

A Critique of Recent Quantitative and Deep-Structure Modeling in Capital Structure Research and Beyond*

Ivo Welch

September 5, 2012

Abstract

My paper highlights shortcomings of recent quantitative and deep-structure models in corporate finance: (1) These models have omitted too many plausible forces not based on evidence, but based on authors' priors. (2) The link between their unobserved structures and their reduced-form empirical evidence has been too weak. (Even orthogonal forces could have affected their inference.) (3) The existing tests have largely ignored many important econometric issues, such as selection and survivorship biases. (4) The models have never been held to reasonable test standards, such as performance in quasi-experimental settings. Constructively, my paper offers two suggestions: The first is to search for more direct empirical proxies instead of relying on "assumed" first-order conditions. The second is to design quasi-experimental tests of structural models. My paper illustrates these points in the context of Hennessy and Whited (2005) and Strebulaev (2007).

*I thank seminar participants at Duke, Washington, USC, UCLA, MIT, and the AEA meetings. I also especially thank Josh Angrist, Patrick Bolton, Alon Brav, Michael Brennan, Chris Hennessy, David Hirshleifer, Thad Jackson, Ross Levine, Michael Weisbach, Toni Whited, Adriano Rampini, Steve Ross, Yonah Rubinstein, and Lu Zhang for arguing with me over various (even more) flawed drafts of this critique. They helped substantially in improving this paper. I owe special thanks to Chris, Ilya, and Toni, who helped me with various aspects of the paper, and who took my critique in the right spirit. They are scholars and gentle(wo)men.

Two macroeconomic papers have had great influence on modern economics. Lucas (1976) critiqued that reduced-form models can be useless for macroeconomic policy evaluation. It pointed out that such interventions (a type of counterfactuals) may require analysis of the model in its deeper micro-foundations (structure) and not in its reduced form. (Of course, incorrect deep structure models are as susceptible to instability as reduced-form models.) Mehra and Prescott (1985) showed that simple models could fit the equity premium only qualitatively but not quantitatively. It pointed out that research should conduct at least basic calibrations to appraise whether a model can explain a phenomenon. Although deep structure and quantitative modeling need not be linked, in practice, they frequently are.

Econometrics has also progressed over the last decades. In particular, it has become clear that identification problems—the confusion in empirical work about what variables are endogenous and what variables are exogenous (Leamer (1983))—can be reduced by focusing on what is now called quasi-experimental methods (Angrist and Pischke (2008)). These methods investigate the evidence in situations in which researchers can identify shocks as “exogenously caused” based on economics. For example, instead of the overall correlation between taxes and capital structure, quasi-experimental research may investigate how a sudden change in federal tax rates influences changes in the capital structures of firms whose net income is either just above vs just below the border where they are taxed. Quasi-experiments can be viewed as examples of realized counterfactuals in Lucas’ sense.

Deep structure and quantitative models have been commanding an increasing market share of theoretical corporate finance, too. Examples of structural quantitative models in corporate finance are Leland (1994), Leland (1998), Leland and Toft (1996), Goldstein, Ju and Leland (2001), Hennessy and Whited (2005), Ju, Parrino, Poteshman and Weisbach (2005), Hennessy and Whited (2007), Strebulaev (2007), Titman and Tsyplakov (2007), and DeAngelo, DeAngelo and Whited (2010b). My paper offers a critical perspective on this quantitative and structural modeling trend in corporate finance. One can disagree about model esthetics, such as the tradeoffs between complexity, realism, and simplicity, especially when it comes to merely highlighting comparative statics and intuition. But my paper argues that there should be little disagreement that existing quantitative and deep-structure models in corporate finance have the following shortcomings:

1. Although built on plausible forces, they have ignored many other plausible forces not based on the data but based on their authors’ priors.

2. The reduced form of structural models that is put to the data often maps into a vast number of plausible alternative explanations that are consistent with the same reduced-form findings. This is particularly problematic in structural models, because even forces orthogonal with the underlying hypothesized structural forces can distort their inference (Section II.A.1). The gap between the theory and the evidence is so unusually large because the model tests lean heavily on inversions of structural first-order conditions, often linking entirely different kinds of variables together. My paper argues that, in the corporate finance context, direct empirical proxies for structural variables can often be found, and direct variables are likely to yield better tests.
3. Their tests have largely ignored other econometric issues of importance in the corporate finance context, such as selection biases, survivorship biases, controls, and so on.
4. The models have not been held to appropriately high test standards (in-difference estimation, quasi-experimental identification, out-of-sample prediction). Few tests have allowed for strong specific alternative explanations.

My paper will discuss these points in great detail. To illustrate my arguments, I will critique two prominent models, Hennessy and Whited (2005) and Strebulaev (2007). I will argue that they convey a confidence that is not justified by an objective evaluation of the evidence. However, my points are broader. I argue that it is unlikely that *any* models in their vein are likely to succeed anytime soon. To be clear, I do not argue that many simpler approach papers have not suffered from many of these weaknesses, too. However, I do argue that the shortcomings are less likely to be overcome in state-of-the-art models in the Hennessy and Whited (2005) and Strebulaev (2007) approach than they are likely to be overcome in state-of-the-art models using simpler approaches.

Some of my points echo similar ones made (independently) in recent critiques by Angrist and Pischke (2010) and Caballero (2010). Angrist and Pischke argue that structural models have failed badly in industrial organizations and labor economics, and propose that the alternative of “design-based” empirical studies (including quasi-experimental designs) has offered real tests of economic forces and causality even for the most simple of models. My paper shares some of their perspective. However, beside the domain and application differences, I argue (1) that the models can be better tested by searching for direct proxies (in lieu of the reliance on inverted first-order conditions in many current tests); and (2) that structural models and quasi-experimental approaches are not substitutes but

complements—the predictions of quantitative and structural models can and should be tested in quasi-experimental settings against reasonable alternatives. One approach does not preclude the other. Caballero (2010) also laments similar issues and the general state of affair in macroeconomics, writing “We are too far from absolute truth to be so specialized and to make the kind of confident quantitative claims that often emerge from the core.”

My paper now proceeds as follows. In Section I, I describe the basic research challenges in corporate finance in general and capital structure in specific. In Section II, I describe a premier deep-structure model in corporate finance, Hennessy and Whited (2005), and critically reevaluate the empirical evidence in its favor. In Section III, I do the same for premier quantitative model, Strebulaev (2007). In Section IV, I contrast their assessment with my own. In Section V, I explain why some areas of asset-pricing, in which there are strong arbitrage conditions (especially derivatives pricing relative to the underlying equity pricing), are better suited to the approach, while other areas (especially equity pricing) share many of the problems of corporate finance. Section VI concludes and speculates how research in corporate finance could progress.

I Corporate Finance

Corporate finance is the study of the behavior of corporations. It faces both common and unique research challenges. This section gives an overview.

It is not uncommon that dozens of economic hypotheses are trying to explain the same phenomenon. The hypotheses are rarely exclusive. Researchers rarely start from the same priors and rarely agree on the first-order effects. The publication process often favors advocacy for a particular economic force, rather than an evaluation of relative strengths and weaknesses. Few papers attempt to reconcile the evidence from multiple earlier papers.

The empirical evidence itself is often “murky.” Sometimes, the evidence is sensitive with respect to the sample, the controls, or the specification. Frequently, the research explains a only small fraction of the variance of the dependent variable of interest.

In contrast to some areas of asset pricing, there are few real-world arbitrage constraints that would allow third parties to force rational behavior onto non-value optimizing corporations. Staggered boards, poison pills, and other mechanisms can allow managers to isolate themselves effectively from any hostile external forces. Even when not in place, the takeover premium required to unseat management is often greater than the value

consequences of poor corporate policies. Darwinian selection is weak. Nevertheless, some corporate finance theories attribute an unconflicted hyper-rationality to managers that contrasts starkly with managers' generally more heuristic and conflicted decision-making process.

Although there is a large number of firms in the cross-section, the data frequency in corporate finance is also usually fairly low, with annual data being most common. Depending on the topic of the study, data availability can be spotty and induce selection biases. When firms disappear mid-year (and they do so on average after 10 years), financial statement data from their final fiscal year is rarely available. This can lead to survivorship biases. Other common problems that plague empirical research in corporate finance are strong firm-size effects that are not fully understood, residual heterogeneity across firms and industries, multicollinearity, and errors in variables due to proxy measurement problems.

On the positive side, corporations are subject to many observable "natural experiments" that can be viewed as exogenous from firms' perspectives. For example, U.S. firms have frequently been subject to tax changes, regulatory changes, aggregate or idiosyncratic technology shocks, changes in external financing availability (e.g., in 2008), personnel deaths, and so on, all of which have been ingredients in some corporate finance models.

To animate my points, my paper now focuses on one concrete subfield, capital structure. However, the reader should remain cognizant that the issues covered readily carry over into other areas of corporate finance (and beyond).

A Capital Structure Theories

Perhaps the most important question in the capital structure literature is how leverage ratios come about.^{1,2} Modigliani and Miller (1958) showed half a century ago that capital structure is not an important choice under fairly strong perfect market conditions. However, these conditions do not hold in the real world. Therefore, many papers have argued that

¹There are more nuanced views of this question. For example, Rauh and Sufi (2010) argue that this is too narrow a question, because firms also have substantial variation in the types of debt they carry. Others, such as DeAngelo, DeAngelo and Stulz (2010a) or DeAngelo et al. (2010b), have argued that capital structure should be viewed together with project choice. This is also perspective adopted in Welch (2008, Chapter 21), which considers firm-scale and capital structure jointly.

²Explaining capital structure is different from explaining corporate issuing behavior and/or payout policy. For example, Welch (2011) shows that the correlation between equity issuing activity and leverage ratio changes is close to zero.

capital structure choice can be value-relevant. Under the additional assumption that corporations or managers optimize, they should choose specific capital structures.

A sampling of forces includes:

Deadweight distress costs favor equity, because equity holders cannot force the firm into financial distress. Sometimes, this is made specific. For instance, Roberts and Sufi (2009) identify the propensity of firms to enter financial distress by the distance from bond covenants.

Corporate taxes favor debt, because the interest is tax-deductible (e.g., Graham (1996)).

First put forth by Robichek and Myers (1966), a “tax vs. distress costs” tradeoff has been the workhorse capital-structure theory. If managers are unconflicted and maximize firm value, the tax-distress cost tradeoff can lead to an optimal leverage ratio. Other proposed explanations include:

Adverse selection favors debt, because the negative signal when firms issue lower-priority securities renders them more expensive. (Adverse selection is also the main ingredient in the pecking-order theory in Myers and Majluf (1984).)

Risk shifting favors equity, because creditors may demand too many restrictions on corporate flexibility (Parrino and Weisbach (1999)).

Unmitigated agency concerns favor equity, especially when corporate governance is weak, because managers prefer less pressure and the opportunity to build empires.

ESOPs favor equity, because they can be used as efficient motivators for employees (Fama and French (2002)).

Credit ratings clientele effects favor equity especially for firms barely or just ranked investment-grade, because an investment grade rating is required to attract some institutional investor clienteles (Kisgen (2006)).

