We quantify the importance of contracting frictions versus taxes in shaping firms’ capital structures. We estimate a dynamic contracting model in which a firm seeks debt financing from an intermediary and is subject to taxation. Because the firm can renege on the financing contract, the optimal contract is self-enforcing, so collateral constraints arise endogenously. Using data from firms in several industries, we find that a model without taxes fits the data as well a model with taxes. Quantitatively, taxes have little effect on optimal leverage but hurt investment. The model is able to detect changes in collateral underlying the changes in leverage surrounding a natural experiment concerning asset reposssession.
Collateral, Taxes, and Leverage

Shaojin Li  Toni M. Whited  Yufeng Wu*

March 25, 2014

*Li is from the Shanghai University of Finance and Economics, School of Finance, Shanghai, 200433 P.R. China. 86-21-65908391. li.shaojin@shufe.edu.cn. Whited is from the Simon Business School, University of Rochester, Rochester, NY 14627. (585)275-3916. toni.whited@simon.rochester.edu. Wu is from the Simon Business School, University of Rochester, Rochester, NY 14627. yufeng.wu@simon.rochester.edu. We thank Adriano Rampini, Shane Heitzman, Erwan Morellec, and seminar participants at Cass Business School, CEMFI, Duke, Northeastern, Princeton, Warwick, Wirtschaftsuniversität Wien, and York University for helpful comments and suggestions.
Collateral, Taxes, and Leverage

Abstract

We quantify the importance of contracting frictions versus taxes for firms’ capital structures. We estimate a dynamic contracting model in which a firm seeks debt financing from an intermediary and is subject to taxation. Because the firm can renego on the financing contract, the optimal contract is self-enforcing, so collateral constraints arise endogenously. Using data from firms in several industries, we find that a model without taxes fits the data as well a model with taxes. Quantitatively, taxes have only a limited effect on leverage. The model can detect changes in the value of collateral underlying a natural experiment concerning asset repossesssion.
1 Introduction

How important are taxes for the determination of corporate capital structures? The traditional discussion of this issue is based on the framework of Modigliani and Miller (1958) in which capital structure is irrelevant for firm value in the absence of taxes. Against this background, taxes and the countervailing friction of the costs of financial distress then provide the tradeoff that determines an optimal amount of debt in a firm’s capital structure. Yet this framework abstracts from two critical ingredients that influence corporate borrowing decisions: the time-varying need for funds to implement new projects and lenders’ desire to be repaid. In addition, many models that embody the traditional tradeoff between taxes and bankruptcy costs take external financial frictions as exogenous, so that firms’ actions and characteristics never affect the terms of their financing.

We depart from this paradigm by quantifying the importance of contracting frictions in shaping corporate financial decisions. In particular, we compare the importance of endogenous financial constraints that arise in a contracting environment with a more traditional friction that is central to most models of corporate capital structure: taxation. To this end, we formulate and estimate a dynamic model of capital structure that is based on limited enforceability of contracts between lenders and firms and that also incorporates taxation of income and deductibility of interest.

The results from our estimation are striking. Models both with and without taxation can be reconciled with average firm leverage, with little change in other model parameters across the two cases. Similarly, our counterfactuals show that changing the corporate tax rate has only a limited effect on optimal firm leverage. In contrast, taxation naturally has a sharp negative effect on payout policy. We are able to reconcile model-implied leverage with actual leverage for samples stratified by industry. Finally, using a natural experiment that affects creditors’ ability to repossess collateral in bankruptcy, we find that the significant changes in leverage surrounding this experiment do indeed stem from movements in the position of the collateral constraint.
The intuition for these results lies in the structure of the model, in which a firm with a possibly infinite lifespan enters the market with a stock of capital that can generate taxable revenues. However, the firm has insufficient funds to scale the project to an optimal level and must obtain financing from an intermediary, which is more patient than the firm. There are no informational asymmetries: the intermediary can observe firm policies. However, financing is not frictionless because the firm can renege on a financing contract, abscond with the firm’s capital, and start over, albeit after losing part of the firm’s capital as a deadweight cost. An additional friction is the firm’s limited liability, which prevents costless equity infusions. The contract maximizes the firm’s equity value, but because the lender must commit to the contract, the only feasible contracts are self-enforcing, so that the firm never has an incentive to renege.

The optimal contract specifies state-contingent financing, payout, and investment policies so that the firm’s long-term benefits from adhering to the contract outweigh the benefits from repudiating it. We show, as in Rampini and Viswanathan (2013) that the constraint that ensures contract enforcement is equivalent to a collateral constraint. The contract is sufficiently flexible that the firm can save in some states of the world, instead of carrying a stock of debt. Thus, borrowing constraints, capital structures, dividends, and investment policy emerge endogenously as a product of an optimal contract.

In this setting, taxes have little effect on leverage. First, the collateral constraint often binds for young, growing firms because the lender is more patient than the firm. Thus, the tax deductibility of interest affects the Lagrange multiplier on the collateral constraint and the shadow value of debt instead of the optimal quantity of debt in the firm’s capital structure. Second, mature firms often reach an interior solution that balances the values of the interest deduction and of free debt capacity. Although the tax rate is a determinant of the position of this interior solution, it is not the only determinant. Even when the collateral constraint does not bind, collateral is still a first-order determinant of leverage because the capacity preservation motive is shaped by the position of the collateral constraint. Traditional bankruptcy costs affect this tradeoff only indirectly by helping to position the
collateral constraint.

We stress that these results are quantitative in nature and come from a model whose parameters we do not choose arbitrarily. Instead, we estimate them via simulated method of moments. The data therefore put tight restrictions on our model parameters and on the predictions that emerge from the model. This feature of our approach is important because the relative magnitudes of the costs and benefits of leverage have been the center of much of the research agenda in capital structure, starting with the horse and rabbit stew of Miller (1977). Our findings extend this analogy in that we find that both the horse (taxes) and rabbit (bankruptcy costs) are of minor importance relative to collateral, which can be likened in size to the proverbial elephant in the room.¹

We find largely reasonable parameter estimates. Our estimates of the firm’s technological characteristics, such as the capital depreciation rate and the extent of decreasing returns to scale are in line with many other structural estimation studies (e.g. Hennessy and Whited 2005, 2007). More importantly, we estimate a parameter that describes the firm’s incentive to renege on the contract: the fraction of the firm’s assets that are lost to creditors in default. We find that this parameter is statistically different from zero and economically important. Finally, we find that the model, although highly stylized, can match important features of the data on firms in several diverse industries. We conclude that an optimal contracting model can characterize broad features of the data, even though the form of real-world contracts deviates from the exact model predictions.

Our findings would have been hard to obtain by more conventional methods. Capital structure and investment are endogenous, and most tax changes are motivated by political economy considerations. Even when they are plausibly exogenous, they affect so many firm decisions that it is hard to pinpoint specific capital structure effects that are motivated directly by interest deductibility. In addition, the main sources of the contracting frictions are unobservable, and proxies for these frictions are unavailable. Using a model helps solve these problems by putting enough structure on the data to identify the effects of interest.

¹We thank Adriano Rampini for this analogy.
Our paper fits into several literatures. The first is a set of theoretical papers that uses limited commitment models such as ours to study such subjects as international trade contracts (Thomas and Worrall 1994), financial constraints (Albuquerque and Hopenhayn 2004), macroeconomic dynamics (Cooley, Marimon, and Quadrini 2004; Jermann and Quadrini 2007, 2012), investment (Lorenzoni and Walentin 2007; Schmid 2011), risk management (Rampini and Viswanathan 2010), and capital structure (Rampini and Viswanathan 2013). The model in our paper is most closely related to that in (Rampini and Viswanathan 2013), but our paper is unique in this group because it uses a limited commitment model as the basis of an explicitly empirical investigation, whereas the rest of these papers are theoretical.

The second literature is the structural estimation of dynamic models in corporate finance, such as Hennessy and Whited (2005, 2007) or Taylor (2010). Our paper departs from these predecessors in one important dimension. Instead of specifying financial constraints or agency concerns as exogenous parameters, we derive financial constraints from an optimal contracting framework, and then estimate the magnitudes of the underlying frictions.

Our paper is also related to dynamic contracting models based on moral hazard (e.g. DeMarzo and Sannikov 2006; Biais, Mariotti, Rochet, and Villeneuve 2010; DeMarzo, Fishman, He, and Wang 2012). In these models, leverage serves solely as a device to incentivize the entrepreneur from consuming private benefits. In contrast, in the limited enforcement model we consider, the level of leverage is set so that the lender can guarantee repayment. In this group of papers, Nikolov and Schmid (2012) also conducts a structural estimation. However, their goal is to quantify private benefits, whereas the goal of our paper is to quantify the relative effects of taxation and contract enforcement on capital structure.

The rest of the paper is organized as follows. Section 2 develops the model. Section 3 describes the data. Section 4 explains the estimation methodology and identification strategy. Section 5 presents the estimation results. Section 6 describes several counterfactual experiments, and Section 7 concludes. The Appendices contain proofs, describe our model solution procedure, and outline the details of the estimation.
2 The Model

This section develops the model, which is a simple discrete-time, infinite-horizon, limited-enforcement contracting problem in the spirit of Albuquerque and Hopenhayn (2004), Rampini and Viswanathan (2010, 2013), or Lorenzoni and Walentin (2007). We first present a model without corporate taxation. Here, we start with a description of the firm’s technology. We then move on to describing the incentive and contracting environment. Next, we characterize the optimal contract. Finally, we explain how we add taxes to the model.

2.1 Technology

We consider an industry that consists of a continuum of firms, each of which produces a homogeneous product. Firms can enter and exit the industry. At time \( t \), an entrant with capital stock \( k_t \) can start to use this capital. Incumbents become unproductive and exit the industry with probability \( \phi \). We assume the mass of firms is fixed and normalize it to be one. Therefore, the mass of entrants each period is also \( \phi \).

Both incumbents and entrants use the production technology \( y_t = z_t k_t^\alpha \), in which \( k_t \) is a capital input, and \( z_t \) is a firm-specific technology shock, which follows a Markov process with finite support \( Z \) and transition matrix \( \Pi \). The entrant also takes its initial productivity draw from the same shock process. The law of motion for \( k_t \) is given by

\[
k_{t+1} = (1 - \delta)k_t + i_t,
\]

in which \( i_t \) is capital investment at time \( t \) and \( \delta \) is the capital depreciation rate.

2.2 Contracting Environment

Although entry is costless, upon entry, the firm has no current profits to fund expansion, and it therefore obtains financing by entering a contractual relationship with a financial intermediary/bank/lender. Three important assumptions shape the contract. First, the entrepreneur has limited liability. Second, the lender commits to the long-run contract, while the firms can choose to default; that is, the long-run contract has one-sided commitment.
Third, firms have a higher discount rate than do banks, as in Lorenzoni and Walentin (2007). Let $\beta$ be the discount factor for the firms, and let $\beta_C$ be the discount factor for the banks, with $\beta_C > \beta$ so that firms are less patient than lenders. As will be seen below, when $\beta = \beta_C$, the firm can sometimes be completely unconstrained, so that capital structure is irrelevant. In order to estimate this model, we require a determinate capital structure, so we must assume that $\beta_C > \beta$.

A plausible friction that might force a wedge between borrowers’ and lenders’ discount factors is the existence of insured deposits, which provide banks with a cheap source of capital and thus induces them to behave patiently. In addition, a natural way to motivate the difference in the discount rates of the principal and agent can be found in Ai, Kiku, and Li (2013). In their general equilibrium economy, both the principal and agent have the same rate of time preference, but the economy grows. In this setting, the discount rate of the agent exceeds that of the principal by expected consumption growth divided by the intertemporal elasticity of substitution. So naturally, the principal is more patient than the agent in growth economies.

The timing of events is as follows. When a new firm is born at time $t$, it receives a draw from the invariant distribution of the productivity shock and an initial capital stock $k_t$. The firm then signs a long-term contract with the lender that provides initial funding. Once the firm enters the contract, production takes place and the firm invests, pays out dividends and makes payments to the lender as required in the contract. At the beginning of next period, the firm first faces an exogenous exit shock. If it survives, the firm can choose whether to renege on the contract after observing the productivity shock and earning current-period profit. If the firm does not renege, the plan defined by the contract continues to be implemented.

We now define the specifics of the contract. Let $z_t$ be the state at time $t$, and let $z^t = (z_0, z_1, \ldots, z_t)$ denote the history of states from times 0 to $t$. A contract between the entrant and the lender at time $t$ is a triple $(i_{t+j}(z^{t+j}), d_{t+j}(z^{t+j}), p_{t+j}(z^{t+j}))_{j=0}^{\infty}$ of sequences specifying the investment $i_{t+j}$, the dividend distribution $d_{t+j}$ and the payment to the lender $p_{t+j}$ as
functions of the firm’s current history. We allow \( p_t \) to be either positive or negative, with positive amounts corresponding to repayments to the intermediary and negative amounts corresponding to additional external financing. The contract is thus fully state contingent.

Of course, real-world contracts do not literally specify policies in this manner. Nonetheless, all loan and debt contracts contain covenants that often specify limits on investment and dividend policies, and these sorts of limits constitute financial frictions. The model therefore captures these sorts of endogenous frictions in an internally consistent manner. The state-contingent nature of the contract seems at first unnatural because debt contracts are typically thought of as being state-incontingent. However, Roberts (2012) documents that most loan contracts get renegotiated multiple times over their lifetime, and the mere existence of debt covenants implies that debt cannot, by definition, be completely state incontingent. Finally, some debt instruments, such as credit lines, are by nature state contingent.

