tag{72}$$

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 3 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. Shown are the average portfolio leverage levels over an event horizon of 20 time periods.

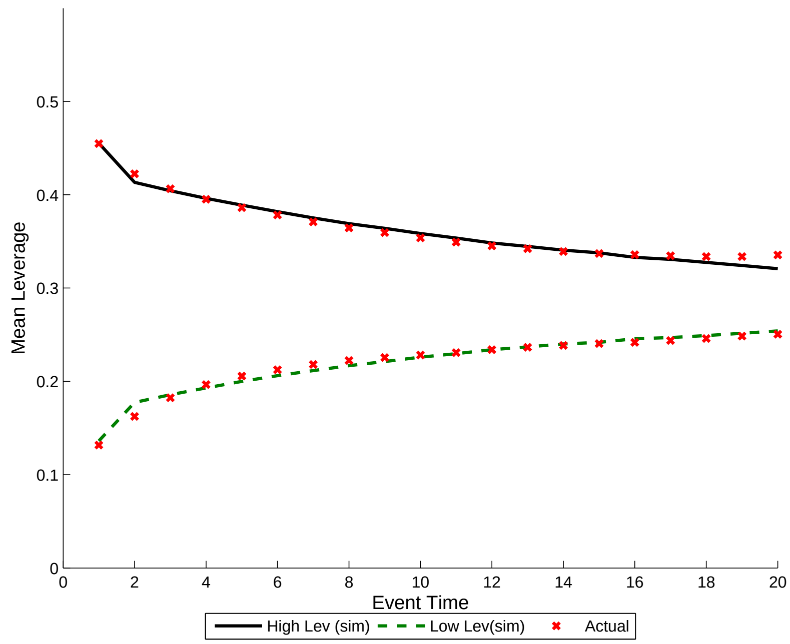


Figure 4: Data-Implied Target Leverage Dynamics.
I model leverage lev as a function of its target $\widehat{lev}$, which in turn is an AR(1) process:

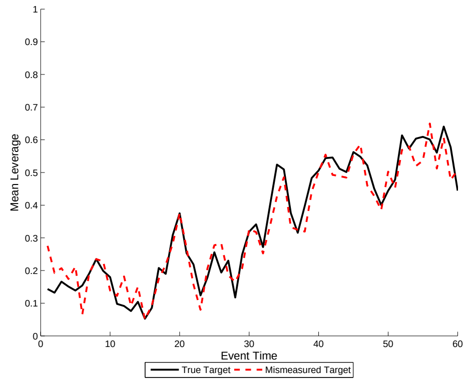
$$lev_t = \widehat{lev}_t + u_t \tag{73}$$
$$\widehat{lev}_t = \varphi_0 + \varphi_1 \widehat{lev}_{t-1} + \varepsilon_t \tag{74}$$

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 resembles the actual data points by minimizing the sum of squared deviations:

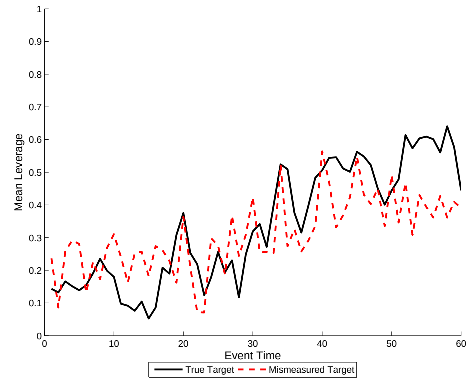
$$\min_{\Phi} \sum_i \sum_t \left(PFlev_{it}^{sim} - PFlev_{it}^{act} \right)^2 \tag{75}$$

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 (73), as well as the intercept, slope and error variance for the AR(1) process governing target leverage in (74). The estimates are as follows:

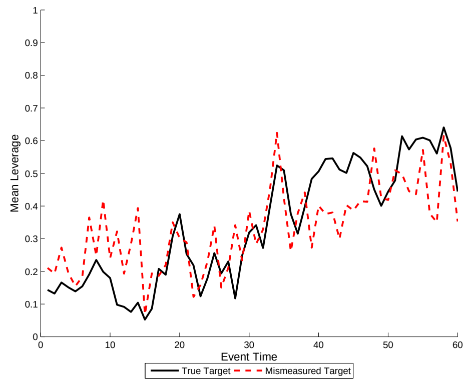
	φ_0	φ_1	σ_ε	σ_u
Estimate	0.021	0.930	0.066	0.080
Std. Error	(0.012)	(0.009)	(0.003)	(0.010)