Tangibility favors debt (Rampini and Viswanathan (2010)), because it reduces the cost of creditors to recover their collateral.

Some of these are forces that can be derived from other forces. Many other economic forces can have situation-specific capital-structure implications. For example,

Frictions (transaction costs to issuing or retiring debt or equity) can induce corporations to be inert most of the time (Fischer, Heinkel and Zechner (1989)).

Stock returns can induce higher or lower market leverage ratios (Welch (2004)). Analogously, operating performance and depreciation rules can induce higher or lower book leverage ratios.

Timing considerations can induce managers to issue debt or equity when they believe that one is cheaper than the other (Baker and Wurgler (2002)).

Consensus concerns can induce managers to issue equity when they are in agreement with their investors and debt otherwise (Dittmar and Thakor (2007)).

M&A capital needs can induce firms to issue securities as needed to facilitate acquisitions (Fama and French (2002)).

Peer effects can influence firms to imitate their industry peers when their capital structures have diverged too much (Welch (2004), Roberts and Leary (2009)).

Project considerations, as in DeAngelo et al. (2010a) (cash needs), in DeAngelo et al. (2010b) (lumpy investment), in Hennessy and Whited (2005) (productivity), or in Shivdasani and Stefanescu (2010) (pension liabilities) can make it relatively more advantageous to issue short-term securities or long-term securities (of which equity is one).

Crowding out by government debt issuing activity can reduce corporate borrowing (Greenwood, Hanson and Stein (2010)).

Precommitments (sinking funds) can favor equity if external financing conditions make it difficult to find equity funding at the time when debt comes due.

Hubris may induce firms to take on too much leverage (Roll (1986), Malmendier and Tate (2005)).

Managerial preferences, measurable as the identity of the manager, may create an idiosyncratic preference for either debt or equity (Bertrand and Schoar (2003)).

Investment bankers may convince CFOs to shift around their capital structures in order to generate fees (and allow the CFOs to look active).

Non-optimal behavior could lead firms to follow arbitrary rules, in which they may not pay attention to bankruptcy costs, tax benefits, transaction costs, and project funding needs, but to something else that may or may not be reasonable and/or observable. For example, they may target a bogie, such as book leverage ratios, even if/when

these are not meaningful for measuring tangible assets or assessing future distress costs. Managers could be “asleep at the switch,” and readjust whenever they wake up (or when a heuristic band is violated).

Many of these economic hypotheses have been published with accompanying favorable empirical evidence. This list of influences is of course not exhaustive, nor are the forces mutually exclusive. They may operate simultaneously, and they may influence different firms differently at different times. The influences may be complementary or substitutive. They may be non-linear.

Tests of any of these research theories can be against specific alternative hypotheses or against unspecific null hypotheses. One reasonable null hypothesis is that firms behave as they did in the past. Another reasonable null hypothesis is that firms behave randomly, i.e., without regard to their own characteristics. Both null hypotheses are necessary, but not sufficient. Beating them are weak tests that any specific theory should be able to pass.

The obvious next question is how research can choose among forces. Ideally, empirical evidence would allow us to assess the relative importance of all these economic hypotheses and how they correlate. Subsequent research could then make informed choices about which economic forces a useful model of capital structure should include and which it can (more or less) ignore. Unfortunately, embedding more than a few hypotheses within the same modeling framework has proven to be difficult. It is thus rare.³

The more common research approach is to rely on priors to identify the subset of forces that are assumed to matter. The choice can be based on economic intuition, or it can be informed by surveys of the individuals that actually decide on the capital structures (Graham and Harvey (2001)), or it can be justified by results from earlier studies. The maintained assumption in tests with ex-ante variable selection is that omitted economic forces are orthogonal to included *measured* variables. Therefore, *ceteris paribus*, when a variable is omitted that has only a weak influence on capital structure, this variable is also less likely to affect inference with respect to the force the research is exploring. In the extreme, if a variable is uncorrelated with capital structure and its changes, and if it is

³An early piece that explicitly tested multiple theories together was Titman and Wessels (1988). Shyam-Sunder and Myers (1999) have entertained the pecking order theory as an embedded competitor—and promptly found it to perform as well as the tax-distress-friction model. Fama and French (2002) test the tradeoff model against the pecking order model; and Huang and Ritter (2009) test the static tradeoff model, a pecking-order model, and the market timing model competitively. Other papers have handled competing explanations implicitly (e.g., with fixed effects), usually in order to control for variables that the researchers are not interested in.

uncorrelated with the independent measured variables, then it has no distortive influence on inference about other forces.

B Other Specification Issues

Empirical capital structure tests also have to struggle with other common problems in this domain, such as:

- Empirical proxies are rarely exact theoretical constructs. For example, empiricists can often measure only average values, even though the theory calls for marginal values. (Tobin's Q is a good example.)
- Firm size, for reasons not fully understood, can be an important determinant.
- Industry, for reasons not fully understood, can be an important determinant.
- Time, for reasons not fully understood, can modulate the strengths of different relations.
- Many firms lack the data necessary to investigate the target hypothesis. The missing data may not be random.
- Public firms last on average only ten years, and final fiscal year data is rarely available. Stock-return data and/or a code that describes the reason for disappearance between the last reported financial statement and the day of disappearance is available, but unfortunately rarely used. Attributing a 100% leverage ratio to firm-years in which stock returns were dismal during the year but the firm never reported a final financial statement would be a reasonable check.
- Many “independent” corporate finance variables seem themselves endogenous.
- Variables may require scaling, though it is not clear whether this should be done by adjusting for assets, tangible assets, sales, or even net income.
- Many independent variables, especially financial statement metrics, are often highly correlated, even after firm-size normalization.
- Because leverage ratios are bounded variables with unusual dynamics, a resampling approach (as in Chang and Dasgupta (2009), Iliev and Welch (2010)) is necessary to establish the properties of common linear estimators whose properties are often familiar only in the context of normally distributed residuals.

Empirical papers can employ tests to assess whether these issues are important, and, if they are, attempt to correct for them. For example, they can perform residual diagnostics. This can lead to the inclusion of further control variables, such as firm-size, time measures, industry controls, or fixed effects. Related to fixed-effect econometric models, some research strategies focus principally on estimating differences, not levels.

B.1 Isolating Managerial Capital Structure Changes

Capital structure theories typically focus on the actions of firms, but leverage ratio changes (dct, “debt-to-capital, total”) are also influenced by corporate performance. A good approach is therefore to take out the part of leverage changes that is due to corporate performance (which I shall call dcp, “debt-to-capital passive”), and focus on predicting the remainder (which I shall call dca, “debt-to-capital active”).

$$\begin{aligned} \text{dct}_{t-1,t} &= \frac{D_t}{D_t + E_t} - \frac{D_{t-1}}{D_{t-1} + E_{t-1}}, \\ \text{dcp}_{t-1,t} &= \frac{D_{t-1}}{D_{t-1} + E_{t-1} \cdot (1 + x_{t-1,t})} - \frac{D_{t-1}}{D_{t-1} + E_{t-1}}, \\ \text{dca}_{t-1,t} &= \frac{D_t}{D_t + E_t} - \frac{D_{t-1}}{D_{t-1} + E_{t-1} \cdot (1 + x_{t-1,t})}, \end{aligned} \quad (1)$$

where D is debt, E is equity, and x is the capital gain of equity over the year. dca is the net effect of all managerial debt and equity issuing and repurchasing activity, including dividends and coupon payments, during the year. A theory of corporate behavior should not be capable of explaining dct as much as it should be capable of explaining dca. (Managerial actions, dca, can of course respond to performance, too.) Moreover, empirical tests with dca instead of dct eliminate not only stock-market induced noise, but also biases arising from stock-market return regularities (such as the book-to-market effect).^{4,5}

⁴Note that a theory could seek to explain on even smaller set of changes, such as equity created by ESOPs or equity not created by ESOPs. This would argue for taking out additional components from dca. However, the procedure of taking out the stock-return induced capital structure change (when evaluating managerial actions) is conservative and thus should be uncontroversial.

⁵The evidence of poor predictability of dca (reported below) is the same if the rate of return instead of the capital gain is used as the performance measure, and/or if debt-due-within-one-year is removed. Not surprisingly, the noise is usually even stronger if dct instead of dca is predicted.

C Empirical Test Approaches

C.1 Levels, Differences, and Quasi-Experimental Tests

In-Sample Evidence: Early tests in the capital structure literature, such as Titman and Wessels (1988), focused on empirical moments in levels. They established such regularities as the negative correlation between profitability and leverage. If the model is correct, an in-sample test uses the data most efficiently. Moreover, in cases in which the dependent variable is conditionally normally distributed (although this is unfortunately not the case for leverage ratios), there is a well-established econometric toolbox for in-sample tests under linearity assumptions. Of course, if the model does not contain the correct variables and/or if the specification assumptions are violated, then the inference can be incorrect.

In-Difference Evidence: Level correlations can easily lead to misidentification—in the Titman and Wessels example, profitability and leverage may be codetermined by another variable. In particular, simultaneous trending of variables can suggest spurious correlations that do not exist. To improve on level tests of capital structure theories, and to reject the null hypothesis that firms behave like they have always behaved, formulating the model in changes is more robust. If a stable panel model holds in levels, it should also hold in differences ($y = a + b \cdot x \Leftrightarrow \Delta y = b \cdot \Delta x$). Suitably adjusted, a test in differences does not present a higher hurdle than a test in levels for a stable model. Of course, estimation in differences ignores some information that estimation in levels retains, and, with a different dependent variable, one cannot expect as high an R^2 for a model in differences as one can expect for a model in levels.

Quasi-Experimental Evidence: However, even regular changes in the dependent and independent variables may be codetermined by an omitted variable, and/or the direction of causality may still be the reverse of what was hypothesized. Thus, many recent tests have shifted to quasi-experimental (QE) methods.⁶ These methods focus on explaining empirical moments in “specific difference” situations only—unusual circumstances, in which the economics of the situation makes it clear that some input variable has experienced a quasi-

⁶Angrist and Pischke (2010) list the main econometric tools as instrumental variables techniques, regression discontinuities, and diff-in-diff. (Out-of-sample prediction, tests of models in differences, and a focus on natural experiments can be viewed as earlier attempts to deal with the same problem.) Their paper appears in the Journal of Economic Perspectives, with comments by Leamer, Sims, Stock, Nevo, and Keane, as well as a de-facto response by Einav and Levin (2009).

exogenous shock. Good candidates for quasi-experimental studies are sudden discrete changes in an input variable of interest, especially if it affects different firms differently.⁷

Before the advent of such quasi-experimental focus, it was often considered impossible to obtain causality inference empirically, leaving researchers with only one method to establish causality: strong faith in the conjectured theory. However, if a quasi-experiment can come close to the ideal of a randomized double-blind real experiment, it can be almost a direct causality test with respect to the intervention and even without a strong deep underlying model. Of course, the credibility of the empirical test depends on the quality of the experiment—and, of course, if the ideal is not met, then the research still needs to adjust for the influences of other simultaneous events and other misspecification concerns.