We define a contract to be feasible if it meets the following two conditions:

\[
\begin{align*}
z_{t+j} k_{t+j}^\alpha - i_{t+j}(z^{t+j}) &\geq d_{t+j}(z^{t+j}) + p_{t+j}(z^{t+j}) \\
d_{t+j}(z^{t+j}) &\geq 0
\end{align*}
\]

for any \( z^{t+j}, j \geq 0 \). The constraint (2) is simply the budget constraint, which requires that net revenue be at least as large as payments to shareholders and the lender. The constraint (3) is the result of limited liability. It prevents the firm from obtaining costless external equity financing from shareholders. Without such a constraint, the contract would be unnecessary. In this detail, the model departs from dynamic investment-based capital structure models such as Hennessy and Whited (2005), in which the firm can extract negative dividends from shareholders, but only after paying them a premium. As such, our model cannot capture the equity issuances we see in the data. However, given that firm initiated equity issuances are both tiny and rare (McKeon 2013), we view this drawback of our model as minor.

We assume that the long-run contract is not fully enforceable. The firm has control of its capital and has the option to renege on the contract and default, leaving the lender with no further payments on the loan and thus setting its liability to zero. If the firm defaults,
it can keep a fraction \((1 - \theta)\) of the capital \(k_t\). At this point, the firm is not excluded from
the market. Instead, it can reinvest the capital and sign a new lending contract. Thus, the
form of punishment for the defaulting firm is only the loss of a fraction \(\theta\) of its assets. This
feature of the model captures Chapter 11 renegotiation, rather than Chapter 7 liquidation.

The interpretation of the parameter \(\theta\) is worth discussion. It can be thought of not only
as the firm’s loss in default, but also as the fraction of assets that can be collateralized,
that is, that can be relinquished to the lender in default. Further, as in Rampini and
Viswanathan (2013), \(\theta\) is the fraction of tangible assets that can be pledged to the lender.
Thus, in our estimation, this parameter captures both the tangibility of the firm’s assets and
the pledgability of those tangible assets.

Let \(E(\cdot)\) be the expectation operator with respect to the transition function \(\Pi\). Then the
total value to the firm of repudiating an active contract at time \(\tau\) is:

\[
D(k_\tau, z_\tau) = E_\tau \sum_{j=0}^{\infty} (\beta(1 - \phi))^j \hat{d}_{\tau+j} (z_{\tau+j}),
\]

in which \(\{\hat{d}_{\tau+j} (z_{\tau+j})\}_{j=0}^{\infty}\) is the dividend stream the firm obtains at time \(\tau\), with a fraction
\((1 - \theta)\) of its capital stock, after it repudiates its original contract and then enters into a new
contract. The diversion value in (4) is a primitive of the model and constitutes the equity
value of reinvesting the diverted capital. Equivalently, this sum is the contract value for the
shareholder with capital \((1 - \theta)k_\tau\). Note that we implicitly assume that the firm absconds
with capital after production.

Because the lender commits to the contract, in order for the contract to be self-enforcing,
the firm cannot have any incentive to deviate from its terms. Therefore, the discounted
dividends from continuing the contract should be no less than the repudiation value. That
is, the firm will not renege on the contract at time \(\tau\) provided that

\[
D(k_\tau, z_\tau) \leq E_\tau \sum_{j=0}^{\infty} (\beta(1 - \phi))^j d_{\tau+j}.
\]

The contract is then self-enforcing/enforceable if (5) is satisfied for all \(\tau > t\).
2.3 Contracting problem

The optimal contract maximizes the equity value of the firm subject to several constraints that define the contract. This problem for an entrant is defined as follows:

\[
\max_{\{d_{t+j}, i_{t+j}, p_{t+j}\}} \mathbb{E}_t \sum_{j=0}^{\infty} (\beta(1-\phi))^j d_{t+j} \quad (6)
\]

subject to

\[
d_{t+j} \geq 0 \quad (7)
\]
\[
z_{t+j} k_{t+j}^\alpha - i_{t+j} - p_{t+j} - d_{t+j} \geq 0 \quad (8)
\]
\[
\mathbb{E}_\tau \sum_{j=0}^{\infty} (\beta(1-\phi))^j d_{\tau+j} \geq D(k_{\tau}, z_{\tau}), \quad \forall \tau > t \quad (9)
\]
\[
\mathbb{E}_t \sum_{j=0}^{\infty} (\beta_C(1-\phi))^j p_{t+j} \geq 0 \quad (10)
\]

Equations (7)–(9) are the dividend nonnegativity constraint, the budget constraint, and the enforcement constraint. Equation (10) is the initial participation constraint for the lender. Intuitively, the lender will only enter the contract with the firm if it expects the present value of its disbursements and repayments to be nonnegative. Note that the lender discounts these payments at a lower rate than the firm.

2.3.1 Collateral constraint

Because of the presence of the future contract value in the enforcement constraint (9), the model given by (6)–(10) is difficult to solve. Therefore, as a first step, we follow Alvarez and Jermann (2000), Bai and Zhang (2010), and Rampini and Viswanathan (2013) by replacing the enforcement constraint with an endogenous borrowing constraint. To start, we define

\[
q_\tau \equiv \mathbb{E}_\tau \sum_{j=0}^{\infty} (\beta_C(1-\phi))^j p_{\tau+j},
\]

which is the contract value for debt owners, or the value of promised debt at time \( \tau \). The collateral constraint is then defined as follows:

\[
\theta k_{\tau}(1-\delta) \geq q_\tau(z_{\tau}), \quad \forall \tau > 0 \quad (11)
\]
Next, we construct a transformed problem that maximizes (6) subject to (7), (8), (11), and (10). The following proposition shows that the solution to this transformed problem equals the solution to the original problem.

**Proposition 1** A sequence of \( \{k_{t+j+1}, \{q_{t+j+1}(z_{t+j+1})\}_{z \in \mathbb{Z}}\}_{j=0}^{\infty} \) is optimal in the original problem given by (6)–(10) if and only if it is optimal in the transformed problem given by (6), (7), (8), (10), and (11).

As stressed in Rampini and Viswanathan (2013), the constraint (11) can also be interpreted as a collateral constraint, so that \( \theta \) represents the fraction of assets that can be pledged as collateral in default. Note that the assumption that the lender seizes some of the firm’s assets in default does not alter the basic form of the problem, because the lender commits to the contract. The interpretation of (11) as a collateral constraint embodies many commonly observed borrowing practices. Most loans are drawn with the specific stated purpose of spending the proceeds on an asset, and some are secured by the asset. In addition, credit lines and term loans often have an upper limit that is contingent on what is called a borrowing base. The base consists of a set of pledgeable assets, usually current assets such as inventory or accounts receivable. The value of this base can vary over time. (Taylor and Sansone 2006) Thus, this collateral constraint conforms to the types of actual financial contracts we observe in the data.

### 2.4 Recursive formulation

Proposition 1 shows the equivalence between the original problem and the transformed problem with (11) in place of (9). As in Spear and Srivastava (1987) and Abreu, Pearce, and Stacchetti (1990), we now rewrite the original problem with the enforcement constraint (9) recursively using \( q_r \) as a state variable.

\[
V(k, q, z) = \max_{k', q'(z')} zk'^\alpha + k(1 - \delta) - q - k' + \beta_C(1 - \phi)E q(z') + \beta(1 - \phi)E V(k', q'(z'), z') \tag{12}
\]
subject to

\[ zk^\alpha + k(1 - \delta) - q - k' + \beta C(1 - \phi)E_q(z') \geq 0 \quad (13) \]

\[ \theta k'(1 - \delta) \geq q(z'), \quad \forall z' \in \mathbb{Z}, \quad (14) \]

in which a prime denotes a variable in the subsequent period, and no prime denotes a variable in the current period.

We now simplify the problem by reducing the dimension of the state space. Define net wealth as

\[ w \equiv zk^\alpha + k(1 - \delta) - q. \]

It is straightforward to show that the solution to the above problem depends only on this variable and not on its individual components. To see this property of the solution, note that without the constraints (13) and (14), the solution to the unconstrained optimization (12) does not depend on both \( k \) and \( q \) because the firm postpones the debt payment and always choose the highest possible debt, \( \bar{q} \), as \( \beta C > \beta \). In this case the total value of the firm does not depend on how much of it is financed with debt. In the case of a constrained problem, \( k \) and \( q \) appear in the constraint (13) only to the extent that they define net wealth. Thus, the recursive problem can be rewritten as follows:

\[ V(w, z) = \max_{k', q(z')} \left( w - k' + \beta C(1 - \phi)E_q(z') + \beta(1 - \phi)E_V(w'(z'), z') \right) \quad (15) \]

subject to

\[ w - k' + \beta C(1 - \phi)E_q(z') \geq 0 \quad (16) \]

\[ \theta k'(1 - \delta) \geq q(z'), \quad \forall z' \in \mathbb{Z}, \quad (17) \]

Next, define the mapping \( T \) in the space of bounded functions.

\[ T(V)(w, z) = \max_{k', q(z')} \left( w - k' + \beta C(1 - \phi)E_q(z') + \beta(1 - \phi)E_V(w'(z'), z') \right) \quad (18) \]

subject to (16) and (17). The following proposition establishes the existence of a solution.

**Proposition 2** Let \( C(X) \) be the space of bounded continuous functions. The operator \( T \) defined in (18), which maps \( C(X) \) to itself, has a unique fixed point \( V^* \in C(X) \); and for all \( v_0 \in C(X) \).
This proposition is also useful because it implies that the solution to the model can then be obtained by iterating on (18). Appendix B contains our numerical solution procedure.

2.5 Optimal Policies

To understand the properties of the model, it is useful to study the first order conditions. To do so, we first assume that \( V(w, z) \) is differentiable. Next, let \( \mu \) be the Lagrange multiplier on the dividend nonnegativity constraint (16), and let \( \beta (1 - \phi) \pi(z'|z) \lambda_{z'} \) be the Lagrange multiplier associated with the enforcement constraint (11) at state \( z' \), where \( \pi(z'|z) \) is the transitional probability from state \( z \) to state \( z' \). The first order condition for \( k' \) is:

\[
\beta (1 - \phi) \sum_{z'} \pi(z'|z) \left( V_w(w', z') \frac{\partial w'}{\partial k'} + \lambda_{z'} (\theta(1 - \delta)) \right) - \mu = 1
\]

where \( \frac{\partial w'}{\partial k'} = z' \alpha k'^{\alpha - 1} + 1 - \delta \). The first term in (19) is the constrained ratio of the marginal product of capital to the user cost. Suppose that the Lagrange multipliers \( \mu \) and \( \beta (1 - \phi) \pi(z'|z) \lambda_{z'} \) are zero (the unconstrained case). Because the envelope theorem implies that \( V_w(w, z) = 1 + \mu \), this first order condition just states that the expected marginal product of capital equals the user cost, as in a standard neoclassical investment model.

We now consider the constrained case. The next term in (19) is the marginal value of capital in relaxing the enforcement constraint. As long as \( \theta > 0 \), and as long the constraint binds in at least one state, this term is strictly positive. The last term is the shadow value of the dividend nonnegativity constraint. Thus, capital has value not only in the production of goods, but also in the relaxation of the enforcement and dividend nonnegativity constraints.

Next, we examine the optimality conditions with respect to the value of payments to the lender. The first order condition for \( q(z') \) for any given value of \( z' \) is

\[
1 + \mu + \frac{\beta}{\beta_C} \left( V_w(w', z') \frac{\partial w'}{\partial q'} - \lambda_{z'} \right) = 0, \quad \forall z' \in \mathbb{Z}
\]

Using the envelope theorem and the condition \( \frac{\partial w'}{\partial q'} = -1 \), we can rewrite (20) as

\[
1 + \mu = \frac{\beta}{\beta_C} (1 + \mu(w', z') + \lambda_{z'}) \quad \forall z' \in \mathbb{Z}
\]
First, note that when \( \mu = \mu(w', z') = 0 \), because \( \beta < \beta_C \), the enforcement constraint binds. In other words, the assumption that the firm is less patient than the lender indicates even mature firms that pay dividends can be constrained because they always want to borrow more. However, if \( \mu = 0 \), but \( \mu(w', z') \neq 0 \) for some \( z' \in \mathbb{Z} \), the collateral constraint does not bind. This situation is likely to occur if the firm is currently in low state. In this case, the contract specifies that the firm conserve debt capacity in those states of the world in which \( z' \) is high because the contract value is maximized by the firm using its resources to invest rather than to pay back the lender.

2.6 Taxes

Thus far we have worked with a model with no taxation. We now consider the possibility that profits are taxed and that interest on debt is tax deductible. The taxation of profits can be modeled simply by replacing the profit function, \( zk^\alpha \) with \((1 - \tau_c)zk^\alpha\), in which \( \tau_c \) is the corporate tax rate.