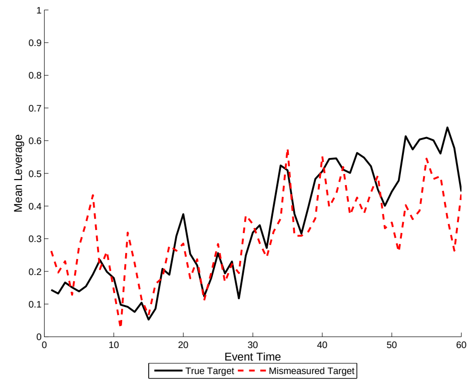
Panel 1: $a = 0.1$



Panel 2: $a = 0.5$



Panel 3: $a = 0.75$



Panel 4: $a = 1.25$

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

$$\widehat{lev^*} = \alpha_0 + \alpha_1 \widehat{lev} + e \tag{76}$$

$$\alpha_0 = (1 - \alpha_1)\mathbb{E}(\widehat{lev}) \tag{77}$$

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

$$\sigma_e^2 = Var(\widehat{lev})\frac{a}{(1+a)^2} \tag{78}$$

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.

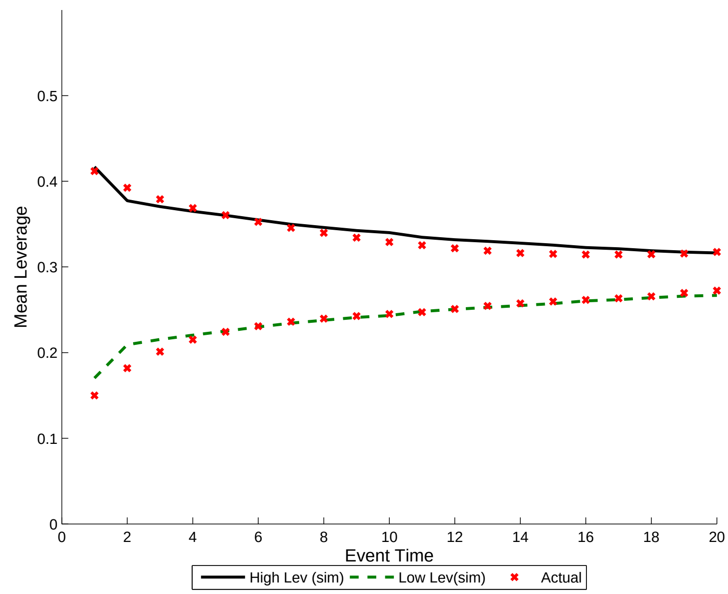


Figure 6: Implied Measurement Error from Residual-Based Sorts.
I recover the implied ratio of measurement noise to variation in the true explanatory variable x. Previously, I obtained the parameter values governing the dynamics of the true target by calibrating simulated leverage portfolios to those obtained by the Lemmon et al. (2008) leverage-based sorts. Knowing the true target then allows the mismeasured target $\widehat{lev^*}$ to be backed out via

$$\widehat{lev^*} = \alpha_0 + \alpha_1 \widehat{lev} + e \tag{79}$$

$$\alpha_0 = (1-\alpha_1)\mathbb{E}(\widehat{lev}) \tag{80}$$

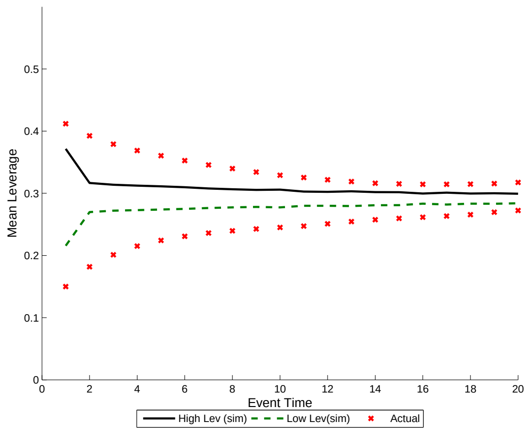
$$\alpha_1 = \frac{1}{1+a}$$

$$\sigma_e^2 = Var(\widehat{lev})\frac{a}{(1+a)^2} \tag{81}$$

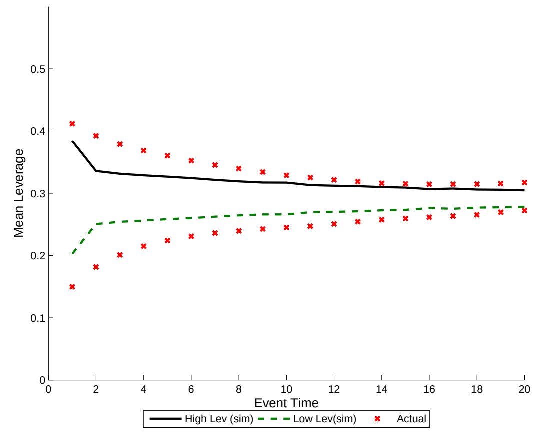
The red crosses correspond to actual portfolio leverage levels. I simulate a panel of 1,000 firms, and choose the noise-to-signal ratio $a = \sigma_\eta^2/\sigma_x^2$ for the system above such that the simulated data most closely resembles the actual data points by minimizing the sum of squared deviations:

$$\min_a \sum_i \sum_t \left(PFlev_{it}^{sim} - PFlev_{it}^{act}\right)^2 \tag{82}$$