Although the quasi-experimental approach has made empirical analyses with even verbal theoretical underpinnings more believable, it has also made empirical work more difficult and constrained. Quasi-experimental tests are feasible only when realized counterfactuals can be found. Some authors, notably Deaton and Heckman, have accused quasi-experimental economics to be driven more by the search for good experiments rather than by the search for answers to good questions, or the search for first-order economic forces.

Fortunately, a shortage of good theory-relevant available quasi-experiments is not (yet) a problem in corporate finance. There have been many identifiable exogenous shocks to the inputs of many corporate finance models over the last fifty years. For example:⁸

Taxes changed repeatedly. These changes also affected different firms differently.⁹

Transaction costs changed repeatedly.¹⁰

⁷Event-studies (Fama, Fisher, Jensen and Roll (1969)) were an early form of quasi-exogenous identification, in which research exploited sharp, unique events. It is usually reasonable to assume that the stock price on a given day is endogenous to an (exogenous) news release, and not the other way around.

⁸Slower changes in many of these variables can allow for in-difference estimations, even though these are more prone to misidentification. For example, the costs of issuing have slowly decreased over the decades and are higher in bear markets than in bull markets. Productivity shocks may have diffused slowly through some industries, and research can exploit different diffusion speeds in different industries in different years.

⁹For example, Hennessy and Whited (2005) mention one possible experiment: they begin their sample in 1993, because Graham (2000) showed that different tax parameters applied earlier. Thus, the Hennessy-Whited model can be tested based on how its inference changed between pre- and post-1993 data. Givoly, Hayn, Ofer and Sarig (1992) looked at the 1986 TRA, which is another natural laboratory.

¹⁰For example, on December 1, 2005, the SEC made it easier to do shelf-registrations for firms with more than \$700 million in equity, or firms that have issued more than \$1 billion in registered non-convertibles. (This practically invites a discontinuity test approach.) In 2008, the disappearance of many investment banks and the differential capital constraints on the surviving investment banks raised the costs of issuing, especially for firms that did not have precommitted line of credits.

Financial distress costs fell in 1986, when prepackaged bankruptcies began to appear. In 1994, the bankruptcy code was overhauled.

Productivity changed dramatically in different industries at different times.¹¹

Financial constraints during the 2008 financial crisis affected firms differently based on their contractual background (Almeida, Campello, Laranjeira and Weisbenner (2009)) and/or based on their existing banking relationships.

There have also been many idiosyncratic exogenous shocks, such as cases in which key personnel suddenly died or became incapacitated.

Quasi-experimental tests also provide an appealing unconditional null conjecture: Any theory should predict better if the known quasi-exogenous shock to the presumed-exogenous variable is *not* ignored. This test is powerful, because the model can now be rejected even if “firms behave as they always have.” A quasi-experimental test is also stronger than a simple in-differences test, because the research should detect changes only in a subset of clearly identified cases (years and firms), and not in others.

C.2 In-Sample and Out-of-Sample Tests

In-sample prediction is often more susceptible to overfitting than out-of-sample prediction. This makes an out-of-sample test a good complement to an in-sample test. Out-of-sample tests lean less on the assumption that the model is correct, and make it easier to assess how stable and predictively useful a model is. Suitably adjusted for sampling variation and estimation uncertainty, an out-of-sample test does not present a higher hurdle than an in-sample test for a stable model.

Like in-sample tests, out-of-sample tests also make it easy to contrast competing models, whether they are nested or competitive. Models can be compared based on a neutral criterion, such as mean-squared-error prediction or the proportion of the variance that a set

¹¹For example, in the transportation sector, costs fell with the wider introduction of GPS, pagers, mobile email, satellite navigation, and containers. (Again, geographic, firm-specific, and industry-specific heterogeneity further strengthen such tests.) SOX imposed new record-keeping costs on many firms, reducing productivity for firms just above the threshold (Iliev (2010)). DNA tests and cures (such as antibiotics) appeared at specific times and lowered the costs in particular sectors of the healthcare sector. Note also that even when the variables of interest (here productivity) are not easily measurable, exogenous shocks to them may be much easier.

of independent variables can explain. (The latter also allows ascertaining which variables are more important [first-order] than others.)

Note that level, differences, and quasi-experimental tests can all be performed either in-sample or out-sample. For the latter, the model parameters would be fit up to the point of the change, and the model would then be used to predict subsequent changes.

II Structural Modeling

Most of the rest of this paper evaluates two specific models in light of the formidable research challenges in corporate finance that were described in the previous section. Hennessy and Whited (2005) is principally a structural model. It leans strongly on its first-order conditions to identify the influence of an unobservable “structural” variable (productivity) on observable variables (liquidity and capital structure). Hennessy-Whited highlight mostly qualitative implications of the model. Strebulaev (2007), discussed in the next section, is somewhat less structurally and more quantitatively oriented. Both papers may well be the most prominent ones in their respective genres. Both won Brattle prizes for the best paper in corporate finance in the Journal of Finance in their respective years. This makes them excellent targets for this critique.

This section discusses structural modeling.¹² My main critique here revolves around the gap between the unobservable theoretical variables and the observable tested relationships (variables). When variation in an underlying unobservable variable is used to infer the variation in empirical variables that are very different and potentially subject to other forces—tied together by hypothetical assumed-correct model first-order conditions—this gap can be huge.¹³

A Unobservable Variables and Test Inference

The main problem when the gap is wide is that multiple underlying models can map into the same empirical reduced-form model. The wider the gap between the structure and reduced-form, the more the research must lean not only on the structural model itself

¹²There are a number of possible definitions of structure (see the appendix), but for purposes of my paper, I do not need to take a stance.

¹³Of course, large gaps could also be caused by excessive noise in simple empirical proxies.

when interpreting the reduced-form model findings, but also on the belief that there are no (or few) plausible alternative explanations for the same evidence.

A.1 Orthogonality and Inference

It is sometimes not recognized that the well-known irrelevance of orthogonal forces holds only in the empirical reduced form. This is worth explaining. For example, consider a scenario in which there are two possible variables, r (right) and w (wrong), and one dependent variable y . Further assume that r and w are orthogonal, $r \perp w$. Assume that the true model is

$$\text{True Model } y = r .$$

w has no influence on y ($w \perp y$).

Now assume that the researcher *incorrectly* believes that w could determine y and puts this hypothesis to the test. If w is observable, then the model is its own empirical reduced-form. The test regression

$$\text{Estimated Model } y = \hat{a} + \hat{b} \cdot w$$

is well known to yield an expected value for \hat{b} of 0, because $w \perp r$. The researcher would correctly reject the wrong model.

Now, assume that there is another observable measure $M = f(r, w)$, specifically $M = w + r$, but the researcher's variable w is not observable (and, of course, the researcher has not thought of r , so it is an omitted variable). With w unobservable, the researcher's model is structural in the sense that it relies on an assumed model relation. (The same inference applies if M is tied to w and r through a first-order condition.) To favor the researcher, assume she knows for sure (and does not merely hypothesize) that $M = w + \epsilon$, where ϵ is uncorrelated with w . This is of course correct in the example (because $M = w + r$ and $r \perp w$). It is easy to verify that the test regression

$$\text{Estimated Model } y = \hat{a} + \hat{b} \cdot M$$

is expected to yield a positive \hat{b} . This researcher would incorrectly conclude that w influences y and thus accept her model.¹⁴ This is the case even though $w \perp r$ and $w \perp y$. Orthogonality between the true force r and the incorrect force w is not enough to guarantee

¹⁴This is a Type-1 error. It is easy to construct other examples in which other unobservable uncorrelated variables induce Type-2 error, instead.

the correct inference that $w \perp y$. Instead, for the inference to be correct, i.e., to yield the inference that $w \perp y$, it must be the case that $w \perp M$. Consequently, “holding other forces constant,” i.e., the standard *ceteris paribus* qualifier, means something quite different in the structural model than it does in the reduced-form model. It means uncorrelatedness of other forces with M , not with r .

This problem is more significant when the gap between the measured variable M and the unobserved structural variable coming out of the model is wide. Call x the researcher hypothesized variable influencing y . In the extreme, where the gap is zero, $M = x$ and we are back to the case where the researcher observes the model variable. Hence, if the model is wrong and $y \perp M$, then $E(\hat{b}) = 0$, and the inference is correct. But if the gap is wide, $M = x + z$, where z is a large wedge between the model proxy and the researcher’s variable x , even with $r \perp z$, any non-zero correlation between z and y distorts the inference.

A.2 Remedies

By definition, if the research does not measure the underlying structural variables, then it is difficult to test empirically (establish based on data ex-post) whether the model was based on the correct (unobserved) structure. In the $M = r + w$ example, no econometrics can distinguish whether it was r or w that determined y . Instead, models that hypothesize about the unobservable variable (r) must lean hard on the assumptions of the model *and* on the assumption that there are no other influences on the reduced form specifications. Thus, a test of a structural model is more likely to provide correct inference if it is *a priori certain* that the structure is correct, and that there are few alternative forces that could drive a wedge between M and r . In some areas of finance, e.g., in the pricing of derivative securities, this is a fairly credible conjecture.

Unfortunately, in corporate finance, correct structure is rarely certain. Structure in this literature is often derived from a behavioral theory’s first-order condition, and corporate behavior is not subject to third-party arbitrage. If the strength of the relations between observed variables in corporate finance are any guide to the strength of relations between unobservable and observable variables, these correlations may well often be in the single digits. (Low explanatory power of variables will also be apparent in the figures below.) If variables are uncorrelated, and the correlation between r and M is 10% and the correlation between r and y is 10%, then the expected correlation between observables M and y if the theory is correct would be only 1%. An optimistic way to interpret such empirical evidence is that the 1% implies much stronger effects between the unobserved variable

and the proxies. This will be the inference from inverting conditions—the model would seem quite powerful. A pessimistic way to interpret the same evidence is that the same 1% correlation has left a 99% gap for other variables (w) to slide in and distort the evidence.

Cautious corporate finance researchers can consider the following:

Narrowing the gap:¹⁵ Ideally, the research would identify as close a reduced-form proxy as possible for the structural variable. In many cases, even a bad direct-measure proxy is not only better than none, but better than inverted first-order-condition proxy from a behavioral hypothesis. In the example, this would suggest searching for a direct measure $w + \epsilon$, in which the error epsilon is not correlated with r .¹⁶ In some cases, it may not be possible to find a good proxy for a structural variable, but it may be possible to identify a good proxy for measurable changes in structural variables.

For the specific example of Hennessy and Whited (2005), while corporate productivity is difficult (but not impossible) to measure, the introduction of new technologies as positive shocks to productivity in particular industries is almost surely measurable. If a researcher believes the theory to be true, the additional data work should be worth the effort. My paper argues that a test of a productivity-based theory with productivity data would likely be more credible than a test of the theory based on a correlation between liquidity and capital structure.

Increasing test stringency: Research can view evidence in favor of deep structure models with wide gaps more skeptically. Sometimes, more stringent empirical tests or a higher bar can help. Although in the example, no empirical test could help distinguish between w and r , in other cases, there may be functional relationships or more empirical moments that can reduce the set of alternative theories that can explain the same set of reduced-form findings.¹⁷

¹⁵In econometrics, instrument misspecification is an analogous problem. In a deep-structure models—and especially in models without closed-form solutions, in which the economic intuition is complex and fit is low—such misspecification can be severe. The question of finding a better instrument is analogous to narrowing the gap between the observed and structural formulation.