Interest deductibility is somewhat more complicated because debt in the model is defined in terms of payments to the intermediary, which can include both principal and interest components. For simplicity, we assume that the firm receives a rebate from the tax authority equal to \( q\tau_c(1 - \beta_C) \), where \( q(1 - \beta_C) \) is defined to be the interest component of the payment to the intermediary. Under this taxation assumption, we define net wealth as

\[
w^* \equiv (1 - \tau_c)zk^\alpha + k(1 - \delta) - (1 - \tau_c(1 - \beta_C))q.
\] (22)

Now we can rewrite the recursive problem as:

\[
V(w^*, z) = \max_{k', q(z')} w^* - k' + (1 - \phi)\mathbb{E}q(z') + (1 - \phi)\mathbb{E}V(w'^*, z')
\] (23)

subject to

\[
w^* - k' + (1 - \phi)\mathbb{E}q(z') \geq 0
\] (24)

\[
\theta k'(1 - \delta) \geq (1 - \tau_c(1 - \beta_C))q(z'), \quad \forall z' \in \mathbb{Z}.
\] (25)

For this model, the proofs of Propositions 1 and 2 proceed with only minor modification.
Taxation affects optimal debt via several different channels. The first is via the collateral constraint (25). Intuitively, for the firm to be indifferent between defaulting and continuing operations, the benefit from shedding debt in default must equal the loss of revenue that follows from the destruction of capital in default. The tax deductibility of interest then loosens the constraint by rendering debt less onerous for the firm.

To see the effect of taxes on an unconstrained firm, we examine the first-order condition for optimal debt, which is given by:

$$1 + \mu = \frac{\beta}{\beta_C} \left( (1 + \mu(w^{*l}, z'))(1 - \tau_c(1 - \beta_C)) + \lambda_{z'} \right) \quad \forall z' \in Z.$$  

(26)

To understand the intuition behind (26), we begin by fixing optimal capital policy relative to a model without taxes. In this setting, increasing the corporate tax has two effects on (26). First, a positive corporate tax rate lowers profits in all states of the world. Therefore the limited liability constraint binds in more states of the world, and the shadow value of dividends, (i.e., the Lagrange multiplier, $\mu(w^{*l}, z')$) is higher for any level of $w^{*l}$. Because $w$ is decreasing in $q$, this effect implies that optimal debt should be lower. The intuition is that when taxes eat away at internal profits, the debt capacity preservation motive becomes stronger. The second effect counteracts the first. The presence of a corporate tax effectively makes the firm more impatient relative to the lender. This effect is akin to the traditional tax-benefit of debt and naturally increases optimal debt. The net effect on debt is theoretically ambiguous. Compounding this ambiguity is the observation that optimal capital policy is affected by corporate taxation, with the optimal size of the firm shrinking. We therefore turn to a quantitative analysis.

### 2.7 Policy Functions

To expand upon the intuition behind the model, we examine the policy functions from two versions of the model. The first uses the parameterization from the estimation in Table 1 in which the tax rate is set to zero. The second uses the same parameterization, except that we set the tax rate to 0.2. This exercise is explicitly quantitative because we compute the
model solution using estimated parameters. Figure 1 contains the policy functions the no-tax model, where we plot several optimal policies as a function of current net wealth for the case of the median current shock. The optimal choices include capital \((k')\), dividends \((d)\), and state contingent debt \((q'(z'))\), where we consider debt contingent on a low, a medium, and a high future state. Each of these variables is scaled by the steady state capital stock, defined as the capital stock that equates the expected marginal product of capital \(E(\alpha z'(k')^{\alpha-1})\) with the user cost, which is given by \(\delta + 1/\beta - 1\).

Three patterns stand out. First, if the firm has low net wealth, the optimal choice of tomorrow’s capital stock is increasing in today’s net wealth. Because of the binding enforcement constraint, the firm can only pick a higher capital stock if it has sufficient internal resources. However, for high levels of net worth, the firm does not expand beyond the point where the marginal product of capital equals the user cost. Second, for low levels of net wealth, the firm does not pay dividends, because resources devoted to capital accumulation earn more than the user cost and because increasing capital helps relax the borrowing constraint. It only pays dividends at high levels of net worth when the firm has more than enough internal resources to fund optimal capital expenditures.

The third and most important feature of Figure 1 pertains to leverage. When the firm has low net wealth, the enforcement constraint always binds, as seen in the proportionality of debt to capital for all three state-contingent debt policies. At low levels of net worth, the constraint binds because the firm has not reached an optimal size and borrows as much as it can to grow to out of its borrowing constraint. At higher levels of net worth, the enforcement constraint does not always bind when debt is contingent on the high or the medium shock. In these cases, the firm knows it will have good investment opportunities, which increase the value of unused debt capacity, which the firm then preserves accordingly. However, the collateral constraint always binds for debt contingent on the low state. In this case, the value of preserving debt capacity is not sufficient to offset the fact that the lender is more patient than the firm, and the firm ends up at a corner solution where the collateral constraint binds. Finally, it is worth noting that in this model debt can take negative values, so that the model
does indeed allow for positive cash holdings.

Figure 2 plots the same policy functions for a model in which all parameters are identical, except that we set the tax rate to 0.2. As in Figure 1, we scale all variables by the capital stock that sets the expected *before-tax* marginal product of capital equal to the user cost. Strikingly, Figures 1 and 2 appear almost identical, except that taxes lower the marginal product of capital and thereby cause the firm to operate on a much smaller scale, as can be seen in the scales of the $x$ and $y$ axes. Except for this difference in scale, the policies for a taxed firm are quantitatively almost identical to those for an untaxed firm. The results for debt indicate that for a parameterization of the model that best fits the data, the various positive and negative effects on taxes on debt discussed in Section 2.6 above have negligible net impact.

### 3 Data

Our data are from the 2013 Compustat files. Following the literature, we remove all regulated utilities (SIC 4900-4999), financial firms (SIC 6000-6999), and quasi-governmental and non-profit firms (SIC 9000-9999). Observations with missing values for the SIC code, total assets, the gross capital stock, market value, debt, and cash are also excluded from the final sample. We also delete firms with fewer than three consecutive years of data. As a result of these selection criteria, we obtain a panel data set with 82,667 observations for the time period between 1965 and 2012 at an annual frequency.

We define total assets as Compustat variable AT, the capital stock as PPEGT, investment as capital expenditures (CAPX) minus sales of capital goods (SPPE), operating income as OIBDP, equity repurchases as PRSTK, dividends as the sum of common and preferred dividends (DVC + DVP), and Tobin’s $q$ as the ratio of $(\text{AT} + \text{PRCC_F} \times \text{CSHO} - \text{TXDB} - \text{CEQ})$ to AT. We define total debt as $(\text{DLTT} + \text{DLC})$ plus the capitalized value of operating leases, which can be substantial relative to traditional debt. For example, for many airlines, the value of leases is much larger than traditional debt. To compute this present value we
discount reported lease payments in years one through five (MRC1–MRC5) at the Baa bond rate. We similarly discount lease payments reported due in years beyond the fifth (MRCTA) by assuming that they are spread out evenly until year ten. Investment, debt, total payout (dividends plus repurchases), and operating profit are expressed as fractions of total assets.

Our definition of leverage inclusive of lease payments is of particular interest because the solutions to dynamic contracting models are typically unique, but the implementations of these solutions in terms of observable variables are not unique. To see the connection between our definition of debt and the uniqueness of an implementation, it useful to consider a slightly different class of contracting models based on dynamic moral hazard (e.g. DeMarzo and Sannikov 2006). In these models, the contract is typically specified as a function of the agent’s continuation utility. A variety of different capital structures can then implement this unobservable level of utility. This feature of dynamic contracting models poses a challenge for structural estimation because it is impossible to know which implementation corresponds to the actual capital structures we observe in the data.2

The estimation of our model faces a similar challenge. For example, Rampini and Viswanathan (2010) show that the contract from a model similar to ours can be implemented with a combination of straight debt and risk management. Similarly, the contract could be implemented with state contingent lines of credit. Fortunately, the issue of a non-unique contract implementation is surmountable in the case of our model, because the contract is specified as a function of the present value of payments to the lender, which is, in principle, observable. To measure this total debt value, we must include debt-like instruments, including lease payments. Because all of the debt instruments included in our measure receive the same tax treatment, we can use this measure of total debt to identify model parameters. In this case, however, we cannot make empirical predictions about the composition of debt, nor can we use the composition of debt to identify model parameters.

2See Bresnahan and Reiss (1991) or, more recently, Aguirregabiria and Mira (2007) for a discussion in the context of industrial organization of how lack of uniqueness poses important problems for structural estimation.
4 Estimation

This section provides an intuitive description of our estimation procedure and discusses the identification of our parameters. Appendix B contains the technical details.

4.1 Simulated Method of Moments

We estimate most of the structural parameters of the model using simulated method of moments. However, we estimate some of the model parameters separately. For example, we estimate $\beta$ as $1/(1 + r_f)$, where $r_f$ is the average real 3-month Treasury bill rate over our sample period. For our simulations we also need to choose the number of years the firm lives. Here, we use the simple average of the length of time a firm is in our sample, which we truncate to the nearest integer, 23. We then estimate the following parameters using simulated method of moments: the depreciation rate, $\delta$; the production function curvature, $\alpha$; the fraction of the capital stock that can be diverted, $\theta$; and the difference between the creditor’s and the firms’ discount factors, $\beta_C - \beta$. To estimate the transition matrix, $\Pi$, we approximate it as an $AR(1)$ process in logs, given by

$$\ln z_{t+1} = \rho \ln z_t + \varepsilon_{t+1}. \quad (27)$$

Here, $\varepsilon_t$ is an i.i.d truncated normal variable with mean 0 and standard deviation $\sigma_z$. With this assumption, we add two more parameters to our list: the standard deviation and serial correlation of the productivity shock, $\rho$ and $\sigma_z$. Finally, we estimate the time-zero capital stock, $k_0$, but we express it as a fraction of the steady state capital stock at which the after-tax marginal profit of capital equals the user cost.\textsuperscript{3}

We define our simulated data variables as follows. Investment is $(k' - (1 - \delta)k)/k$; future leverage in state $z'$ is $q(z')/k'$; current leverage is $q/k$; dividends are $d/k$, and Tobin’s $q$ is $V(w, z)/k$.

\textsuperscript{3}We set the parameter $\phi$ to 0.0001 because attempts to estimate this parameter nearly always result in a small parameter estimate and always result in a large standard error. Also, the parameter $\phi$ has little effect on any of the moments we use in our estimation.
Simulated method of moments, although computationally cumbersome, is conceptually simple. First, we generate a panel of simulated data using the numerical solution to the model. Next, we calculate interesting moments using both these simulated data and actual data. The objective of SMM is then to pick the model parameters that make the actual and simulated moments as close to each other as possible.

4.2 Identification

The success of this procedure relies on model identification, which requires that we choose moments that are sensitive to variations in the structural parameters such as the diversion parameter, \( \theta \). On the other hand, we do not “cherry-pick” moments. Instead, we examine moments of all of the observable variables contained in our model.

We now describe and rationalize the 10 moments that we match. The first 8 are the means and standard deviations of the rate of investment, \((k' - (1 - \delta)k)/k\), the ratio of profits to assets \((zk^{\alpha-1})\), the leverage ratio \((q/k)\), and the ratio of dividends to assets \(d/k\). We also include the mean of Tobin’s \(q \,(V/k)\), but not the standard deviation because the model-implied standard deviation of Tobin’s \(q\) is much smaller than the values seen in the data. This result is shared by many production-based asset pricing models, which cannot reconcile the observed high volatility of asset prices. Finally, we also include the serial correlation of operating profits.

In this list several of these moments are particularly useful for identification of specific parameters. We start with the technological parameters, all of which are straightforward to identify. First, the mean rate of investment is the moment most useful for pinning down the depreciation rate, with higher rates of depreciation naturally leading to higher rates of contractual capital replacement. Next, the standard deviation and autocorrelation of profits are directly related to the parameters \(\sigma_z\) and \(\rho\). Finally, the curvature of the production function, \(\alpha\) is most directly related to average profits and Tobin’s \(q\). As \(\alpha\) decreases, the firm faces more severe decreasing returns to scale, which, all else held constant, results in lower average profits. However, rents to capital increase, so Tobin’s \(q\) rises.
The identification of $\theta$ is also straightforward because it is strictly increasing in the average leverage ratio. The identification of $k_0$ involves firm dynamics. If $k_0$ is small, then the firm starts its life far away from its desired capital stock. Because the firm pays no dividends while it is young, average dividends naturally rise with $k_0$, as the firm spends less time in a region in which it is constrained. The most difficult parameter to identify is the difference between the firm and lender discount factors $\beta_C - \beta$. Although leverage does naturally rise as this difference moves away from zero, as shown below in Section 6, these effects occur largely at the transition from a zero difference to a positive difference. In addition, none of the other moments are affected by this parameter. Although this parameter is hard to identify, the positive side of this problem is that its value matters little for our basic conclusions.

5 Results

Table 1 contains the results from our estimation. We consider two versions of the model: one in which we set the corporate tax rate to zero and one in which we set it to 20%. Panel A contains estimates of the real-data moments, the simulated moments, and the t-statistics for the difference between the two. Panel B contains the parameter estimates.

Two main results stand out in Panel A. First, both versions of the model fit the data reasonably well. Across the two estimations, just under half of the simulated moments are statistically significantly different from their real-data counterparts, but only a few are economically different. Both versions of the model do a good job of matching the means of leverage, investment, operating profit, and Tobin’s $q$. Although the average of simulated distributions is well-matched in the no-tax model, it lies at half its actual level in the model with taxes. In contrast, the average of Tobins’ $q$ is well-match in the model with taxes, but not in the model without taxes, in which is it too high. Both models struggle more with standard deviations. Both models underestimate the standard deviations of distributions and profits, but overestimate the standard deviation of investment. The standard deviation of leverage is only slightly underestimated. In the end, by fitting a large number of moments,
we have stress tested the model to determine whether and where it succeeds in matching important features of the data.