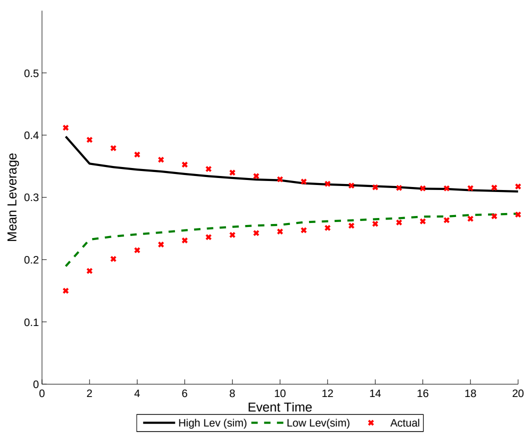
where i indexes whether a data point belongs to a high or low leverage portfolio at time t. The minimum of the objective function is reached at $a = 1.42$ (std. error = 0.12). The resulting fit is shown above.



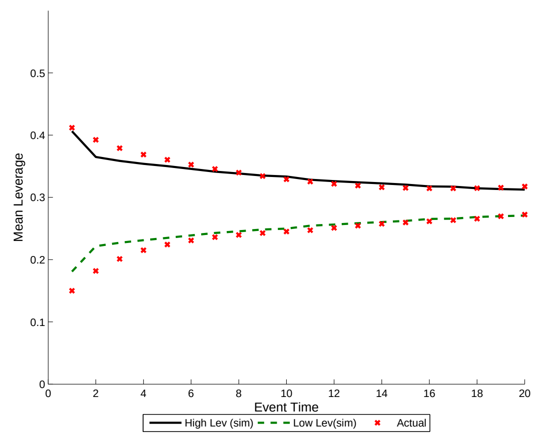
Panel 1: $a = 0.1$



Panel 2: $a = 0.25$



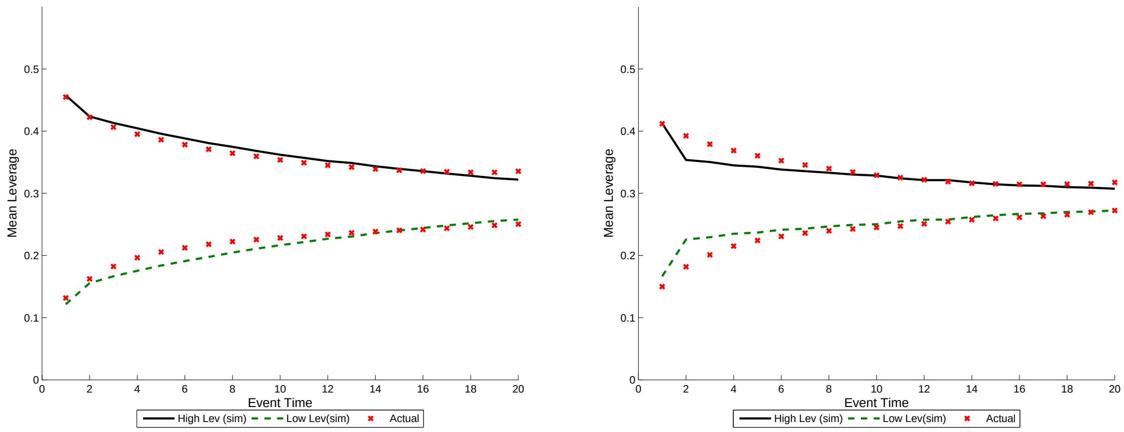
Panel 3: $a = 0.5$



Panel 4: $a = 0.75$

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:
Sort on Actual Leverage

Panel B:
Sort on Unexpected Leverage

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 Sections 4.3. 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.

Tables

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 1: Summary Statistics.
Summary statistics over the sample period 1965-2003 for nonfinancial firms on Compustat. Variable definitions are provided in Appendix A.

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	β_0	0.129	0.014
Parameters	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 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 modeled 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 chosen moments are described in Section 4.3.1.

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 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 4.3.1. The latent structural values are obtained with the estimated structural parameter values from Table 2, and are included here again for ease of comparison.

	(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

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.