¹⁶An even better approach may be to extract one proxy from the first-order conditions and another from direct measurement. The researcher could then report their correlations.

¹⁷This can conflict with the popular view that not all implications of a theory need to hold. Note that this is also a stricter view than that in Friedman (1966, p.8) that a theory need perform only in the arena for which it was designed. I interpret Friedman's view to mean that a model does not have to offer implications on every economic phenomenon, not that when a model has strong implications on other variables, these need not hold.

Enumerating alternative interpretations: Papers can enumerate plausible alternative explanations for the observed reduced-form relationships. If there are too many plausible influences to enumerate, the empirical evidence in favor of the deep-structure model is weak (move one's prior very little), even if the estimated moments all have the correct sign.

B The In-Sample Performance of the Hennessy and Whited (2005) Model Revisited

The Hennessy-Whited project funding hypothesis is principally a deep structure model. (It is not principally a quantitative model in the sense that its authors highlighted the qualitative implications more than the quantitative implications.) The theory is based on (unobserved) productivity shocks, which have to be backed out (through the model first-order conditions) from variables that are economically quite different. Both capital structure and observable liquidity invert back into the unobservable productivity that the model is built on. Although the model's forces are intuitive (and the model provides a wealth of moment predictions), it is not easy to understand. Important tax assumptions (Lewellen and Lewellen (2006)) work in concert with productivity shocks. The model is not solved in closed-form.

The model is too involved to allow a brief characterization here. Thus, to reassess its empirical meaningfulness, my paper follows Hennessy and Whited (2005) itself. On page 1131, they state

We highlight the main empirical implications. First, absent any invocation of market timing or adverse selection premia, the model generates a negative relationship between leverage and lagged measures of liquidity, consistent with the evidence in Titman and Wessels (1988), Rajan and Zingales (1995), and Fama and French (2002).

(The model's broader conceptual point was to help the profession understand that a negative relation between liquidity and capital structure can arise even in the absence of transaction costs and asymmetric information.) In my earlier notation, the unobserved structural productivity variable (r) drives the empirically observable reduced-form, with two variables, liquidity (M), and capital structure (y).

I will now present empirical evidence that the empirical link between liquidity and capital structure changes is weak. This is not a direct test of the model, as much as it is evidence to assess whether liquidity or changes therein are a first-order influence explaining managerial capital-structure activity, dca (eq. 1). This is the empirical regularity which Hennessy and Whited (2005) suggest should help shift the reader's posterior in favor of their model.

The first question is whether managers actively delever when their operating cash flows are high. Of course, an empirical “cash-flow vs. active leverage changes” relation would hardly qualify as strong evidence in favor of an optimal-behavior-to-productivity-shocks theory of capital structure. Any theory in which firms use operating cash flows partly to pay back (or assume less) debt could predict this. But such evidence would be in line with their productivity theory prediction. The left panel in Figure 1 shows that the cross-sectional relation between *active* managerial changes (dca) and operating cash flows is strongly negative (with a T of -35), as predicted, but with an R^2 of only 1.2%. 98.8% of the variation in managerial capital structure activity remains unexplained.

Although Hennessy and Whited (2005) use cash flows in their own tests (presumably more related to exogenous productivity shocks experienced by firms than the more endogenous cash holdings), Bolton, Chen and Wang (2010) suggest that cash balances are more suitable. A firm that already has a lot of cash on hand can reduce its leverage ratio, even if it has no particularly positive cash flow. Thus, the second question is whether managers actively delever when their cash holdings are high. Of course, an empirical “cash-holdings vs. active leverage changes” relation would again hardly qualify as strong evidence in favor of an optimal-behavior productivity shock theory of capital structure. Any theory in which firms pay more interest on loans than what they can earn on cash would suggest that managers should pay off some of their own debt—lend money to themselves instead of to others. Similarly, any theory in which well-performing firms both pay down some debt and accumulate some cash can predict this. The middle panel in Figure 1 shows that there is no relation between *active* managerial changes (dca) and cash holdings normalized by capital. The R^2 is practically zero.¹⁸

¹⁸In other specifications in earlier Compustat samples, I obtained a T statistic of -11 and an R^2 of 0.0008. The point of my test is simply that the relation is very weak and noisy, not that it is not negative.

The third question is whether managers actively delever when their cash holdings increase. Omitting the obvious alternative explanations, the right panel shows that there is no relation here, either, with an R^2 that is practically zero, too.¹⁹

It is important for the reader to realize that, even if the explanatory power had been superior, for such empirical evidence to move beliefs greatly in favor of Hennessy and Whited (2005), would also have required a belief that other variables—orthogonal to productivity or not—would not have influenced liquidity (M). This is simply not plausible.

Therefore, as already mentioned, my constructive suggestion for future tests is to search proxies for empirical productivity shocks. For example, the introduction of GPS into the shipping industry is less likely to be correlated with other forces listed in Section I.A than liquidity is (the indirect “first-order-condition inversion approach”). Such tests would seem to speak more to whether productivity influences capital structure than tests that relate liquidity and capital structure. Note also that a productivity proxy could be used to test not only the hypothesized influence of productivity on capital structure, but also the influence of productivity on liquidity.

Hennessy-Whited continue with

Second, even though the model features single-period debt, leverage exhibits hysteresis, in that firms with high lagged debt use more debt than otherwise identical firms. This is because firms with high lagged debt are more likely to find themselves at the debt versus external equity margin.

In its most stringent interpretation, this is the hypothesis that capital structure has a persistent component—that leverage ratios are not independent draws. It is difficult to think of any theory in which this would not be the case. Such a test would not be very informative.

A more relaxed and thus more interesting interpretation is the question of how levels of capital structure correlate with changes of capital structure. The prediction then is the exact opposite of readjustment, although it may apply only to high-debt firms. The left panel in Figure 2 shows that the empirical relationship is negative, not positive.²⁰ Still,

¹⁹When the dependent variable is dct instead of dca , the R^2 increase to 0.42%, mostly explained by the fact that firms that have a good year experience increases in both their cash holdings and their stock prices.

²⁰The reader should not draw a strong inference from the negative sign. Iliev and Welch (2010) document that the relation between lagged leverage and current leverage is very weak. The negative pattern in the right panel here may well have come about only [a] because firms with zero lagged leverage can only increase leverage, while firms with 100% lagged leverage can only decrease leverage; and [b] because the stock-induced change, dcp , is also a function of the past leverage ratio.

it leaves 96% of managerial capital structure actions unexplained. Another interesting question is whether there is much hysteresis in dca itself. The right panel shows that this is not the case. Managers previous year's actions are uninformative about their current actions.

C Assessment and Further Challenges

The empirical evidence suggests that even if liquidity has a marginal influence and even if there is some hysteresis, simple measures of either leave almost all capital-structure relevant managerial activities unexplained. This does not mean that Hennessy and Whited (2005) can be rejected. If the productivity link to liquidity and lagged leverage is noisy, the theory may even predict weak relationships among reduced-form observables. My point is not that the data reject the model. Instead, my points are first that the correlations between the empirical variables are not so strong as to give the reader great confidence that the model has captured the first-order determinants of managerial capital-structure decisions; and second that, even if the correlations had been strong, they could have easily been driven by very different forces, orthogonal to the productivity forces assumed to be important. To the extent that the burden is on a theory—in this case that unmeasurable productivity differences systematically drive the observed correlations—the existing empirical evidence in favor of Hennessy and Whited (2005) seems almost non-existent. In the absence of better empirical support, the Hennessy and Whited (2005) model should not be a primary lens through which capital structure should be viewed.

Hennessy and Whited (2005) also did not consider other challenges enumerated in the previous section. Their paper focused on specific forces (productivity, funding needs, and taxes) *a priori*, and thus ignored many other potentially important forces. Liquidity and leverage may well be related to other economic forces that were described in Section I.A. Hennessy and Whited (2005) offered few controls or residual diagnostics. It did not consider the effects of selection and survivorship biases. It omitted firms with low assets, capital stocks, or sales, leaving only about 20% of the total Compustat population in its tests. It did not offer estimation in differences. It did not attempt to reduce the depth of its structure. As explained, although it is true that productivity is difficult to measure, even a bad direct measure could have helped pin down productivity as the most likely influence (compared to what is in effect no direct empirical measure). It offered no in-differences evidence—when *changes* in leverage ratios would be expected. It offered no quasi-experimental tests. As mentioned in footnote 11, many sudden, large changes in

productivity were measurable—and, even better, varied across different firms and industries in different years—as were changes in tax rates. Such evidence could greatly aid the reader in assessing the empirical validity of the theory. It offered no out-of-sample evidence.

III Quantitative Modeling

Strebulaev (2007) is a quantitative model that puts forth a dynamic version of the tax-distress-friction tradeoff theory in Fischer et al. (1989), in which corporate taxes favor debt, distress costs favor equity, and frictions favor inaction. In addition to its quantitative aspects, Strebulaev (2007) points out that the tax-distress-friction framework predicts that when firms experience positive profitability shocks, and they do not adjust their capital structure (due to transaction costs), more profitable firms can have relatively more equity in their capital structures. This correlation was widely and incorrectly viewed as inconsistent with the tradeoff theory. Although this is an important insight in itself, because this intuition can also be obtained in a one-page sketch model (or using Fischer et al. (1989)), it is usually the quantitative aspect of Strebulaev (2007) that is viewed as its hallmark.

A A Brief Sketch of the Strebulaev (2007) Model

In Strebulaev (2007), firms can optimize their capital structures dynamically. The complexity of the model is an order of magnitude higher than that of its predecessor (Fischer et al. (1989)). It has 23 parameters and 9 (non-linear) equations, many with unreduceable integrals and embedded optimizations. The economic intuition of the model is easy to follow. However, its algebraic intuition is not: The model has no closed-form solutions or comparative statics. Although the Strebulaev model conveys some intuition through numerical solutions of reduced-form variable, in terms of economic intuition, it is a step backwards from its predecessors. Yet, in exchange, its quantitative formulation also offers great advantages. It can use the data more efficiently. The model can be tested not only more precisely on specific magnitudes, but even in cases in which the comparative statics are not unambiguous. Qualitative models often simply ignore ambiguous comparative statics in empirical tests.²¹ Of course, relative to qualitative directional testing, quantitative

²¹On occasion, qualitative tests do multi-moment testing—e.g., predicting that one moment has a certain direction only if another moment has a certain direction.

tests must lean harder on the model and its empirical proxies in order to identify an exact location in the parameter space.

In sum, in this and in many other quantitative models, it is often the case that at the cost of loss in intuition for the reader, a quantitative approach can take the model more seriously, be more ambitious, and if successful, be more empirically useful than a qualitative model.²²

B The In-Sample Performance of the Tax-Distress-Friction Model Revisited

I will now show that the main problem of the Strebulaev (2007) theory is that it put all its eggs into one basket (the tax-distress-friction model)—*but that this is the wrong basket*. I will show that the tax-distress-friction model cannot possibly explain the empirical non-adjustment evidence in corporate leverage ratios that motivated it. My evidence speaks to my first point—that papers in this genre are not built on forces that are known to be the main first-order mechanisms determining capital structures. Although such a critique must be specific to one model being critiqued, it is not difficult to find similar problems in other papers, too.