Our second main result is that adding taxes to the model does little to help reconcile the model with the data, especially average leverage, which is well matched in both models. Intuitively, debt is, by nature, akin to a constant returns technology for transferring resources through time, that is, a storage technology. Constant returns to scale implies that even the tiniest difference between the lender and borrower discount factors makes borrowing attractive. Constant returns then also implies that the additional attractiveness of debt conferred by the interest tax deduction is then quantitatively limited.

Panel B shows that our estimates of fundamental contracting frictions are significantly different from zero. In the no-tax model, we estimate that the average firm loses 36% of its assets as a deadweight cost when it repudiates a contract. Equivalently, 36% of the assets can serve as collateral. This figure is slightly higher for the model with taxes. Both of these parameter estimates are noticeably higher than our estimates of average leverage. This result implies that firms do not always hug the collateral constraint. These estimates are somewhat higher than the direct estimates deadweight bankruptcy costs in Andrade and Kaplan (1998). We attribute the difference to our wider definition of leverage that also includes leases.

We also find that the starting capital stock in both models is quite low, with the firm starting life with less than 25% of its eventual steady state capital stock. To assess the plausibility of this parameter, we calculate the ratio of a firm’s real assets in the final year in our sample to real assets the in the first year the firm appears in the sample. We find that for firms in our data set for at least 23 years (the median length in our sample), this number is 0.083. Although slightly lower than the number we estimate from the model, this figure is qualitatively comparable to our model estimate, with both figures being low. Interestingly, although we find positive estimates for the difference between the lenders’ and firms’ discount rates, the estimates are insignificantly different from zero. The large standard errors arise because this parameter is extremely hard to identify, but the difficult
in identification also has a positive side. The moment estimates are largely insensitive to the value of this parameter.

Our estimates of the technological parameters, \( \delta, \alpha, \rho, \) and \( \sigma_z \) are slightly different from those found in previous studies such as Hennessy and Whited (2005, 2007). In particular, the estimated depreciation rate is much lower. This result makes sense, because the models used in Hennessy and Whited (2005, 2007) are of mature firms, in which the depreciation rate roughly equals average model-generated investment. In contrast, in the current model, firms start out at a suboptimally low level and then grow into mature firms. In this setting, young firms invest at enormous rates, while mature firms invest at very low rates. The result is that mean investment lies far above median investment, with the latter approximately equal to the depreciation rate. The production function curvature, \( \alpha \), at over 0.9, is somewhat higher than the estimates from these previous studies. The other two technological parameters, \( \rho \) and \( \sigma_z \), are largely in line with the estimates from this previous work.

\[ 5.1 \quad \text{Industry Estimation} \]

To examine groups of firms that are ex ante homogeneous, we also estimate the model on the twenty four two-digit industries that have at least 1,000 firm-year observations. These industries are listed in Table 2. The results for moment matching are in Figure 3, which shows that the model can match the wide variation across industries in leverage: 0.52 in Air Transportation to 0.18 in Metal Mining. Interestingly, the low-leverage industries have the lowest asset tangibility, as measured by the ratio of net property, plant, and equipment to total assets, and the highest leverage industry, Air Transportation, has the third highest measure of asset tangibility. Indeed, the correlation between this measure of asset tangibility and our estimates of \( \theta \) is 0.61. This piece of ancillary evidence is comforting in that one of the implications of our model is that collateral is one of the main determinants of leverage. It is also in accord with the reduced form evidence in Erickson, Jiang, and Whited (2013) that the coefficient on collateral in a standard leverage regression is near and insignificantly different from 1. Although the model can match the wide variation in investment and Tobin’s
$q$ across industries, it struggles with matching average distributions, with just under half of the industries well matched. For the rest of the industries, simulated average distributions fall short of their real-data counterparts.\footnote{For brevity, we omit the analogous plots for the rest of our moments. The results largely mirror those in Table 1.}

Although Table 2 presents all of the parameter estimates accompanying Figure 3, we highlight the main result in this table in Figure 4. This figure plots average simulated leverage versus the collateral constraint, which is given by $\theta(1 - \delta)$. We see that leverage in about one third of the industries hugs the collateral constraint. However, in most of the industries average leverage lies below the constraint, with the largest gaps in Trucking (42) and Oil and Gas Exploration (10). At around 0.08, these gaps are substantial. To explore the reasons for the variation in this conservation of debt capacity across industries, we examine the other parameter estimates. Those industries with high capacity preservation are those with highly serially correlated profits and with less curvature in their production functions. This second feature is particularly important because it induces a highly variable optimal investment policy. Thus, in those industries where investment opportunities are highly persistent and in which optimal investments themselves are highly variable, debt capacity is naturally more valuable.

It is worth examining the amount of debt capacity preservation in our model versus the class of closely-related neoclassical investment models, such as (Hennessy and Whited 2005) or DeAngelo, DeAngelo, and Whited (2011). In these models, leverage lies much further below the collateral constraint. One main difference between our model and these two models underlies our result. The solution to the contracting problem in our model is state-contingent debt, whereas neoclassical investment models typically contain straight debt. When the firm can only use straight debt, it has a greater incentive to conserve debt capacity because it has no ability to alter its repayment schedule as a function of its future productivity. In contrast, with state-contingent debt, the need to manage risk by delevering is much lower because the firm can engage in explicit risk management.
5.2 Natural Experiment

In this section, we reestimate the model using data surrounding a quasi-natural experiment involving the repossession of assets in bankruptcy, and thus the value of these assets as collateral. The purpose of this exercise is twofold. First, we wish to determine whether our model can detect this change in the value of collateral. Conversely, we want to use our model to ascertain whether other important determinants of leverage changed at the same time. Thus, we both use the natural experiment to test the model and allow the model to illuminate the results from the natural experiment.

The experiment we consider is the enactment of anti-recharacterization laws in seven states in the late 1990s and early 2000s. These laws are relevant for any firm that uses a Special Purpose Vehicle (SPV) to conduct secured borrowing. Instead of borrowing directly from the lender, the firm (the originator) first transfers collateral to an SPV, who has limited risk exposure and remains solvent even if the originator files for bankruptcy. The advantage of this transfer is that in bankruptcy, the secured lender is protected from the automatic stay because of the bankruptcy-remote status of the SPV. This status then enables the lender to seize the collateral without any delay. However, prior to the laws we consider, this strategy was not guaranteed to succeed because bankruptcy courts had the discretion to recharacterize the asset transfer to the SPV as a loan instead of a true sale. Once the recharacterization takes place, the lender becomes a secured creditor of the originator instead of the SPV. In addition, the collateral is protected by the automatic stay and goes back to the bankrupt firm. Recharacterization is likely to take place under Chapter 11 if the court believes that the collateral plays a key role in the originator’s operation during the reorganization process. In this case, the lender cannot recover its claim until the originator is liquidated or emerges from the restructuring. The automatic stay delays the lenders’ seizure of the collateral and creates substantial uncertainty regarding the value of the collateral that is eventually recouped by the lenders.

The anti-recharacterization laws that we explore require collateral transfers to SPVs to
be treated as true sales if they are labeled as such. These laws thus strengthen creditors’ rights by enabling the swift seizure of collateral. In contrast, the laws add to the costs of bankruptcy ($\theta$ in our model) and relax their enforcement constraint (25). This set of laws was officially introduced in Texas and Louisiana in 1997, followed by Alabama in 2001, Delaware in 2002, South Dakota in 2003, Virginia in 2004, and Nevada in 2005.

The passage of the state laws enhances the pledgability of assets for firms incorporated in those states. At least as relevant for the repossession of assets that have been transferred to SPVs is the Reaves Brokerage Co. v. Sunbelt Fruit and Veggie. Co. case in 2003. In this case, the court recharacterized the debtor’s transfer and prevented the creditor from seeking recovery after the debtor filed for bankruptcy. The importance of the case is that the court completely ignored the anti-recharacterization statute of Texas and used a federal standard to determine the nature of the sale. This specific court decision increases the likelihood that the federal law will preempt state-level property rights when the debtor goes bankrupt. Therefore, the effect of passing an anti-recharacterization law at a state level shortly before or after this case law should be limited.

This institutional setting allows us to construct a difference-in-difference specification as follows. We classify a firm as “treated” if it is incorporated in the three states that passed the law before the year 2002. We then restrict our “after” period to those years after a specific state passed the law but before 2003.

Although the timing of the court case is plausibly exogenous to the firms we study, the enactment of the laws is likely to be influenced by lobbying activities. However, as discussed in Kettering (2008), these lobbying efforts were mainly concentrated among the banking and especially the securitization industries, rather than the industrial firms in our sample. Kettering (2008) describes the large role played by the securitization industry in forcing the legislative change, which elevated the popularity of the asset backed collateral and provided a larger market for structured financial products. Similarly, Janger (2003) argues that lobbying activities by banks, bond lawyers, and rating agencies contributed significantly to the passage of state-level anti-recharacterization statutes. Interestingly, the impetus for this
lobbying was not economic in nature, but legal, specifically, the LTV bankruptcy case in 1993, in which the bankruptcy court allowed the firm to pull back its collateralized assets and use them as working capital during the early part of its Chapter 11 process. In contrast, the industrial firms who act as originators and borrowers appear to have had a limited role in the law change. Indeed, Janger (2003) explains that they are likely to be the source of challenges to their own previous asset-transfer transactions when they ask the bankruptcy courts to recharacterize their sales of collateral to SPVs. In the end, although we cannot completely rule out political economy considerations in our difference-in-difference experiment, the background for the enactment of anti-recharacterization laws makes it difficult to believe that political economy considerations affect our experimental design.

As a first step in this investigation, we conduct a standard reduced-form difference-in-difference exercise. Without any controls, the difference-in-difference effect is 0.04 with a \( t \)-statistic of 2.5. This effect is economically quite large, and its statistical significance is notable, considering that we have fewer than 200 firm-year observations in the “treated-after” group. To assuage concerns that this result is driven by unobservable heterogeneity, we re-run the analysis using firm and year fixed effects, as well as the following four firm characteristics from Rajan and Zingales (1995) as controls: the log of sales, the market-to-book ratio, the ratio of operating income to assets, and the ratio of net property plant and equipment to assets. In this case, we find that the difference-in-difference effect is again 0.04 with a \( t \)-statistic of 2.0.

We now use our model to ascertain the forces driving this change in leverage by estimating the model (with taxes) on the treated and control groups in the before and after periods.\(^5\) The results are in Table 3. First, note that in all four samples, leverage is well matched. Without this result, it would be hard to claim that our model could illuminate the results of the experiment. Second, the estimates of \( \theta \) show that a change in the value of collateral is indeed behind the large difference-in-difference effect on leverage. If we do a similar analysis on the collateral parameter, \( \theta \), we find that experiment produces a difference-in-difference

\(^{5}\) The results from the model without taxes are nearly identical.
effect of 0.07 with a $t$-statistic of 2.7. Why is the effect on collateral larger than the effect on leverage? Answering this question is where the interplay between the model and the data becomes useful. We see in Panel A of Table 3 that although the level of leverage stays the same both before and after the treatment in the control group, the standard deviation of leverage falls. In the model, leverage is more variable the further the firm situates itself from the collateral constraint. Thus, for the control firms in the after period, the value of collateral falls, but the firms hug the collateral constraint somewhat more closely, with no change in leverage.

We draw two conclusions from this experiment. First, our model is able to detect a plausibly exogenous change in the value of collateralized assets. Second, we are able to add texture to our reduced-form different-in-difference exercise, by using the model to measure the change in collateral, which we find is somewhat larger than the change in leverage. The reason for the difference is a concurrent change in debt capacity preservation. This second effect would have been impossible to uncover via reduced-form estimation alone.

6 Counterfactuals

We now examine what would happen to optimal financing if firms had different fundamental characteristics than those implied by the parameter estimates from Table 1. To this end, we consider a baseline simulated firm from the model without corporate taxation, i.e., $\tau_c = 0$. We then investigate results of changing three key parameters: $\tau_c$, $\beta_C - \beta$, and $\theta$, the latter two of which are the difference between the lender’s and the borrower’s discount factors and the collateralizable fraction of the firm’s assets.

The first set of results from these exercises is in Figure 5. To construct this figures, we pick a grid for the parameter in question. We then solve the model for each of the different parameter values, simulate the model for 7000 firms over 50 time periods, and then plot average leverage, investment, Tobin’s $q$, and dividends as functions of either the parameter in question. We perform separate analyses for two groups of simulated firm-year
observations. “Young” firms are those in years 1-10 of life, and “old” firms are those in years 11 through 50.

One main result stands out in Figure 5. In the top panel we see that leverage is largely unresponsive to taxes, either for young or old firms. However, mature firm leverage is slightly less than young firm leverage because the mature firms conserve debt capacity, while the young firms hug the collateral constraint.

