# Endogenous Financial Constraints, Taxes, and Leverage

Shaojin Li and Toni M. Whited*

October 20, 2013

*Li is from the Shanghai University of Economics and Finance, School of Finance, Shanghai, 200433 P.R. China. 86-21-65908391. li.shaojin@shufe.edu.cn. Whited is from the Simon School of Business, University of Rochester, Rochester, NY 14627. (585)275-3916. toni.whited@simon.rochester.edu. We thank Adriano Rampini and seminar participants at Cass Business School and Wirtschaftsuniversität Wien for helpful comments and suggestions.

**Endogenous Financial Constraints, Taxes, and Leverage**


**Abstract**

We quantify the relative importance of contracting frictions and taxes in shaping firms' capital structures. We estimate a dynamic contracting model based on limited entrepreneurial commitment. In the model 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 that financial constraints arise endogenously. We estimate the model's parameters using data from firms in several industries. We find that taxes have no effect on optimal leverage because they affect neither the lender's discount factor nor the constraints that allow the contract to be self-enforcing. In contrast, taxes have a sharp deleterious effect on investment.

# 1 Introduction

It is safe to say that credit supply affects firms' financial and real economic decisions. Hundreds of studies have demonstrated that the cross-sectional and time-series patterns in investment, employment, and finance are nearly impossible to reconcile with a world in which external finance is easy to obtain. Yet much of the empirical and structural research in corporate finance is based on models in which external financial frictions are 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.

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 little effect on optimal firm leverage. In contrast, taxation has a sharp negative effect on investment. The intuition for these results requires a brief description of the model.

In the model, a firm with a possibly infinite lifespan enters the market with a valuable investment opportunity that generates taxable revenues. However, the firm has insufficient funds to start the project 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. 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. 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, we find almost no effect of taxes on leverage. Taxes naturally affect the firm's budget constraint and the contract enforcement constraint, which stipulates that the firm must be better off repaying the debt than repudiating the contract and starting over. However changing *firm* taxes does not affect the rate at which the *lender* discounts the payments it receives from the firm. Quantitatively, we find that the main effect of taxes operates via the enforcement constraint, because this constraint nearly always binds. 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 renders debt less onerous for the firm but it also makes capital less valuable. The end result is that the taxation has little net effect on this basic tradeoff between debt and capital that makes the contract self-enforcing.

We stress that these results are quantitative in nature and come from a model whose parameters we estimate via simulated method of moments. The data therefore put tight restrictions on our model parameters and on the predictions that emerge from the model.

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 can survive default and that the firm can use to start over after contract repudiation. 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 both large and small firms, as well as 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. Instruments or quasi-experimental settings for these issues are scarce. In addition, the main sources of the contracting frictions are unobservable, and proxies for these frictions are unavailable. Quantifying their effects therefore requires a model, which puts enough structure on the data to identify the effects of interest.

Our paper fits into several literatures. The first is a set of theoretical papers that uses limited commitment models 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). 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 largely theoretical.

The second literature is the structural estimation of dynamic models in corporate finance, such as Hennessy and Whited (2005, 2007), DeAngelo, DeAngelo, and Whited (2011), and Morellec, Nikolov, and Schürhoff (2012). Our paper departs from these predecessors in one important dimension. Instead of specifying financial constraints or agency concerns as an exogenous parameter, we derive financial constraints from an optimal contracting framework, and then estimate the magnitudes of the underlying frictions. Expressing finance constraints exogenously has the advantage that one can model a richer set of financial decisions. Unfortunately, this approach has the drawback that it is hard to assess the effects of these constraints because the parameters that describe financial constraints are not likely to be invariant to firms' actions or characteristics. In contrast, estimating an actual contracting model allows one to examine counterfactual questions with respect to deeper parameters.

Our paper is most closely related to Nikolov and Schmid (2012), which also estimates a dynamic contracting model. However, our paper examines a different question. While we focus primarily on capital structure, they concentrate on investment. In addition, they work with a different class of contracting models based on moral hazard (e.g. DeMarzo and Sannikov 2006; DeMarzo, Fishman, He, and Wang 2012; Biais, Mariotti, Rochet, and Villeneuve 2010). Although ideal for understanding the effects of agency on investment, these models are less amenable to the empirical study of capital structure because the implementation of the contract in terms of capital structures is not unique. Therefore, it is impossible to know which implementation corresponds to the actual capital structures we observe in the data.

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 $\mathbb{Z}$ 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

$$k_{t+1} = (1-\delta)k_t + i_t, \tag{1}$$

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 $\beta_C > \beta$.