The principal mechanism through which the tax-distress-friction model explains non-adjustment is inertia. Firms that experience shocks are reluctant to adjust their capital structures, because the costs of doing so exceed the benefits. In the theory, firms' inertia causes non-adjustment. Although inertia implies non-adjustment, non-adjustment does not imply inertia. If managers randomly increase or decrease their leverage ratios following shocks, they do not adjust—but they are also not inert. The empirical evidence that Strebulaev (2007) claims in support of the model are about non-adjustment, although the prediction of the theory and thus his interpretation is that there is inertia. Inertia is not a “nuisance moment”—on the contrary. It is the primary cause of *all* the other quantitative implications in the Strebulaev (2007) model. I will now provide evidence that firms are not inert. They are very active—the variance of managerial leverage changes is very high—and moreover their managerial capital structure activity does not push them towards readjusting.

²²Calibration can serve two purposes: it can help assess whether a model's effects are of the right order of magnitude. Or it can help assess whether the model *alone* can explain the empirical evidence. The latter is more objectionable than the former.

The tax-distress-friction inertia hypothesis predicts the following:

1. dca, i.e., all changes caused by managerial issuing and repurchasing activity, should be low most of the time. Strebulaev can quantify its magnitude: active capital structure changes (dca) should be 0% about 85% of the time, and large (10-50%) otherwise.

Figure 3 shows a histogram of dca. There is indeed a small spike at 0% leverage. (Because leverage is defined on the unit domain and 20% of all firms have zero leverage [many remaining there], dca can of course not be normally distributed.) However, the spike is nowhere near the 85% predicted by the Strebulaev model. There is lots of managerial activity at, say, -3% and +3%.

Table 2 provides replicated model and actual empirical statistics. The model is a variant of Goldstein et al. (2001) and Strebulaev (2007), and generously made available to me by Ilya Strebulaev. His original model used a 5% friction parameter, the second data column in the table. A number of moments in this Goldstein-version are different from the original Strebulaev paper and the current data, because this version of the model has stronger stationarity requirements. For me, the important moments are those on dca. More than 90% of firm-year observations should be exactly zero (no managerial adjustment), with the remaining firm-years seeing very large changes. (10% of non-zero firm-years are large enough to produce an unconditional standard deviation of about 10%!) In contrast, in the data, more than 90% of firm-years have dca's that are *not* zero. The distribution of changes seem quite smooth—nothing like the rare sudden changes prescribed by frictions. This may not be surprising—after all, credit lines and debt and equity repurchases may well make adjustments fairly cheap. However, this is precisely contrary to the friction assumption in the model.

At this point, the friction aspect of the specific model in Strebulaev (2007) can be viewed as inconsistent with the data. dca contains a lot of activity, which rejects the (tax-distress-) friction perspective. (Welch (2004, Table 5) further shows that managerial activity occurs in many different claims, too: about half is long-term debt related, about a third is short-term debt related, and about ten percent is equity related.) And if anything, the evidence provided here is conservative. Managers could also be very active within the categories themselves. For example, if managers exchanged one form of debt for another (Rauh and Sufi (2010)), or if managers repurchased and reissued equity, or if managers changed debt and equity proportionally, such activity would not even show up in dca as it was defined here.

Again, at this point, one can object that other factors, such as precommitted ESOPs or sinking funds could help explain these year-to-year changes. This is correct, but this is precisely *not* in the Strebulaev (2007) model. The omission of other forces leave most of the variation in managerial leverage structure activities unexplained.

Although the central friction aspect is already rejected, we can examine the readjustment aspect under the additional assumption that the optimal tax-distress-induced leverage ratio target is moving only slowly. This assumption is reasonable for most publicly-traded corporations. (Many theories and tests have even assumed this to be perfectly static from year-to-year.) Unfortunately, the data also strongly reject the readjustment aspect:

2. The variability of total capital structure changes (dct) should be lower than the variability of capital structure changes in the absence of managerial intervention (dcp). That is, managers should use their capital structure activity to dampen swings in leverage, not worsen them. In Table 2, the calibrated model estimates suggest around -30% as a reasonable correlation between dca and dcp .

Table 1 shows that $sd(dct) \not\leq sd(dcp)$. In fact, dca and dcp are approximately orthogonal. Thus, the variance of dct is about $\sqrt{2}$ times the variance of either.

When stock returns reduce the leverage ratio ($dcp \ll 0$), managers should work, on average, to increase it again ($dca > 0$). Conversely, when stock returns increase the leverage ratio ($dcp \gg 0$), managers should work to decrease it again ($dca < 0$). These are *on average* predictions, not *in all cases* predictions.

Table 2 and Figure 4 show that there is no systematic response of managers to stock price movements.

In sum, the empirical evidence strongly suggests not only that managers are very active, but also that they do not use their activity to dampen fluctuations in their capital structures. Instead, managers add to capital structure volatility. Importantly, for any future theory that is based on a moving target to explain the evidence, this target must not only show tremendous variability, but the target's movements must also be orthogonal to what stock returns do to firms' leverage structures.²³ This simply does not seem plausible for a wide range of existing theories.

²³A different way to state this is that average outcome leverage is quantitatively too eerily close to what stock returns do in the absence of managerial intervention. It is highly unlikely that these revised capital structures also happen to be on the numerically exact target. The evidence strongly suggests that managers are actively maximizing something that we have not yet identified.

In sum, even if managers are less capital-structure active than they would be in the absence of capital-structure transaction costs, inertia is simply not a first-order determinant of capital structure. Managers are active. And whatever their motives, managers show no interest in dampening the volatility of their leverage ratios.

C Assessment and Further Challenges

Like all tax-distress-friction models, the Strebulaev model is fundamentally incapable of explaining high levels of non-adjusting managerial activity. Nevertheless, Strebulaev (2007) declared victory based on two moments (a negative correlation between stock returns and leverage, and a negative correlation between productivity and leverage), even though the theory's causal link from inertia to non-adjustment is absent. Thus, the model fails even the weakest test, in-sample scrutiny.

Strebulaev (2007) also did not consider the other challenges enumerated in Section I. It focused on the three tax-distress-friction forces *a priori*, and thus ignored a host of other potentially important forces. The forces it chose are weak, at best. It is not clear that the other forces mentioned in Section I are orthogonal to those in the model, and how this would affect the inference. Strebulaev (2007) offered few controls or residual diagnostics. It did not consider the effects of selection and survivorship biases. Its reduced-form inference was based on 75 years of simulated quarterly data, a marked contrast to the average 10-year life span of publicly traded firms. It did not use mid-year stock returns to predict impending disappearance. It offered no estimation in differences—here, using the theory to predict exactly when large changes in capital structure are expected. (The model does offer such predictions.) It offered no quasi-experimental tests. Good quasi-experimental tests of the tax-distress-friction theory are readily available—all three ingredients (taxes, distress costs, frictions) experienced regulatory changes in recent decades that can be viewed as exogenous from the perspective of individual corporations. If anything, quantitative models are better suited than qualitative models to tests using such natural experiments. For example, specific changes in corporate income tax rates should have led to specific quantitative changes in leverage ratio. (Givoly et al. (1992) investigated the 1986 Tax Reform Act.) Can the Strebulaev model predict the quantitative response better than a more naïve capital structure version?

IV Contrasts in Perspectives

The following three simple empirical regularities about managerial capital structure activities have been defying explanation:

1. Managers are active in issuing and retiring debt and equity.
2. This managerial (net) capital structure activity is orthogonal to the (non-linear) influence of stock returns on leverage ratios.

Combined with the facts that stock returns have significant volatility and that there is variation in lagged leverage ratios, the first two regularities imply that stock returns can explain about 45% of the variation in year-to-year market-leverage changes. (Long-term debt net issuing can explain about 30%, other debt changes can explain about 20%, and equity net issuing can explain about 5%).

3. Although there are many statistically significant variables, some with large economic coefficients, no known variables can explain much of the variation (R^2) in managerial capital structure activities.

In my opinion, managerial capital-structure behavior largely remains a mystery.

In contrast, the two models described above conclude with much more optimistic perspectives. Hennessy and Whited (2005) write

Our theoretical and empirical results underline the importance of understanding corporate financial decisions in dynamic settings, as well as the importance of having a tight connection between theory and empirical work. Given the power of our theoretical and empirical framework to explain observed leverage phenomena, it appears likely that similar success is possible in other areas of corporate finance.

Strebulaev (2007) writes that

Research that combines these two strands [real cash flow models and capital structure models] is likely to be a fruitful avenue for future research in capital structure, and more generally, corporate finance.

Similar confident sentiment can be found in many other structural corporate finance papers. For example, Li, Livdan and Zhang (2009), the lead article in the November 2009 issue of the *Review of Financial Studies*, begins both its abstract and introduction with the statement

We take a simple q-model and ask how well it can explain external financing anomalies both qualitatively and quantitatively. Our central insight is that optimal investment is an important driving force of these anomalies.²⁴

My assessment is more pessimistic. In practice, deep-structure theories in corporate finance, with their wide gaps between structure and reduced form, have only passed empirical tests that would be analogous to judging qualitative reduced-form theories by the t-statistic of a selected variable in an *in-sample* regression, without controls for competitive explanations and confounding variables, and without diagnostics and corrections for a whole range of possible misspecification errors. Moreover, some of these theories were designed to explain a particular correlation observed in earlier papers. In such cases, viewed from a global perspective, the research approach was analogous to searching for a good t-statistic first, followed by finding support for a specific model constructed to deliver this statistic. No models considered alternatives more powerful than the simplest strawman. In sum, all existing tests of deep structure quantitative theories that I am aware of have been perfunctory.

In my opinion, it is no surprise that there have been no quantitative and no deep-structure models that have performed well in explaining empirical corporate finance data. Given the challenges, it would have been a miracle if they had. For now, much managerial capital-structure activity (and many other corporate actions) remain largely a mystery. We still need to learn what the first-order effects are—the noise is deafening. The points in my paper suggest to me that instead of building quantitative and/or deep structure models now, greater rewards can be reaped by making progress on two more basic issues first:

²⁴LLZ reach this conclusion based on the fact that their model can be calibrated in-sample to some of the moments in the data, although it fails on others. The conclusion is a logical error. If theory A makes two predictions, B and C, it is wrong to credit it as a success when we observe “B” and “not C.” The correct conclusion is “not A.” (The Ptolemaic geocentric theory could explain some but not all moments, too.) In a deep-structure theory, such a skeptical approach is all the more important. Note that this logical inference issue is different from the fact that if a model suffers noise, it may fit all the empirical moments only moderately well. Moreover, LLZ conduct no out-of-sample tests and do not consider the hypotheses put forth in earlier literature (i.e., non-rational behavior) as the null hypothesis that their own model needs to reject with 95% confidence. Instead the burden of proof is pushed onto alternatives to their model, which have not been considered. Their model has in effect usurped the null hypothesis.