Figure 5 contains several additional results. First, increasing the deadweight costs of repudiating a contract, $\theta$, leads directly to higher leverage. This intuition lies in sharp contrast to the traditional intuition that large default costs should lead to lower leverage. The difference is that if the firm can start over anew after contract repudiation with few consequences (i.e. low $\theta$), lenders have little incentive to extend the firm much credit. Equivalently, a lower level of $\theta$ tightens the collateral constraint. Second, the level of leverage does increase with the difference in discount factors, but only for mature firms, and largely when the difference between the discount factors starts to diverge from zero. For the young firms, the difference in discount factors does not affect optimal leverage because these firms always face a binding collateral constraint. However, for mature firms, as the lender’s patience approaches that of the firm, the opportunity cost of conserving debt capacity falls. It thus engages in more of this risk management behavior, and the average level of leverage falls. However, as the gap between $\beta_C$ and $\beta$ widens and the relative cost of external financing falls, this benefit quickly overshadows the value of preserving debt capacity, and average leverage rises, although never to the level of the collateral constraint. Thus, nearly all of the increase in leverage occurs when $\beta_C - \beta$ is less than 0.003.

This last result is particularly instructive about the role of taxes. The lender’s discount factor is $\beta_C$, and, as seen in (26), the borrower’s effective discount factor is $\beta(1 - \tau_c(1 - \beta_C))$. Indeed, Rampini and Viswanathan (2013) suggest that one reason for the difference between lenders’ and borrows’ discount factors might well be taxes. To see the tax implications, suppose that instead of requiring $\beta_C > \beta$ and setting $\tau_c = 0$, as we have done so far in Section 6, we set $\beta = \beta_C$ but require $\tau_c$ to be strictly positive. In this case, the borrower’s and
lender’s effective discount factors differ only because of taxes. We then can examine the role of taxes for leverage without any confounding factors. The results from this counterfactual are in Figure 6. Here we see that increasing the tax rate does affect optimal leverage, but that most of the effect of taxes on leverage occurs for tax rates less than 3%. Importantly, this result implies that if $\beta_C > \beta$ by even a small amount for any reason other than taxes, increasing the tax rate will do little to optimal leverage. As discussed above, frictions such as deposit insurance or the general equilibrium considerations in Ai et al. (2013) make the assumption that $\beta_C > \beta$ quite plausible.

The intuition is that debt can be thought of as a constant-returns storage technology that earns the after-tax rate risk-free rate. Thus, whenever the tax rate deviates from zero, because the firm discounts at the before-tax risk-free rate, the rate on the storage technology deviates from the discount rate. Because the storage technology has constant returns to scale, the model has an attenuated corner solution, with the effects on debt being large the moment the tax rate deviates from zero. However, the debt capacity preservation motive in the model implies that the solution is not always a strict corner solution, with debt hugging the collateral constraint.

Interestingly, these leverage counterfactuals are in line with other, more elaborate, neoclassical investment models (Hennessy and Whited 2005, 2007). For example, using the canonical model of this type in ?, we find that changing the corporate tax rate affects leverage only when the tax rate is low, as is the case here. The intuition is identical.

6.1 Robustness

Our conclusions are based on a very simple model. There are no equity issuances or lumpy investment, and the tax code is much simpler than the tax codes observed in the real world. To allay concerns that our results concerning taxes are due to these simplifications, we add these features, one by one, to the our model. Here, we modify the collateral-constraint version of the model instead of the original contracting problem, which has the value function in the constraints. First, instead of requiring dividends to be positive, we allow them to be negative,
but require the firm to pay a 5% equity issuance cost in this case. Second, investment in our model is not lumpy, as in the case of many real options models. To induce lumpiness in the model, we add a cost of adjustment equal to 0.5% of the current period capital stock. Because this cost is independent of the amount of investment, it induces optimal lumpy behavior, where the firm is inactive for long spells before investing a great deal. Third, we add a convex tax schedule to the model, in which the tax rate is 0.1 for when $z k^{a \alpha}_\alpha - q (1 - \tau_c (1 - \beta_C)) < 0.05$, and 0.2 otherwise. This gradation approximates the convexity of the U.S. tax code, which results in approximately 25% of listed firms paying a substantially lower tax rate than the statutory 0.35. In our simulated panel, approximately 25% of the firms also pay the lower rate. In all three cases, we find that changing the tax rates has almost no affect on optimal leverage. The one difference is that in the model with lumpy investment, the firm preserves debt capacity to a greater extent. We conclude that our results concerning taxes are robust to these concerns.

7 Conclusion

We have sought to deepen our understanding of whether corporate taxes or agency concerns are more important for capital structure. To this end, we estimate a dynamic contracting model in which financial constraints and capital structures arise endogenously as the result of contracting frictions. This approach departs from much of the structural estimation literature, which models financial constraints as exogenous to firm actions and policies. We produce four main findings. First, a model without taxes fits the data as well as a model with taxes. Second, when we parameterize the model according to our estimation results, we find that counterfactually varying the corporate tax rate has almost no effect on leverage. Third, our model can reconcile model-generated leverage with actual leverage across a broad spectrum of industries, and the estimated position of the collateral constraint is the primary determinant of the level of leverage. Finally, using a natural experiment in which the value of collateral changes exogenously, we find that our model can detect this change.
These results are the product of two intuitive features of the model. First, taxes have little quantitative effect on the constraint that allows the contract between the lender and the firm to be self-enforcing. This constraint is a collateral constraint, and taxes do not have an effect on the level of collateralizable assets. Although taxes do make debt less onerous, so that the constraint is less likely to bind, this effect is quantitatively small. Second, and more importantly, the optimal contract in our model calls for state contingent debt. When debt is state contingent, the firm has less need to conserve debt capacity than when debt cannot be adjusted as a function of the state of the world. Thus, even though the firm does not completely exhaust its debt capacity, it optimally conserves less debt capacity than it would if debt were not state contingent. Thus, the position of the collateral constraint determines the amount of debt capacity preservation, with the end result that taxes have little effect on leverage. Turning outside the model, we often see this type of risk management even when firms do not literally use derivatives. For example, as documented in Roberts (2012), firms often renegotiate their loans before maturity, and lines of credit are by nature state contingent.

Although our results are not in accord with two recent studies that find effects of taxation on capital structure (Heider and Ljungqvist 2012; Pérez-González, Panier, and Villanueva 2012), the results are in line with the vast majority of empirical studies, which find no effects of taxes on leverage (Graham 2007). The result is also in line with the recent evidence in Graham, Leary, and Roberts (2013) that the sharp tax increases of the 1940s were followed by only a gradual upward drift in leverage. Similarly, our results support the evidence in Bargeron, Denis, and Lehn (2013) that the introduction of personal and corporate taxes accompanying World War I saw little change in corporate leverage.

Estimating an optimal contracting model has given us a new perspective on the relative importance of taxes versus agency issues for capital structure. In a world where firms borrow only to fine tune an optimal capital structure, taxes are clearly a first-order consideration. However, in a world where firms borrow to finance investment and in which lenders wish to be repaid, taxes are not nearly as important. We speculate that estimating optimal contracting
models can be used as bases for deepening our understanding of a variety of corporate finance questions. One obvious candidate is executive compensation, but others include managerial incentives in a conglomerate, mergers, and banking decisions.
Appendix A

This appendix proves the propositions in the main text.

Preliminaries

To streamline the following proofs, use a simplified characterization of the problem given by (6)–(10). Let
\[ x = \{k_{j+1}, q_{j+1}(z)\}_{z \in Z} \] denote a sequence of capital stocks and state-contingent debt holdings. To simplify notation, we assume the contract starts at \( t = 0 \) so that the initial state is \( x_0 = (k_0, q_0, z_0) \). First, we show that the budget constraint (8) and participation constraint (10) are both binding.

**Lemma 1** The budget constraint (8) and the participation constraint (10) hold with strict equality at any solution of (6)-(10).

**Proof.** The proof proceeds by contradiction.

1. Budget constraint: Let \( \{d^*_j, i^*_j, p^*_j\}_{j=0}^{\infty} \) be a solution with strict inequality of (8) for some period \( \tau \). Pick \( d_\tau = d^*_\tau + \epsilon \). \( \epsilon \) is a small positive number such that \( z_\tau(k^*_\tau)^\alpha - i^*_\tau - p^*_\tau - d_\tau \geq 0 \). Thus, we can construct a feasible plan equal to \( \{d^*_j, i^*_j, p^*_j\}_{j=0}^{\infty} \) with \( d_\tau = d^*_\tau + \epsilon \), which achieves a higher contract value.

2. Participation constraint: Let \( \{d^*_j, i^*_j, p^*_j\}_{j=0}^{\infty} \) be a solution with strict inequality of (10). Pick \( \epsilon > 0 \), \( p_0 = p^*_0 - \epsilon \) such that \( p_0 + \mathbb{E}_1 \sum_{j=0}^{\infty} (\beta C(1 - \phi))^j p^*_{j+1} \geq 0 \). Let \( d_0 = d^*_0 + \epsilon \) so that the budget constraint does not change and the rest of constraints are satisfied. Thus, we can construct a feasible plan equal to \( \{d^*_j, i^*_j, p^*_j\}_{j=0}^{\infty} \) with \( p_0 < p^*_0 \) and \( d_0 = d^*_0 + \epsilon \), which is associated with a higher contract value.

The original problem of an entrant is then stated as follows:

\[
\max_x U(x) = \mathbb{E}_0 \sum_{j=0}^{\infty} \{((\beta(1 - \phi))^j(z_j k^\alpha + k_j (1 - \delta) - k_{j+1} - q_j + \beta C(1 - \phi) j+1 q_{j+1}(z_{j+1})}\} \quad (A.1)
\]
subject to
\[ z_j k_j^\alpha + k_j (1 - \delta) - k_{j+1} - q_j + \beta_C (1 - \phi) \mathbb{E}_{j+1} q_{j+1} (z_{j+1}) \geq 0 \quad \text{(A.2)} \]
\[ U(x; z^\tau) \geq D(k_j; z^\tau), \quad \forall z^\tau, \tau > 0 \quad \text{(A.3)} \]
\[ \text{given } k_0, z_0, \text{ and } q_0 = 0. \quad \text{(A.4)} \]

where \( U(x; z^\tau) \) describes the continuation value with allocation \( x \) and history \( z^\tau \) and \( D(k_j; z^\tau) \) represents the diversion value with capital stock \( k_j \) and history \( z^\tau \).

**Proof of Proposition 1**

First, we show that the solution of the original problem \( \{ k_{j+1}, \{ q_{j+1}(z) \}_{z \in \mathbb{Z}} \}_{j=0}^\infty \) is feasible in the transformed problem. As the constraints (A.2) and (A.4) are the same in the two problems, we focus on the feasibility of the collateral constraint (11). We prove feasibility by contradiction. Suppose (11) is violated for some \( \tau \): \( q_\tau > \theta k_\tau (1 - \delta) \). Let \( \{ k_{j+1}^D, \{ q_{j+1}^D(z) \}_{z \in \mathbb{Z}} \}_{j=\tau}^\infty = \{ k_{j+1}, \{ q_{j+1}(z) \}_{z \in \mathbb{Z}} \}_{j=\tau}^\infty \). Thus, dividend policies are the same except in period \( \tau \). At period \( \tau \), \( d_\tau = z_\tau k_\tau^\alpha + k_\tau (1 - \delta) - k_{\tau+1} - q_\tau + \beta_C (1 - \phi) \mathbb{E}(q_{\tau+1}) \), while \( d_{\tau}^D = z_\tau k_\tau^\alpha + k_\tau (1 - \theta)(1 - \delta) - k_{\tau+1} + \beta_C (1 - \phi) \mathbb{E}(q_{\tau+1}) \). Because \( q_\tau > \theta k_\tau (1 - \delta) \), we have \( d_{\tau}^D > d_\tau \geq 0 \). Thus, the newly constructed sequence is a feasible solution of the firm’s problem after repudiation, and \( \mathbb{E}_\tau \sum_{j=0}^\infty d_{\tau+j}^D > \mathbb{E}_\tau \sum_{j=0}^\infty d_{\tau+j} \). Then the diversion value \( \mathbb{E}_\tau \sum_{j=0}^\infty \hat{d}_{\tau+j} \geq \mathbb{E}_\tau \sum_{j=0}^\infty d_{\tau+j}^D > \mathbb{E}_\tau \sum_{j=0}^\infty d_{\tau+j} \). This inequality then violates the assumption that \( \{ k_{j+1}, \{ q_{j+1}(z) \}_{z \in \mathbb{Z}} \}_{j=0}^\infty \) is optimal in the original problem.

Next, we show that the solution of the transformed problem \( \{ \hat{k}_{j+1}, \{ \hat{q}_{j+1}(z) \}_{z \in \mathbb{Z}} \}_{j=0}^\infty \) is feasible in the original problem. Suppose the enforcement constraint at time \( \tau \) is violated: \( \mathbb{E}_\tau \sum_{j=0}^\infty \hat{d}_{\tau+j} < \mathbb{E}_\tau \sum_{j=0}^\infty d_{\tau+j} \). We construct a sequence \( \{ k_{j+1}^N, \{ q_{j+1}^N(z) \}_{z \in \mathbb{Z}} \}_{j=\tau}^\infty \) equal to the optimal policies of repudiating at \( \tau \), \( \{ \hat{k}_{j+1}, \{ \hat{q}_{j+1}(z) \}_{z \in \mathbb{Z}} \}_{j=\tau}^\infty \). Then \( d_{\tau+j}^N = \hat{d}_{\tau+j}, \forall j \geq \tau \). Let \( k_{\tau}^N = \hat{k}_{\tau} \) and \( q_{\tau}^N = \hat{q}_{\tau} \). Accordingly, \( d_{\tau}^N = z_\tau k_{\tau}^\alpha + \hat{k}_{\tau} (1 - \theta)(1 - \delta) - \hat{k}_{\tau+1} + \beta_C (1 - \phi) \mathbb{E}(\hat{q}_{\tau+1}) \).

As \( \hat{q}_{\tau} = 0 \), \( \hat{d}_{\tau} = z_\tau k_{\tau}^\alpha + \hat{k}_{\tau} (1 - \theta)(1 - \delta) - \hat{k}_{\tau+1} + \beta_C (1 - \phi) \mathbb{E}(\hat{q}_{\tau+1}) \). According to the collateral constraint, \( d_{\tau}^N \geq \hat{d}_{\tau} \). Thus, \( \mathbb{E}_\tau \sum_{j=0}^\infty d_{\tau+j}^N \geq \mathbb{E}_\tau \sum_{j=0}^\infty \hat{d}_{\tau+j}, \) which combined with
the violation of enforcement constraint, implies that $E \tau \sum_{j=0}^{\infty} d_{\tau+j}^N > E \tau \sum_{j=0}^{\infty} d_{\tau+j}$. Because the sequence $\{k_{j+1}^N, \{q_{j+1}^N(z)\}_{z \in Z}\}_{j=\tau}$ is feasible in the original problem, it is also feasible in the corresponding transformed problem. Thus, the sequence $\{k_{j+1}^N, \{q_{j+1}^N(z)\}_{z \in Z}\}_{j=\tau-1}$ is a feasible solution in the transformed problem. Then $E \tau \sum_{j=0}^{\infty} d_{\tau+j}^N > E \tau \sum_{j=0}^{\infty} d_{\tau+j}$ indicates $\{\tilde{k}_{j+1}, \{\tilde{q}_{j+1}(z)\}_{z \in Z}\}_{j=0}$ is not optimal in the transformed problem.

Finally, we show the solution of the original problem $\{k_{j+1}, \{q_{j+1}(z)\}_{z \in Z}\}_{j=0}$ is also optimal in the transformed problem. Suppose not. Then there exists a new sequence which achieves a higher value because the solution of the transformed problem is also feasible in the original problem. This new sequence then generates higher value in the original problem, which violates the condition that $\{k_{j+1}, \{q_{j+1}(z)\}_{z \in Z}\}_{j=0}$ be optimal in the original problem. Similarly, we can prove that the solution of the transformed problem is also a solution of the original problem. ■

**Proof of Proposition 2**

We apply Theorem 9.6 in Stokey, Lucas, and Prescott (1989) as Assumption 9.4-9.7 hold. The only nontrivial of these assumptions is that the constraint set is non-empty, compact-valued, and continuous. We now establish this.

Assume that $z \in Z$, which is a Borel set in $\mathbb{R}$, and that the transition function has the Feller property. For simplicity and without loss of generality, we assume $k' \in [0, \bar{k}]$ and $q'(z) \in [\underline{q}, \bar{q}]$. Here the upper bound for the capital stock, $\bar{k}$, can be derived from $\bar{k} = z h \bar{k}^\alpha + \bar{k}(1 - \delta)$ because if the capital stock is larger than $\bar{k}$, profit is negative. The upper bound for $q$ can then be derived from the collateral constraint. The lower bound for $q$ is also well defined. First, the limited liability constraint implies that $p_j$ has an upper bound $\forall j$. Second, the initial participation constraint then implies that the lower bound of $p_j$ is well-defined $\forall j$. Therefore the lower bound of $q$, $\underline{q}$, must also be well defined. Let $\bar{w}$ and $\underline{w}$ be the corresponding boundaries for net wealth. Define the domain $X = Z \times [\underline{w}, \bar{w}]$. $Y = [0, \bar{k}] \times [\underline{q}, \bar{q}]$. Define $\Gamma(w, z)$ as follows.
Lemma 2 $\Gamma : \mathcal{X} \rightarrow \mathcal{Y}$ is non-empty, compact-valued, and continuous.

Proof. Pick $(w,z) \in \mathcal{X}$.

1. $k' = 0$ and $\{q'(z')\}_{z' \in \mathbb{Z}} = 0$ belongs to $\Gamma(w,z)$. Thus, the constraint set is non-empty.

2. $\Gamma(w,z)$ is bounded as a subset of $\mathcal{Y}$. Pick $\{k_n', \{q_n'(z')\}_{z' \in \mathbb{Z}}\} \in \Gamma(w,z)$ and $\{k_n', \{q_n'(z')\}_{z' \in \mathbb{Z}}\} \rightarrow (k', \{q'(z')\}_{z' \in \mathbb{Z}})$. Because of the continuity of the constraints, $(k', \{q'(z')\}_{z' \in \mathbb{Z}}) \in \Gamma(w,z)$.

3. Lower hemi-continuity: Pick $\{k_n', \{q_n'(z')\}_{z' \in \mathbb{Z}}\} \in \Gamma(w,z)$ and $\{k_n', \{q_n'(z')\}_{z' \in \mathbb{Z}}\} \rightarrow (k', \{q'(z')\}_{z' \in \mathbb{Z}})$. $\exists M_1$ such that $|k_n' - k'| < \delta_1$, $|q_n'(z') - q'(z')| < \delta_2$, $\forall z'$, and $w - k_n' + \beta(1 - \delta)\mathbb{E}q_n'(z') > 0$. Pick $w_n \rightarrow w$ and $z_n \rightarrow z$, $\exists M_2$ such that $|w_n - w| < \delta_3$, $|z_n - z| < \delta_4$, and $w_n - k_n' + \beta(1 - \delta)\mathbb{E}(q_n'(z')|z_n) \geq 0$. Let $N = \max(M_1, M_2)$. Then $\{k_n', \{q_n'(z')\}_{z' \in \mathbb{Z}}\} \in \Gamma(w_n, z_n)$. Thus, $\forall(k', \{q'(z')\}_{z' \in \mathbb{Z}}) \in \Gamma(w,z)$, $\forall(w_n, z_n) \rightarrow (w,z)$, we can find an $N$ such that $\{k_n', \{q_n'(z')\}_{z' \in \mathbb{Z}}\}_{N} \rightarrow (k', \{q'(z')\}_{z' \in \mathbb{Z}})$, $\{k_n', \{q_n'(z')\}_{z' \in \mathbb{Z}}\} \in \Gamma(w_n, z_n)$. The upper hemi-continuity can be proved by applying Theorem 3.4 in Stokey et al. (1989). As $(w,z)$ is chosen arbitrarily, the arguments apply to all $(w,z) \in \mathcal{X}$.

Appendix B

This appendix summarizes the numerical methods used to solve the model in the main text. The basic procedure follows Proposition 2 by using value function iteration. We use the enumeration method to search for a set of feasible boundaries of net wealth $w$, the capital stock $k$, and debt $q$. In particular, the upper bound of $k$, $\bar{k}$, can be derived from the first-best
solution. By fixing the lower bound of \( q, \bar{q} \), the upper bound of \( w, \bar{w} \), can be derived using \( \bar{k} \) from the definition of the wealth. The upper bound of debt, \( \bar{q} \), can be also be derived from \( w \) and \( \bar{k} \). Thus, we only need to search for the feasible values of \( w \) and \( \bar{k} \) in addition to \( q \). The details of this computational method are as follows:

(a) We first define the sets of possible values of \( w \) and \( \bar{k} \). As \( w \) and \( k \) are both strictly positive\(^6\), we set the lower bounds of \( w \) and \( k \) to be small positive numbers. The upper bounds of \( w \) and \( k \) can be derived from the first-best solution. Then, the possible values of \( w \) and \( k \) are divided into two vectors each with 30 equal-spaced grid points. As a starting point, we let \( q = 0 \).

(b) Next we search over the two vectors for \( w \) and \( k \) in an ascending order. Give one set of boundaries, the state space of \( w \) has 128 log-spaced grid points. The control space of \( k \) has 100 equally spaced grid points. The control space of \( q(z') \) has 80 equally spaced grid points. We then re-compute the first-best case as an initial guess. The \( AR(1) \) process of the idiosyncratic shock is discretized using using the algorithm in Tauchen and Hussey (1991) with the number grid points equal to 3.

(c) We then iterate the value function for 4 times. The value function iteration method is described in detail below. During this process, if at least one of the policy functions hits the lower bound of \( k \) or the upper bound of \( q(z') \) the search process will go back to step (b) and repeat this process for the next set of boundaries. If the iteration survives after 4 times, we proceed to step (d).

(d) If steps (b) through (c) produces one feasible pair of \( w \) and \( k \) but the lower bound of \( q \) is hit by the policy function for optimal debt, \( q \) needs to be updated and the procedure goes back to step (a). Otherwise, we say the procedure find a feasible set of boundaries.

\(^6\)See Lemma 6 in Rampini and Viswanathan (2013).
If the above procedure does find a set of feasible boundaries, we then keep solving the Bellman equation using the value function iteration method and accelerating the process with McQueen Porteus bounds. At each grid point of the state space, the maximum is achieved by searching the control spaces of $k$ and $\{q(z')\}$ plus two extra sets of possible control variables. One set of control variables is associated with all the binding enforcement constraints where $k'$ belongs to the grids of the control space. The other set of control variables is associated with all the binding enforcement constraints and the binding non-negative dividend constraint. In the later case, the binding non-negative dividend constraint is transformed into one equation with one unknown $k'$. We then use bisection method to solve for the zero root. The off-grid value of $w$ is interpolated using a cubic spline with a “not-a-knot” condition as stated in Khan and Thomas (2008).

Appendix C

We now give a brief outline of the estimation procedure, which draws from Ingram and Lee (1991) Duffie and Singleton (1993), but which is adapted to our panel setting. Suppose we have $J$ variables contained in the data vector $x_{it}, i = 1, \ldots, n; t = 1, \ldots, T$. We assume that the $J \times T$ matrix $x_i$ is i.i.d., but we allow for possible dependence among the elements of $x_i$. Let $y_{itk}(b)$ be a data vector from simulation $k, i = 1, \ldots, n, t = 1, \ldots, T, \text{ and } k = 1, \ldots, K$. Here, $K$ is the number of times the model is simulated, i.e. the simulated sample size divided by the actual sample size).

The simulated data, $y_{itk}(b)$, depend on a vector of structural parameters, $b$. In our application $b \equiv (\delta, \alpha, \rho, \sigma_z, \phi, \theta, k_0, \beta_C - \beta)$. The goal is to estimate $b$ by matching a set of simulated moments, denoted as $h(y_{itk}(b))$, with the corresponding set of actual data moments, denoted as $h(x_{it})$. Our moments are listed in the text, and we denote the number of moments as $H$. Define the sample moment vector

$$g (x_{it}, b) = (nT)^{-1} \sum_{i=1}^{n} \sum_{t=1}^{T} \left[ h (x_{it}) - K^{-1} \sum_{k=1}^{K} h (y_{itk}(b)) \right].$$
The simulated moments estimator of $b$ is then defined as the solution to the minimization of

$$\hat{b} = \arg\min_b g(x,b)' \hat{W} g(x,b),$$

in which $\hat{W}$ is a positive definite matrix that converges in probability to a deterministic positive definite matrix $W$.

Our weight matrix, $\hat{W}$, differs from that given in Ingram and Lee (1991). First, we calculate it using the influence function approach in Erickson and Whited (2002). Second, it is not the optimal weight matrix, and we justify this choice as follows. First, because our model is of an individual firm, we want the influence functions to reflect within-firm variation. Because our data contain a great deal of heterogeneity, we therefore demean each of our variables at the firm level and then calculate the influence functions for each moment using the demeaned data. We then covary the influence functions (summing over both $i$ and $t$) to obtain an estimate of the covariance matrix of the moments. The estimated weight matrix, $\hat{W}$, is the inverse of this covariance matrix. Note that the weight matrix does not depend on the parameter vector, $b$.

Two details regarding this issue are important. First, neither the influence functions for the autocorrelation coefficients nor the coefficients themselves are calculated using demeaned data because we obtain them using the double-differencing estimator in Han and Phillips (2010). Thus, we remove heterogeneity by differencing rather than by demeaning. Second, although we cannot use firm-demeaned data to calculate the means in the moment vector, we do use demeaned data to calculate the influence functions for these moments. Otherwise, the influence functions for the means would reflect primarily cross sectional variation, whereas the influence functions for the rest of the moments would reflect within-firm variation. In this case, the estimation would put the least weight on the mean moments, which does not appear to be a sensible economic objective.

The above described weight matrix does achieve our goal of reflecting within-firm variation. However, it does not account for any temporal dependence in the data. We therefore calculate our standard errors using the optimal weight matrix, which is the inverse of a
clustered moment covariance matrix. We calculate the estimate of this covariance matrix, denoted $\hat{\Omega}$, as follows. Let $\phi_{it}$ be the influence function of the moment vector $g(x_{it}, b)$ for firm $i$ at time $t$. $\phi_{it}$ then has dimension $H$. Note that this influence function is of the actual moment vector $g(x_{it}, b)$, which implies that we do not use demeaned data to calculate the influence functions for the means or autocorrelation coefficients, but that we do use demeaned data to calculate the rest of the moments. The estimate of $\Omega$ is

$$\frac{1}{nT} \sum_{i=1}^{n} \left( \sum_{t=1}^{T} \phi_{it} \right) \left( \sum_{t=1}^{T} \phi_{it} \right)^{\prime}$$

Note that this estimate does not depend on $b$. Note also that if we were to use demeaned data, the elements corresponding to the mean moments would be zero.