1. Empirically, we need to learn what the first-order associations are. This is mostly an exploratory and descriptive exercise, though it should be theoretically informed. The initial goal can be to learn which variables have a (possibly linearized) strong association with the dependent variable—a question of economic magnitude (and statistical standard error) of the coefficient. The more ambitious goal is to learn which variables can explain much of the variation (a question about R^2) in the dependent variable.
2. Empirically, we need to learn whether there is good evidence that changes in such variables associate with changes in the dependent variables (usually a managerial action). This applies both to variables that seem economically significant in levels, and to variables that seem to explain much of the variation (R^2) in changes.
3. Empirically, we need to determine which of the first-order variables exert some influence in simple quasi-experimental tests.

In the capital structure arena, the appropriate dependent variable is the managerially induced change in capital structure, not the stock-market (or accounting) induced total change in capital structure. Surveys of the actual decision-makers, such as those in Graham and Harvey (2001), can help inform this research.²⁵ Once we have a better idea, it will be more productive than it is today to build more elaborate theories, including quantitative and deep-structure theories; and to put them to the test, taking advantage of the many quasi-experimental tests that the corporate finance context offers.

²⁵Unfortunately, they neither correlate survey answers the actual capital structure actions of these firms, nor do they debrief their CFOs when their activities obviously deviate from their stated motives.

V Beyond Corporate Finance

Quantitative and structural modeling has also been popular outside of corporate finance. As mentioned in the introduction, Angrist and Pischke (2010) and Caballero (2010) point out related issues in industrial organization and macroeconomics (though with different perspectives and conclusions). It is a natural question to ask whether asset pricing is better suited to quantitative and deep-structure modeling than corporate finance. There are some differences.

First, the data frequency in asset-pricing is often much higher. This can sometimes make it easier to uncover the correct mechanism.

Second, and most importantly, there are much stronger market forces. When agents make mistakes, near arbitrage can push prices back toward their correct values. Arbitrage can in effect convert deep mechanisms into shallow ones. If the mechanism is known, the first point of my critique is no longer applicable. Not surprisingly, structural models in asset pricing have generally performed better where arbitrage bounds are fairly strong, such as in derivatives pricing (or possibly fixed-income pricing with a strictly timed final convergence condition) relative to the underlying equity pricing. (In the extreme, this is the domain that Summers (1985) termed “ketchup economics.”)

Yet arbitrage bounds are so weak in equities that it does not seem surprising that quantitative structural models (usually based on risk and market imperfections) have also had a poor track-record explaining even first-order empirical regularities, such as the value and momentum effects. Without strong arbitrage-bounds in equity pricing models, further structural “enhancements” to the model run the risk of overfitting the data—creating more flexible reduced forms that can be attributed to many models beyond the original structural model itself.

Third, testing approaches have been different. It has not been uncommon in asset pricing to conduct out-of-sample testing. However, there have been few or no attempts at conducting good and more stringent quasi-experimental tests. On the one hand, this may be because, in contrast to corporate finance, there are few natural experiments in asset pricing that come to mind. For example, it is difficult to think of a clearly identifiable exogenous shock to factor exposures. On the other hand, it seems ironic that micro foundations are as popular in structural theory modeling, as they are unpopular in the empirical tests of these models. For example, habit formation has been a popular modeling primitive, but using individual trading data to measure these habit preferences within the same data set used

to test the expected return predictions of habit formation has not been equally popular. Of course, the proper conclusion from the lack of a tight link between unobservables and observables (the large gap) and from the lack of quasi-experimental tests is not that we can adopt such models even without such evidence, but that we should view empirical evidence in their favor as very weak. We need to remain even more skeptical of them.

VI Conclusion

It is tempting to infer from the increasing number of award-winning quantitative and deep-structure corporate finance papers that we have largely settled on theories and modern modeling approaches, that there is good supporting empirical evidence, and that we only need to develop these models further if we want to understand corporate behavior better. Eventually, with more bells and whistles (degrees of freedom), the models should fit.

My paper has disputed this. It has argued that no quantitative and (deep) structural model has been subjected to as much empirical scrutiny as the best alternatives. None has succeeded in an empirical test that was more than perfunctory. In the future, quantitative and deep-structure models deserve at least as much scrutiny as simpler models. The modern standard for testing economic theories is quasi-experimental, and quasi-experimental tests are often just as feasible in structural and quantitative contexts as they are in alternative approaches. Moreover, direct empirical proxies for underlying constructs would often make tests more credible than correlations among very different types of variables, tied together only by assumed behavioral first-order conditions. *De facto*, theories which do not allow for either will remain forever in the realm of conjecture, remaining empirically meaningless.

To be clear, my paper has not argued in favor of theory-free empiricism. Models help inform empiricists where to look and where interpretation pitfalls loom. However, my paper has argued that, at the current stage of knowledge in corporate finance, comparative statics from simple “marginal” sketch models with quasi-experimental tests seem better suited to learning what the most important forces are, compared to the more ambitious “comprehensive models.” My paper has also not argued against the presence of first-order conditions in models. However, it has argued against relying on their inverted forms to infer (identify) causal effects without more direct empirical proxies—i.e., when it is likely that these first-order conditions leave a wide gap between an unobserved variable and the reduced-form variable. Finally, my paper has not argued that quantitative and deep-structure models are intrinsically inferior. On the contrary. If such models can explain the

data better than their simpler counterparts (and this paper has made specific suggestions on how to improve on this aspect), then these models can be more useful, despite the additional cost to the reader to understand their economic intuition. However, my paper has argued that quantitative and deep-structure models are more likely to succeed in simpler settings, in which the first-order mechanisms are well understood. In financial economics, this is really only the case in situations in which there is strong arbitrage. Therefore, I have little hope that structural and quantitative approaches can succeed in corporate finance.

Reasonable people can disagree with my pessimistic assessment. I would be thrilled to be proven wrong.

A Figures and Tables

Table 1: Descriptive Statistics of Capital Structure Changes, in %

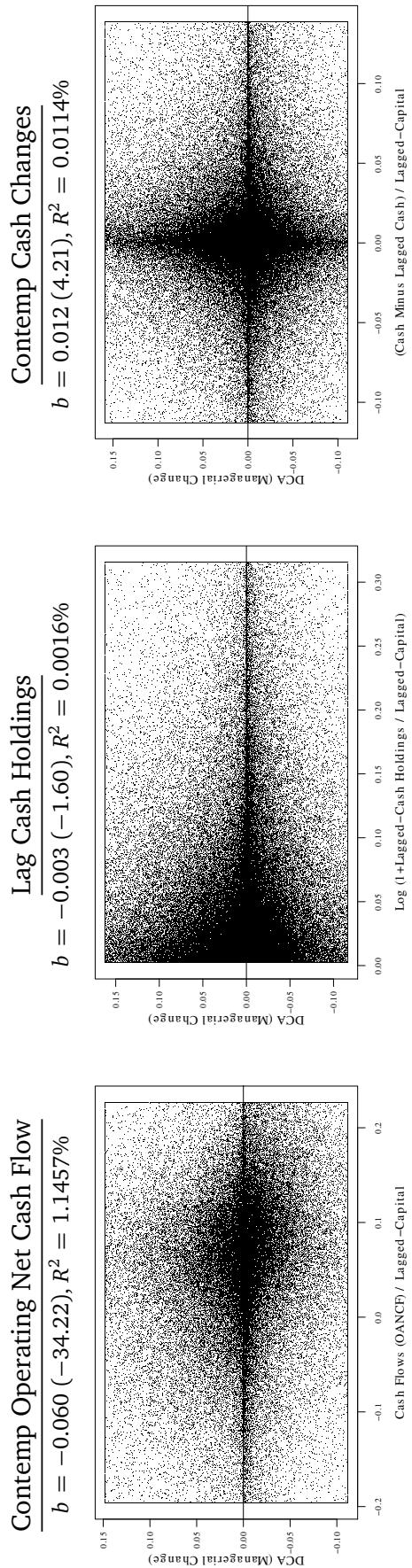
	Unwinsorized						Winsorized at 5% and 95%			
	Min	Median	Max	Mean	Sdv	Range	Mean	Sdv	N	
dca (excl. divs)	-99.6	0	98.5	0.9	10.8	-11.5...16.3	0.9	6.3	179,917	
Contemp dct	-99.6	0	98.5	1.2	13.7	-18.6...24.1	1.2	10.2	179,954	
Contemp dcp	-66.1	0	90.4	0.3	8.8	-12.4...15.5	0.1	6.5	179,954	
Contemp log(1+r)	-601.8	2.7	519.1	-2.9	59.7	-106.1...81.8	-2.0	47.8	179,954	
Lagged dca	-99.6	0.0	98.5	1.0	10.5	-11.0...16.2	1.0	6.2	162,480	
Lagged dc	0.0	21.8	100.0	27.8	25.3	0.0...76.3	27.4	24.3	179,954	
Lagged cash	0.0	3.9	274.9	8.2	12.7	0.2...31.6	7.3	8.5	158,928	
Contemp cash changes	-965.8	0.1	219,229.4	4.9	637.6	-11.3...13.9	0.5	5.4	154,941	
Contemp operating CF	-4,248.8	5.7	30,195.0	5.5	131.2	-19.6...22.7	4.4	10.3	101,032	

Description: The capital structure change variables dca, dct, and dcp are defined in (1) on page 9, and quoted in percent. In brief, dct are total debt-to-capital ratio changes, dcp are the part due to capital gains, and dca are the part not due to capital gains, i.e., due to managerial actions. The dependent variable is usually dca.

dc is the market-value-based debt-to-capital ratio. Cash is Compustat CH. Operating Cash flows are Compustat OANCE Cash, cash changes, and operating cash flows are normalized by lagged financial capital (debt plus market value of equity). Lagged cash is also quoted in “one-plus logs.” “Contemp” means contemporaneous with dca. “Lagged” means at the beginning of the period for stock variables dc and cash, and one year-lagged for flow variable dca.

The sample are all firm-years from Compustat from 1963 to 2007, excluding firms with lagged financial capital of less than \$1,000,000. All reported R^2 in subsequent tables are from regressions on the winsorized variables, and all figures display the winsorized observations at their borders.