The standard errors are then given by the usual GMM formula, adjusted for simulation error. Letting $G \equiv \partial g(x_{it}, b)/\partial b$, the asymptotic distribution of $b$ is

$$\text{avar}(\hat{b}) \equiv \left( 1 + \frac{1}{K} \right) [GWG']^{-1} [GW\Omega WG'] [GWG']^{-1}.$$
References


Ai, Hengjie, Dana Kiku, and Rui Li, 2013, A mechanism design model of firm dynamics: A mechanism design model of firm dynamics: The case of limited commitment, Manuscript, University of Minnesota.


Bargeron, Leonce, David Denis, and Kenneth Lehn, 2013, Taxes, investment, and capital structure: A study of u.s. firms in taxes, investment, and capital structure: A study of u.s. firms in the early 1900s, Manuscript, University of Pittsburgh.


Heider, Florian, and Alexander Ljungqvist, 2012, As certain as debt and taxes: Estimating the tax sensitivity of leverage from exogenous state tax changes, Manuscript, NYU.


Lorenzoni, Guido, and Karl Walentin, 2007, Financial frictions, investment and Tobin’s q, Manuscript, MIT.


of Financial Economics* 109, 466–492.

Roberts, Michael R., 2012, The role of dynamic renegotiation and asymmetric information
in financial contracting, Manuscript, University of Pennsylvania.

Schmid, Lukas, 2011, A quantitative dynamic agency model of financing constraints,
Manuscript, Duke University.

Spear, Stephen E., and Sanjay Srivastava, 1987, On repeated moral hazard with discounting,


Tauchen, George, and Robert Hussey, 1991, Quadrature-based methods for obtaining ap-

Taylor, Allison, and Alicia Sansone, 2006, *The Handbook of Loan Syndications and Trading*
(McGraw-Hill).

Taylor, Lucian A., 2010, Why are CEOs rarely fired? Evidence from structural estimation,
*Journal of Finance* 65, 2051–2087.

Thomas, Jonathan, and Tim Worrall, 1994, Foreign direct investment and the risk of expro-
Calculations are based on a sample of nonfinancial firms from the annual 2013 Compustat industrial files. The sample period is from 1965 to 2012. The estimation is done with SMM, which chooses structural model parameters by matching the moments from a simulated panel of firms to the corresponding moments from the data. The table contains estimates from two models: one without corporate taxation and one with both taxation of corporate income and an interest tax deduction. Panel A reports the simulated and actual moments and the clustered $t$-statistics for the differences between the corresponding moments. Panel B reports the estimated structural parameters, with clustered standard errors in parentheses. $\delta$ is the rate of capital depreciation. $\alpha$ is the curvature of the profit function. $\rho_z$ and $\sigma_z$ are the serial correlation and the standard deviation of the innovation to the profitability shock. $\theta$ is fraction of the capital stock lost when a contract is repudiated. $k_0$ is the initial capital stock. $\beta_C - \beta$ is the difference in discount factors between lenders and borrowers.

### A. Moments

<table>
<thead>
<tr>
<th></th>
<th>Actual</th>
<th>No Taxes</th>
<th>Taxes</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Sim.</td>
<td>T-stat.</td>
</tr>
<tr>
<td>Average debt</td>
<td>0.289</td>
<td>0.279</td>
<td>1.749</td>
</tr>
<tr>
<td>Standard deviation of debt</td>
<td>0.115</td>
<td>0.097</td>
<td>1.911</td>
</tr>
<tr>
<td>Average investment</td>
<td>0.107</td>
<td>0.109</td>
<td>-0.116</td>
</tr>
<tr>
<td>Standard deviation of investment</td>
<td>0.066</td>
<td>0.102</td>
<td>-4.071</td>
</tr>
<tr>
<td>Average profits</td>
<td>0.120</td>
<td>0.138</td>
<td>-1.772</td>
</tr>
<tr>
<td>Standard deviation of profits</td>
<td>0.088</td>
<td>0.022</td>
<td>3.244</td>
</tr>
<tr>
<td>Serial correlation profits</td>
<td>0.651</td>
<td>0.520</td>
<td>1.747</td>
</tr>
<tr>
<td>Average Tobin’s q</td>
<td>1.588</td>
<td>1.814</td>
<td>-2.226</td>
</tr>
<tr>
<td>Average distributions</td>
<td>0.026</td>
<td>0.029</td>
<td>-0.505</td>
</tr>
<tr>
<td>Standard deviation of distributions</td>
<td>0.032</td>
<td>0.018</td>
<td>1.935</td>
</tr>
</tbody>
</table>

### B. Parameter estimates

<table>
<thead>
<tr>
<th>$\delta$</th>
<th>$\alpha$</th>
<th>$\rho_z$</th>
<th>$\sigma_z$</th>
<th>$\theta$</th>
<th>$k_0/k^*$</th>
<th>$\beta_C - \beta$</th>
</tr>
</thead>
<tbody>
<tr>
<td>No Taxes</td>
<td>0.077</td>
<td>0.938</td>
<td>0.550</td>
<td>0.158</td>
<td>0.318</td>
<td>0.134</td>
</tr>
<tr>
<td></td>
<td>(0.007)</td>
<td>(0.042)</td>
<td>(0.096)</td>
<td>(0.036)</td>
<td>(0.001)</td>
<td>(0.047)</td>
</tr>
<tr>
<td>Taxes</td>
<td>0.029</td>
<td>0.892</td>
<td>0.570</td>
<td>0.215</td>
<td>0.356</td>
<td>0.236</td>
</tr>
<tr>
<td></td>
<td>(0.006)</td>
<td>(0.021)</td>
<td>(0.058)</td>
<td>(0.029)</td>
<td>(0.002)</td>
<td>(0.028)</td>
</tr>
</tbody>
</table>
Table 2: Industry Parameter Estimates

Calculations are based on a sample of nonfinancial firms from the annual 2013 COMPUSTAT industrial files. The sample period is from 1965 to 2012. The estimation is done with SMM, which chooses structural model parameters by matching the moments from a simulated panel of firms to the corresponding moments from the data. The table contains the parameter estimates from estimations done on several industries. We estimate the model with taxes. $\delta$ is the rate of capital depreciation. $\alpha$ is the curvature of the profit function. $\rho_z$ and $\sigma_z$ are the serial correlation and the standard deviation of the innovation to the profitability shock. $\theta$ is fraction of the capital stock lost when a contract is repudiated. $k_0$ is the initial capital stock, and $k^*$ is the steady-state capital stock. $\beta_C - \beta$ is the difference in discount factors between lenders and borrowers. Clustered standard errors are in parentheses.

<table>
<thead>
<tr>
<th>Industry</th>
<th>$\delta$</th>
<th>$\alpha$</th>
<th>$\rho_z$</th>
<th>$\sigma_z$</th>
<th>$\theta$</th>
<th>$k_0/k^*$</th>
<th>$\beta_C - \beta$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Metal Mining (10)</td>
<td>0.064</td>
<td>0.944</td>
<td>0.588</td>
<td>0.164</td>
<td>0.242</td>
<td>0.215</td>
<td>0.027</td>
</tr>
<tr>
<td></td>
<td>(0.046)</td>
<td>(0.152)</td>
<td>(0.322)</td>
<td>(0.179)</td>
<td>(0.007)</td>
<td>(0.222)</td>
<td>(0.135)</td>
</tr>
<tr>
<td>Oil and Gas (13)</td>
<td>0.105</td>
<td>0.949</td>
<td>0.605</td>
<td>0.157</td>
<td>0.368</td>
<td>0.240</td>
<td>0.011</td>
</tr>
<tr>
<td></td>
<td>(0.020)</td>
<td>(0.082)</td>
<td>(0.252)</td>
<td>(0.067)</td>
<td>(0.004)</td>
<td>(0.125)</td>
<td>(0.121)</td>
</tr>
<tr>
<td>Food (20)</td>
<td>0.061</td>
<td>0.811</td>
<td>0.521</td>
<td>0.100</td>
<td>0.268</td>
<td>0.359</td>
<td>0.024</td>
</tr>
<tr>
<td></td>
<td>(0.037)</td>
<td>(0.094)</td>
<td>(0.290)</td>
<td>(0.181)</td>
<td>(0.006)</td>
<td>(0.128)</td>
<td>(0.139)</td>
</tr>
<tr>
<td>Paper (26)</td>
<td>0.070</td>
<td>0.919</td>
<td>0.732</td>
<td>0.122</td>
<td>0.322</td>
<td>0.268</td>
<td>0.040</td>
</tr>
<tr>
<td></td>
<td>(0.037)</td>
<td>(0.137)</td>
<td>(0.367)</td>
<td>(0.184)</td>
<td>(0.007)</td>
<td>(0.119)</td>
<td>(0.149)</td>
</tr>
<tr>
<td>Printing (27)</td>
<td>0.057</td>
<td>0.810</td>
<td>0.531</td>
<td>0.097</td>
<td>0.263</td>
<td>0.457</td>
<td>0.022</td>
</tr>
<tr>
<td></td>
<td>(0.042)</td>
<td>(0.114)</td>
<td>(0.310)</td>
<td>(0.195)</td>
<td>(0.007)</td>
<td>(0.205)</td>
<td>(0.139)</td>
</tr>
<tr>
<td>Chemicals (28)</td>
<td>0.062</td>
<td>0.931</td>
<td>0.541</td>
<td>0.202</td>
<td>0.307</td>
<td>0.147</td>
<td>0.022</td>
</tr>
<tr>
<td></td>
<td>(0.022)</td>
<td>(0.111)</td>
<td>(0.393)</td>
<td>(0.076)</td>
<td>(0.003)</td>
<td>(0.060)</td>
<td>(0.247)</td>
</tr>
<tr>
<td>Oil Refining (29)</td>
<td>0.076</td>
<td>0.921</td>
<td>0.588</td>
<td>0.106</td>
<td>0.275</td>
<td>0.273</td>
<td>0.039</td>
</tr>
<tr>
<td></td>
<td>(0.042)</td>
<td>(0.165)</td>
<td>(0.369)</td>
<td>(0.124)</td>
<td>(0.008)</td>
<td>(0.211)</td>
<td>(0.150)</td>
</tr>
<tr>
<td>Rubber and Plastics (30)</td>
<td>0.060</td>
<td>0.911</td>
<td>0.885</td>
<td>0.107</td>
<td>0.309</td>
<td>0.494</td>
<td>0.032</td>
</tr>
<tr>
<td></td>
<td>(0.047)</td>
<td>(0.154)</td>
<td>(0.455)</td>
<td>(0.113)</td>
<td>(0.007)</td>
<td>(0.201)</td>
<td>(0.166)</td>
</tr>
<tr>
<td>Stone, Clay, Glass, Concrete (32)</td>
<td>0.061</td>
<td>0.927</td>
<td>0.888</td>
<td>0.112</td>
<td>0.350</td>
<td>0.270</td>
<td>0.032</td>
</tr>
<tr>
<td></td>
<td>(0.048)</td>
<td>(0.123)</td>
<td>(0.416)</td>
<td>(0.166)</td>
<td>(0.008)</td>
<td>(0.221)</td>
<td>(0.123)</td>
</tr>
<tr>
<td>Primary Metal (33)</td>
<td>0.034</td>
<td>0.947</td>
<td>0.737</td>
<td>0.135</td>
<td>0.312</td>
<td>0.254</td>
<td>0.022</td>
</tr>
<tr>
<td></td>
<td>(0.037)</td>
<td>(0.195)</td>
<td>(0.413)</td>
<td>(0.169)</td>
<td>(0.006)</td>
<td>(0.143)</td>
<td>(0.113)</td>
</tr>
<tr>
<td>Fabricated Metal (34)</td>
<td>0.059</td>
<td>0.890</td>
<td>0.526</td>
<td>0.104</td>
<td>0.270</td>
<td>0.311</td>
<td>0.020</td>
</tr>
<tr>
<td></td>
<td>(0.030)</td>
<td>(0.127)</td>
<td>(0.360)</td>
<td>(0.179)</td>
<td>(0.006)</td>
<td>(0.206)</td>
<td>(0.153)</td>
</tr>
<tr>
<td>Machinery (35)</td>
<td>0.064</td>
<td>0.928</td>
<td>0.560</td>
<td>0.221</td>
<td>0.287</td>
<td>0.297</td>
<td>0.031</td>
</tr>
<tr>
<td></td>
<td>(0.020)</td>
<td>(0.128)</td>
<td>(0.305)</td>
<td>(0.064)</td>
<td>(0.004)</td>
<td>(0.088)</td>
<td>(0.136)</td>
</tr>
<tr>
<td>Industry</td>
<td>$\delta$</td>
<td>$\alpha$</td>
<td>$\rho_z$</td>
<td>$\sigma_z$</td>
<td>$\theta$</td>
<td>$k_0/k^*$</td>
<td>$\beta_C - \beta$</td>
</tr>
<tr>
<td>-------------------------------</td>
<td>----------</td>
<td>----------</td>
<td>----------</td>
<td>-----------</td>
<td>----------</td>
<td>-----------</td>
<td>-------------------</td>
</tr>
<tr>
<td>Electronics (36)</td>
<td>0.062</td>
<td>0.942</td>
<td>0.571</td>
<td>0.125</td>
<td>0.259</td>
<td>0.134</td>
<td>0.011</td>
</tr>
<tr>
<td></td>
<td>(0.025)</td>
<td>(0.176)</td>
<td>(0.331)</td>
<td>(0.073)</td>
<td>(0.004)</td>
<td>(0.100)</td>
<td>(0.119)</td>
</tr>
<tr>
<td>Transportation Equipment (37)</td>
<td>0.049</td>
<td>0.914</td>
<td>0.745</td>
<td>0.112</td>
<td>0.311</td>
<td>0.413</td>
<td>0.037</td>
</tr>
<tr>
<td></td>
<td>(0.032)</td>
<td>(0.218)</td>
<td>(0.404)</td>
<td>(0.114)</td>
<td>(0.006)</td>
<td>(0.121)</td>
<td>(0.125)</td>
</tr>
<tr>
<td>Instrumentation (38)</td>
<td>0.055</td>
<td>0.922</td>
<td>0.622</td>
<td>0.110</td>
<td>0.294</td>
<td>0.110</td>
<td>0.029</td>
</tr>
<tr>
<td></td>
<td>(0.021)</td>
<td>(0.088)</td>
<td>(0.305)</td>
<td>(0.061)</td>
<td>(0.004)</td>
<td>(0.154)</td>
<td>(0.139)</td>
</tr>
<tr>
<td>Trucking (42)</td>
<td>0.110</td>
<td>0.945</td>
<td>0.849</td>
<td>0.101</td>
<td>0.410</td>
<td>0.398</td>
<td>0.039</td>
</tr>
<tr>
<td></td>
<td>(0.047)</td>
<td>(0.174)</td>
<td>(0.333)</td>
<td>(0.197)</td>
<td>(0.009)</td>
<td>(0.291)</td>
<td>(0.143)</td>
</tr>
<tr>
<td>Air Transportation (45)</td>
<td>0.056</td>
<td>0.940</td>
<td>0.623</td>
<td>0.161</td>
<td>0.573</td>
<td>0.220</td>
<td>0.011</td>
</tr>
<tr>
<td></td>
<td>(0.056)</td>
<td>(0.186)</td>
<td>(0.447)</td>
<td>(0.185)</td>
<td>(0.009)</td>
<td>(0.158)</td>
<td>(0.101)</td>
</tr>
<tr>
<td>Communications (48)</td>
<td>0.058</td>
<td>0.915</td>
<td>0.781</td>
<td>0.155</td>
<td>0.478</td>
<td>0.308</td>
<td>0.030</td>
</tr>
<tr>
<td></td>
<td>(0.028)</td>
<td>(0.136)</td>
<td>(0.318)</td>
<td>(0.160)</td>
<td>(0.005)</td>
<td>(0.126)</td>
<td>(0.148)</td>
</tr>
<tr>
<td>Wholesale Durable (50)</td>
<td>0.011</td>
<td>0.950</td>
<td>0.605</td>
<td>0.117</td>
<td>0.347</td>
<td>0.279</td>
<td>0.013</td>
</tr>
<tr>
<td></td>
<td>(0.041)</td>
<td>(0.125)</td>
<td>(0.346)</td>
<td>(0.115)</td>
<td>(0.006)</td>
<td>(0.187)</td>
<td>(0.226)</td>
</tr>
<tr>
<td>Wholesale Nondurable (51)</td>
<td>0.033</td>
<td>0.915</td>
<td>0.699</td>
<td>0.106</td>
<td>0.315</td>
<td>0.297</td>
<td>0.034</td>
</tr>
<tr>
<td></td>
<td>(0.053)</td>
<td>(0.183)</td>
<td>(0.426)</td>
<td>(0.260)</td>
<td>(0.009)</td>
<td>(0.223)</td>
<td>(0.126)</td>
</tr>
<tr>
<td>Restaurants (58)</td>
<td>0.060</td>
<td>0.934</td>
<td>0.601</td>
<td>0.117</td>
<td>0.556</td>
<td>0.146</td>
<td>0.019</td>
</tr>
<tr>
<td></td>
<td>(0.058)</td>
<td>(0.185)</td>
<td>(0.442)</td>
<td>(0.250)</td>
<td>(0.009)</td>
<td>(0.219)</td>
<td>(0.146)</td>
</tr>
<tr>
<td>Business Services (73)</td>
<td>0.044</td>
<td>0.924</td>
<td>0.606</td>
<td>0.118</td>
<td>0.272</td>
<td>0.132</td>
<td>0.037</td>
</tr>
<tr>
<td></td>
<td>(0.028)</td>
<td>(0.073)</td>
<td>(0.203)</td>
<td>(0.059)</td>
<td>(0.003)</td>
<td>(0.064)</td>
<td>(0.128)</td>
</tr>
<tr>
<td>Health Services (80)</td>
<td>0.025</td>
<td>0.926</td>
<td>0.542</td>
<td>0.127</td>
<td>0.417</td>
<td>0.251</td>
<td>0.024</td>
</tr>
<tr>
<td></td>
<td>(0.041)</td>
<td>(0.187)</td>
<td>(0.229)</td>
<td>(0.183)</td>
<td>(0.008)</td>
<td>(0.180)</td>
<td>(0.125)</td>
</tr>
<tr>
<td>Engineering, Accounting, Research (87)</td>
<td>0.031</td>
<td>0.937</td>
<td>0.600</td>
<td>0.110</td>
<td>0.305</td>
<td>0.087</td>
<td>0.012</td>
</tr>
<tr>
<td></td>
<td>(0.042)</td>
<td>(0.153)</td>
<td>(0.350)</td>
<td>(0.175)</td>
<td>(0.008)</td>
<td>(0.208)</td>
<td>(0.117)</td>
</tr>
</tbody>
</table>
Table 3: Difference-in-Difference Simulated Moments Estimation

Calculations are based on a sample of nonfinancial firms from the annual 2013 Compustat industrial files. The sample period is from 1965 to 2012. The estimation is done with SMM, which chooses structural model parameters by matching the moments from a simulated panel of firms to the corresponding moments from the data. Panel A reports the simulated and actual moments and the clustered t-statistics for the differences between the corresponding moments. Panel B reports the estimated structural parameters, with clustered standard errors in parentheses. $\delta$ is the rate of capital depreciation. $\alpha$ is the curvature of the profit function. $\rho_z$ and $\sigma_z$ are the serial correlation and the standard deviation of the innovation to the profitability shock. $\theta$ is fraction of the capital stock lost when a contract is repudiated. $k_0$ is the initial capital stock. $\beta_C - \beta$ is the difference in discount factors between lenders and borrowers. After is defined as after the passage of an anti-recharacterization law but before 2003. Treated is defined as incorporation in Texas, Louisiana, or Alabama.

### A. Moments

<table>
<thead>
<tr>
<th></th>
<th>Control Before</th>
<th>Control After</th>
<th>Treated Before</th>
<th>Treated After</th>
</tr>
</thead>
<tbody>
<tr>
<td>Average debt</td>
<td>0.291</td>
<td>0.293</td>
<td>-0.696</td>
<td>0.298</td>
</tr>
<tr>
<td>Standard deviation of debt</td>
<td>0.119</td>
<td>0.109</td>
<td>1.719</td>
<td>0.082</td>
</tr>
<tr>
<td>Average investment</td>
<td>0.106</td>
<td>0.120</td>
<td>-1.120</td>
<td>0.097</td>
</tr>
<tr>
<td>Standard deviation of investment</td>
<td>0.065</td>
<td>0.093</td>
<td>-3.710</td>
<td>0.047</td>
</tr>
<tr>
<td>Average profits</td>
<td>0.117</td>
<td>0.137</td>
<td>-2.830</td>
<td>0.096</td>
</tr>
<tr>
<td>Standard deviation of profits</td>
<td>0.090</td>
<td>0.018</td>
<td>3.766</td>
<td>0.072</td>
</tr>
<tr>
<td>Serial correlation profits</td>
<td>0.666</td>
<td>0.590</td>
<td>1.993</td>
<td>0.576</td>
</tr>
<tr>
<td>Average Tobin’s $q$</td>
<td>1.645</td>
<td>1.642</td>
<td>0.792</td>
<td>1.776</td>
</tr>
<tr>
<td>Average distributions</td>
<td>0.027</td>
<td>0.009</td>
<td>3.433</td>
<td>0.026</td>
</tr>
<tr>
<td>Standard deviation of distributions</td>
<td>0.033</td>
<td>0.009</td>
<td>2.159</td>
<td>0.030</td>
</tr>
</tbody>
</table>

### B. Parameter estimates

<table>
<thead>
<tr>
<th></th>
<th>$\delta$</th>
<th>$\alpha$</th>
<th>$\rho_z$</th>
<th>$\sigma_z$</th>
<th>$\theta$</th>
<th>$k_0/k^*$</th>
<th>$\beta_C - \beta$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Control Before</td>
<td>0.060</td>
<td>0.948</td>
<td>0.635</td>
<td>0.126</td>
<td>0.340</td>
<td>0.226</td>
<td>0.026</td>
</tr>
<tr>
<td></td>
<td>(0.008)</td>
<td>(0.056)</td>
<td>(0.176)</td>
<td>(0.039)</td>
<td>(0.003)</td>
<td>(0.146)</td>
<td>(0.021)</td>
</tr>
<tr>
<td>Control After</td>
<td>0.040</td>
<td>0.901</td>
<td>0.585</td>
<td>0.119</td>
<td>0.325</td>
<td>0.147</td>
<td>0.037</td>
</tr>
<tr>
<td></td>
<td>(0.044)</td>
<td>(0.337)</td>
<td>(0.644)</td>
<td>(0.105)</td>
<td>(0.006)</td>
<td>(0.517)</td>
<td>(0.098)</td>
</tr>
<tr>
<td>Treated Before</td>
<td>0.080</td>
<td>0.956</td>
<td>0.876</td>
<td>0.100</td>
<td>0.374</td>
<td>0.241</td>
<td>0.015</td>
</tr>
<tr>
<td></td>
<td>(0.042)</td>
<td>(0.433)</td>
<td>(1.072)</td>
<td>(0.204)</td>
<td>(0.009)</td>
<td>(0.938)</td>
<td>(0.157)</td>
</tr>
<tr>
<td>Treated After</td>
<td>0.048</td>
<td>0.939</td>
<td>0.703</td>
<td>0.152</td>
<td>0.430</td>
<td>0.424</td>
<td>0.027</td>
</tr>
<tr>
<td></td>
<td>(0.142)</td>
<td>(0.944)</td>
<td>(2.981)</td>
<td>(0.027)</td>
<td>(0.023)</td>
<td>(2.222)</td>
<td>(0.258)</td>
</tr>
</tbody>
</table>
This figure depicts the policy functions for the model of Section 2, with the corporate tax rate set to zero and with all other parameters from the estimation of the no-tax model on the full sample. All of the variables are scaled by the unconstrained steady-state capital stock. The horizontal axis contains net wealth. On the vertical axis are the capital stock, dividends, and debt contingent on a low future state, a medium future state, and a high future state.
This figure depicts the policy functions for the model of Section 2, with the corporate tax rate set to 0.2 and with all other parameters from the estimation of the no-tax model on the full sample. We use the no-tax model parameters for comparability with Figure 1. All of the variables are scaled by the unconstrained steady-state capital stock. The horizontal axis contains net wealth. On the vertical axis are the capital stock, dividends, and debt contingent on a low future state, a medium future state, and a high future state.
Calculations are based on a sample of nonfinancial firms from the annual 2013 Compustat industrial files. The sample period is from 1965 to 2012. The sample is split into 24 industries, listed in Table 2. The estimation is done with SMM, which chooses structural model parameters by matching the moments from a simulated panel of firms to the corresponding moments from the data. This figure plots data averages versus simulated averages for four variables: leverage, the rate of investment, Tobin’s $Q$, and dividends.
Calculations are based on a sample of nonfinancial firms from the annual 2013 Compustat industrial files. The sample period is from 1965 to 2012. The sample is split into 24 industries, listed in Table 2. The estimation is done with SMM, which chooses structural model parameters by matching the moments from a simulated panel of firms to the corresponding moments from the data. This figure plots the estimated level of the collateral constraint, $\theta(1 - \delta)$ versus average leverage for each industry.
This figure is constructed as follows. We pick a grid for $\theta$, the fraction of the capital stock lost to the firm when it repudiates a contract, $\tau_c$, the corporate tax rate, and $\beta_C - \beta$, the difference in the discount factors between borrows and lenders. We then solve the model for each of the different parameter values, simulate the model for 50 time periods, and then plot average leverage. We use a parameterization from the estimates of the no-tax model from Table 1. “Young” firms are those in the first 10 periods of life, and “old” firms are those in periods 11 to 50.
Figure 6: Taxes and leverage in a model with $\beta_C = \beta$

This figure is constructed as follows. We pick a grid for $\tau_c$, the corporate tax rate. We then solve the model for each of the different parameter values, simulate the model for 50 time periods, and then plot average leverage. We use a parameterization from the estimates of the no-tax model from Table 1. “Young” firms are those in the first 10 periods of life, and “old” firms are those in periods 11 to 50.