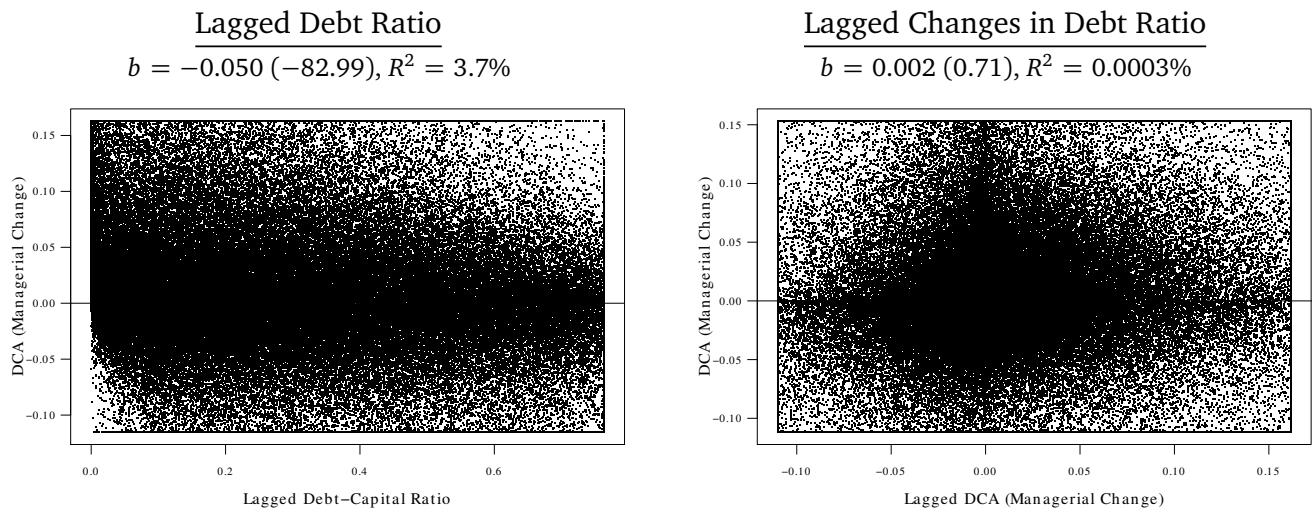
Figure 1: Managerial Changes of Debt-To-Capital Ratios (dca, in %) Vs Liquidity



Description: Sample and variables are described in Table 1. In brief, the independent variables are measures of liquidity. The dependent variable (dca) is the active managerial capital structure change. T-statistics are in parentheses following the coefficient.

Interpretation: The plots suggest that knowledge of how “cash rich” firms are does not explain much of managerial capital structure actions. Even the best of the three variables, *contemporaneous operating net cash flow*, leaves 99% of managerial capital-structure activity unexplained.

Figure 2: Hysteresis and Managerial Changes of Debt-To-Capital Ratios (dca, in %)



Description: Sample and variables are described in Table 1. In brief, the independent variables are lagged measures of capital structure (left) and managerial capital structure changes (right). The dependent variable (dca) is the active managerial capital structure change. T-statistics are in parentheses following the coefficient.

Interpretation: The left plot suggest that knowledge of firm's lagged capital ratios has marginal explanatory power for managerial capital-structure activity, but it leaves 96% unexplained. Note that this is optimistic, because firms that failed to reduce their leverage so may have gone to a 100% debt-ratio (bankruptcy) and dropped out. The right plot shows that there is no hysteresis in managerial capital-structure activities—to predict managerial leverage-related actions, past managerial leverage-related actions are not very informative.

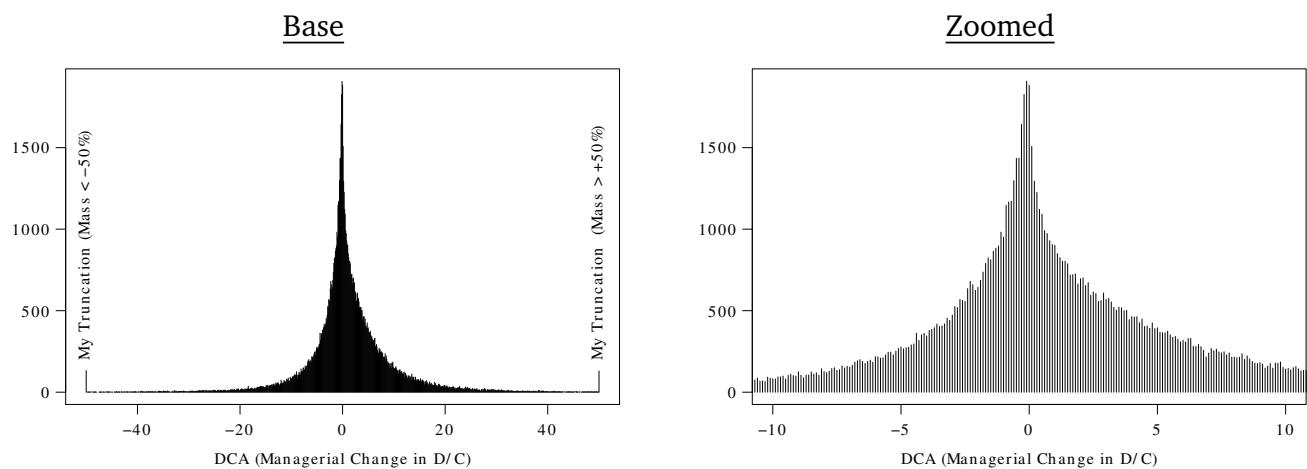
Table 2: Model Moments and Data Moments

	1%	Tax-Distress-Friction Model γ				Current Data
		5%	10%	30%	∞	
Current Debt QMLC						
Mean	59.9%	55.6%	50.5%	38.1%		28.5%
Sd	5.9%	9.9%	12.4%	15.3%		26.4%
Frac=0	None	None	None	None		0.120
Welch Regression Predicting QMLC (with unreported constant)						
QMLL	-0.36	-0.38	-0.45	-0.50	0	-0.042
IDR	0.56	0.93	1.147	1.343	1	0.945
dca \equiv QMLC - IDR						
Mean	0.034	0.030	0.026	0.020	0.00	0.010
Sd	0.074	0.095	0.101	0.099	0.00	0.108
Frac = 0	83%	91%	94%	96%	100%	8%
					Frac \leq 1%	31%
					Frac \leq 2%	45%
					Frac \leq 3%	55%
Correlations						
corr(dca,dcp)	-40.5%	-30.9%	-27.7%	-26.3%		≥ 0
corr(dca,dcp ₋₁)	-39.3%	-25.4%	-18.6%	-10.0%		≥ 0

Description: The simulations are based on a model similar to that in Strebulaev (2007) (based on Goldstein et al. (2001)), and kindly made available to me by Ilya Strebulaev. The friction parameter γ closest to the original model is 5%. Simulations are based on 3,000 firms and 140 quarters. The Welch regression is run as a simple pooled regression, which delivers practically the same results as the Fama-Macbeth regressions in the original paper.

Interpretation: dca adjustments are too large (and too smooth) to warrant the friction-induced discrete characterization of capital structure changes. Moreover, the correlations between dca and dcp suggest that managers do not change their capital structures primarily to stabilize their capital structures.

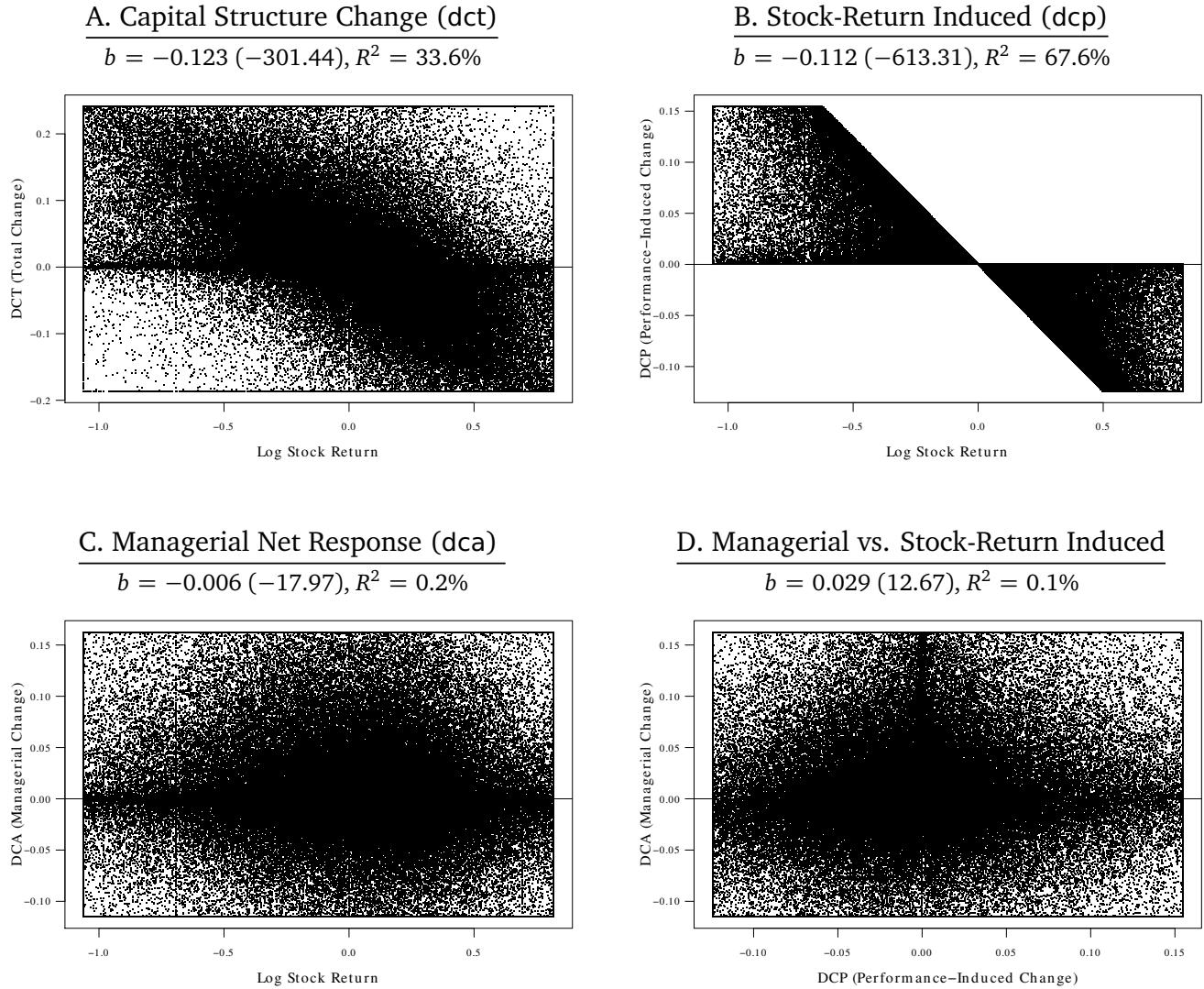
Figure 3: Histogram of Managerial Changes to Debt-To-Capital Ratios (dca), in %



Description: Sample and variables are described in Table 1. dca are active managerial changes to capital structure.

Interpretation: Inertia (lack of activity) is not a first-order characteristic of managerial behavior. There is plenty of activity off the zero-center.

Figure 4: Changes of Debt-To-Capital Ratios, in %, and Stock Returns



Description: Sample and variables are described in Table 1. In brief, the independent variable in Panels A through C are contemporaneous stock returns. In Panel D, it is the change in capital structure caused by these stock returns. The dependent variable (dca) are active managerial capital structure changes. (Horizontal bands at zero in Panels A, B, and C occur because about 10% of all firms had zero debt, so their leverage ratios cannot be affected by stock returns.) T-statistics are in parentheses.

Interpretation: Panel A illustrates that stock returns explain at least one-third of the year-to-year variation in capital structure. Stock returns are a first-order determinant of capital structure changes. Panel B illustrates stock returns do not explain 100% of stock-return induced leverage variation, because firms' own lagged capital structures modulate the effect of the same stock return differently for different firms. Panels C and D illustrate that, although there is significant managerial capital structure activity (wide spreads in dca), there is no clear directional managerial counteracting response to stock returns.

A Defining Structure

From the perspective of my critique, it is not essential how structure is defined. My arguments are about the lack of a strong link between unobservable constructs and the tested measured constructs. To clarify some ambiguity in different authors' interpretations, this appendix briefly describes three definitions of "structure."

One common definition of a structural model is that it is one whose estimated equations are the result of a first-order condition. The standard example of a reduced-form model to contrast with a structural model is Fama and French (1992). This is not a useful definition for my paper. It is not impossible that some as-yet-unknown model exists whose first-order conditions imply exactly the Fama-French specification. Thus, the model could be reduced-form in the minds of some researchers, but structural in the minds of others. The same published model with the same specifications and data could be reduced-form today and structural tomorrow. There is not even a guarantee that *any* reduced-form models exist in this definition.

Another definition of structure, due to Lucas himself, is a self-fulfilling rational-expectations model. Here the Fama-French model is not structural if it is the case that the relation would change if investors used the model itself. In the strict Lucas sense, a model is structural if only preferences and technology are specified. And in this very strict sense, there are *no* structural models in corporate finance.

Yet, another definition of structure might emphasize the functional form of the relationship. However, functional identification leans so strongly on the model functional input assumptions that it has fallen out of favor.

References

- Almeida, Heitor, Murillo Campello, Bruno A. Laranjeira, and Scott J. Weisbenner**, “Corporate Debt Maturity and the Real Effects of the 2007 Credit Crisis,” Technical Report, University of Illinois at Urbana-Champaign 2009.
- Angrist, Joshua D. and Jörn-Steffen Pischke**, *Mostly Harmless Econometrics*, Princeton University Press, 2008.
- _____, **and Jörn-Steffen Pischke**, “The Credibility Revolution in Empirical Economics: How Better Research Design is Taking the Con Out of Econometrics,” *Journal of Economic Perspectives*, Spring 2010, 42 (2), 3–30.
- Baker, Malcolm and Jeffrey Wurgler**, “Market Timing and Capital Structure,” *The Journal of Finance*, February 2002, 57, 1–32.
- Bertrand, Marianne and Antoinette Schoar**, “Managing With Style: The Effect of Managers on Firm Policies,” *Quarterly Journal of Economics*, November 2003, 118 (4), 1169–1208.
- Bolton, Patrick, Hui Chen, and Neng Wang**, “A Unified Theory of Tobin’s q, Corporate Investment, Financing, and Risk Management,” Technical Report, Columbia University and Massachusetts Institute of Technology and Columbia University August 2010.
- Caballero, Ricardo J.**, “Macroeconomics after the Crisis: Time to Deal with the Pretense-of-Knowledge Syndrome,” Technical Report, MIT September 2010.
- Chang, Xin and Sudipto Dasgupta**, “Target Behavior and Financing: How Conclusive is the Evidence,” *The Journal of Finance*, 2009, 64 (4), 1767–1796.
- DeAngelo, Harry, Linda DeAngelo, and Rene M. Stulz**, “Seasoned Equity Offerings, Market Timing, and the Corporate Lifecycle,” *Journal of Financial Economics*, March 2010, 95, 275–295.
- _____, _____, **and Toni M. Whited**, “Capital Structure Dynamics and Transitory Debt,” *Journal of Financial Economics*, 2010, ?, forthcoming.
- Dittmar, Amy K. and Anjan V. Thakor**, “Why do firms issue equity?,” *The Journal of Finance*, February 2007, 62 (1).
- Einav, Liran and Jonathan Levin**, “Industrial Organization: A Progress Report,” Technical Report, Stanford University and NBER December 2009.
- Fama, Eugene F. and Kenneth R. French**, “The Cross-Section of Expected Stock Returns,” *The Journal of Finance*, June 1992, 68 (2), 427–465.
- _____, **and _____**, “Testing Trade-Off and Pecking Order Predictions about Dividends and Debt,” *Review of Financial Studies*, 2002, 15, 1–33.
- _____, **Lawrence Fisher, Michael C. Jensen, and Richard Roll**, “The Adjustment of Stock Prices to New Information,” *International Economic Review*, February 1969, 10 (1), 1–21.
- Fischer, Edwin O., Robert Heinkel, and Josef Zechner**, “Optimal dynamic capital structure choice: Theory and tests,” *The Journal of Finance*, 1989, 44, 19–40.

- Friedman, Milton**, “The Methodology of Positive Economics,” in “Essays in Positive Economics” 1966, pp. 3–16.
- Givoly, Dan, Carla Hayn, Aharon R Ofer, and Oded Sarig**, “Taxes and Capital Structure: Evidence from Firms’ Response to the Tax Reform Act of 1986,” *Review of Financial Studies*, 1992, 5 (2), 331–55.
- Goldstein, Robert, Nengjiu Ju, and Hayne E. Leland**, “An EBIT-based model of dynamic capital structure,” *Journal of Business*, 2001, 74, 483–512.
- Graham, John and Campbell Harvey**, “The Theory and Practice of Corporate Finance: Evidence from the Field,” *Journal of Financial Economics*, May 2001, 61 (2-3), 187–243.
- Graham, John R.**, “Debt and the Marginal Tax Rate,” *Journal of Financial Economics*, 1996, 41, 41–73.
- _____, “How Big Are the Tax Benefits of Debt?,” *The Journal of Finance*, October 2000, 55 (5), 1901–1941.
- Greenwood, Robin M., Samuel Hanson, and Jeremy C. Stein**, “A Gap-Filling Theory of Corporate Debt Maturity Choice,” *The Journal of Finance*, 2010, 70, ?
- Hennessy, Christopher A. and Toni M. Whited**, “Debt dynamics,” *The Journal of Finance*, June 2005, 60 (3), 1129–1165.
- _____, and _____, “How Costly is External Financing? Evidence from a Structural Estimation,” *The Journal of Finance*, August 2007, 62 (4), 1705–1745.
- Huang, Rongbing and Jay R. Ritter**, “Testing Theories of Capital Structure and Estimating the Speed of Adjustment,” *Journal of Financial and Quantitative Analysis*, May 2009, 44 (2), 237–271.
- Iliev, Peter**, “The Effect of SOX Section 404: Costs, Earnings Quality, and Stock Prices,” *The Journal of Finance*, June 2010, 65, 1163–1196.
- _____, and Ivo Welch, “Reconciling Estimates of the Speed of Adjustment of Leverage Ratios,” Technical Report January 2010.
- Ju, Nengjiu, Robert Parrino, Allen M. Poshman, and Michael S. Weisbach**, “Horses and Rabbits? Optimal Dynamic Capital Structure from Shareholder and Manager Perspectives,” *Journal of Financial and Quantitative Analysis*, 2005, 40, 259–281.
- Kisgen, Darren J.**, “Credit Ratings and Capital Structure,” *The Journal of Finance*, June 2006, 61 (3), 1035–1072.
- Leamer, Edward**, “Let’s Take the Con Out of Econometrics,” *American Economic Review*, 1983, 73 (1), 31–43.
- Leland, Hayne E.**, “Corporate Debt Value, Bond Covenants, and Optimal Capital Structure,” *The Journal of Finance*, September 1994, 45 (4), 1213–1252.
- _____, “Agency Costs, Risk Management, and Capital Structure,” *The Journal of Finance*, August 1998, 53 (4), 1213–1243.
- _____, and Klaus Toft, “Optimal capital structure, endogenous bankruptcy, and the term structure of credit spreads,” *The Journal of Finance*, 1996, 51, 987–1019.
- Lewellen, Jonathan and Katharina Lewellen**, “Internal Equity, Taxes, and Capital Structure,” Technical Report, Dartmouth College and NBER March 2006.

- Li, Erica X. N., Dmitry Livdan, and Lu Zhang**, “Anomalies,” *Review of Financial Studies*, 2009, 22 (11), 4301–4334.
- Lucas, Robert E.**, “Econometric Policy Evaluation: A Critique,” *Carnegie-Rochester Conference Series on Public Policy*, 1976, 1, 19–46.
- Malmendier, Ulrike and Geoffrey Tate**, “CEO Overconfidence and Corporate Investment,” *The Journal of Finance*, December 2005, 60 (6), 2661–2700.
- Mehra, Rajneesh and Edward C. Prescott**, “The Equity Premium: A Puzzle,” *Journal of Monetary Economics*, March 1985, 15, 145–161.
- Modigliani, Franco and Merton H. Miller**, “The Cost of Capital, Corporation Finance and the Theory of Investment,” *American Economic Review*, June 1958, 48 (3), 261–297.
- Myers, Stewart C. and Nicholas S. Majluf**, “Corporate financing and investment decisions when firms have information that investors do not have,” *Journal of Financial Economics*, June 1984, 13, 187–221.
- Parrino, Robert and Michael Steven Weisbach**, “Measuring Investment Distortions Arising from Stockholder-Bondholder Conflicts,” *Journal of Financial Economics*, July 1999, 53, 3–42.
- Rampini, Adriano A and S. Viswanathan**, “Collateral and Capital Structure,” ?, 2010, ?, ?–?
- Rauh, Joshua D. and Amir Sufi**, “Capital Structure and Debt Structure,” *Review of Financial Studies*, 2010, 23, forthcoming.
- Roberts, Michael R. and Amir Sufi**, “Control Rights and Capital Structure: An Empirical Investigation,” *The Journal of Finance*, 2009, 64, 1657–1695.
- _____, and Mark Leary, “Do Peer Firms Affect Corporate Financial Policy,” Technical Report, University of Pennsylvania and Cornell University 2009.
- Robichek, Alexander A. and Stewart C. Myers**, “Problems in the Theory of Optimal Capital Structure,” *Journal of Financial and Quantitative Analysis*, June 1966, 1 (2), 1–35.
- Roll, Richard**, “The Hubris Hypothesis of Corporate Takeovers,” *Journal of Business*, April 1986, 59 (2), 197–216.
- Shivdasani, Anil and Irina Stefanescu**, “How do Pensions Affect Corporate Capital Structure Decisions,” *Review of Financial Studies*, 2010, 23, forthcoming.
- Shyam-Sunder, Lakshmi and Stewart C. Myers**, “Testing Static Tradeoff Against Pecking Order Models of Capital Structure,” *Journal of Financial Economics*, 1999, 51, 219–244.
- Strebulaev, Ilya**, “Do Tests of Capital Structure Theory Mean What They Say?,” *The Journal of Finance*, 2007, 62 (1747-1787).
- Summers, Lawrence H.**, “On Economics and Finance,” *The Journal of Finance*, July 1985, 60 (3), 633–635.
- Titman, Sheridan and Roberto Wessels**, “The Determinants of Capital Structure Choice,” *The Journal of Finance*, 1988, 43 (1), 1–19.
- _____, and Sergey Tsyplakov, “A dynamic model of optimal capital structure,” *Review of Finance*, 2007, 11, 401–451.

- Welch, Ivo**, “Capital Structure and Stock Returns,” *Journal of Political Economy*, 2004, 112, 106–131.
- _____, *Corporate Finance : An Introduction*, Prentice-Hall, 2008.
- _____, “Two Common Problems in Capital Structure Research: The Financial-Debt-To-Asset Ratio and Issuing Activity Versus Leverage Changes,” *International Review of Finance*, 2011, ?, 1–17.