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. 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 1 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.

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

$$z_{t+j} k_{t+j}^{\alpha} - i_{t+j}(z^{t+j}) \geq d_{t+j}(z^{t+j}) + p_{t+j}(z^{t+j}) \tag{2}$$

$$d_{t+j}(z^{t+j}) \geq 0 \tag{3}$$

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.[1] This feature of the model captures Chapter 11 renegotiation, rather than Chapter 7 liquidation.

Let $\mathbb{E}(\cdot)$ 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:

$$D(k_\tau, z_\tau) = \mathbb{E}_\tau \sum_{j=0}^{\infty} (\beta(1-\phi))^j \hat{d}_{\tau+j}(z_{\tau+j}), \tag{4}$$

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 $(1-\theta)k_\tau$.

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

$$D(k_\tau, z_\tau) \leq \mathbb{E}_\tau \sum_{j=0}^{\infty} (\beta(1-\phi))^j d_{\tau+j}. \tag{5}$$

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:

[1] The form of diversion is adopted from Lorenzoni and Walentin (2007).

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

subject to

$$d_{t+j} \geq 0 \tag{7}$$

$$z_{t+j} k_{t+j}^{\alpha} - i_{t+j} - p_{t+j} - d_{t+j} \geq 0 \tag{8}$$

$$\mathbb{E}_\tau \sum_{j=0}^{\infty} (\beta(1-\phi))^j d_{\tau+j} \geq D(k_\tau, z_\tau), \quad \forall \tau > t \tag{9}$$

$$\mathbb{E}_t \sum_{j=0}^{\infty} (\beta_C(1-\phi))^j p_{t+j} \geq 0 \tag{10}$$

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]

## 2.4 Recursive formulation

This section reformulates the contracting problem recursively, with the solution characterized as the Pareto frontier between the value entitlements of the firm and the lender, given by (6) and (10), respectively. 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 promised debt at time $\tau$. Following Spear and Srivastava (1987) and Abreu, Pearce, and Stacchetti (1990), the contractual problem can be stated recursively using $q_\tau$ as a state variable. The argument proceeds as follows. We first propose a recursive formulation. Then we show that this formulation

[2] It is possible to assume that the lender seizes some of the firm's assets in default. However, because the lender commits to the contract, its participation constraint never binds, so we omit this possibility for simplicity.

can be simplified. Finally, we prove that the simplified recursive formulation is equivalent to the original problem.

Let a prime denote a variable in the subsequent period, and let no prime denote a variable in the current period, and define $D(k', z') \equiv V((1-\theta)k', 0, z')$. We propose that the Bellman equation for the problem can be expressed as:

$$V(k,q,z) = \max_{k',q(z')} \quad zk^\alpha + k(1-\delta) - q - k' + \beta_C(1-\phi)\mathbb{E}q(z') + \beta(1-\phi)\mathbb{E}V(k', q'(z'), z') \quad (11)$$

subject to

$$zk^\alpha + k(1-\delta) - q - k' + \beta_C(1-\phi)\mathbb{E}q(z') \geq 0 \quad (12)$$

$$V(k', q'(z'), z') \geq D(k', z'), \quad \forall z' \in \mathbb{Z}, \quad (13)$$

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 (12) and (13), the solution to the unconstrained optimization (11) 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 (12) only to the extent that they define net wealth. Thus, the recursive problem can be rewritten as follows:

**Proposition 1**

$$V(w,z) = \max_{k',q(z')} \quad w - k' + \beta_C(1-\phi)\mathbb{E}q(z') + \beta(1-\phi)\mathbb{E}V(w'(z'), z') \quad (14)$$

*subject to*

$$w - k' + \beta_C(1-\phi)\mathbb{E}q(z') \geq 0 \quad (15)$$

$$V(w'(z'), z') \geq D(k', z'), \quad \forall z' \in \mathbb{Z}, \quad (16)$$

*where* $D(k', z') = V(\hat{w}(z'), z')$ *and* $\hat{w}(z') = z'(k'(1-\theta))^\alpha + k'(1-\theta)(1-\delta)$.

Solving this problem is difficult for two reasons. First, for some sets of model parameters, the constraint set is empty, so that the model has no solution. We therefore we assume that under some parameterizations, the constraint set is not empty. This assumption is innocuous given that we can obtain numerical solutions for many parameterizations. Second, in contrast to a standard dynamic programming problem, the value function appears in the constraints. To deal with this second issue, we note, as in Rampini and Viswanathan (2013), that it is possible to rewrite the constraint (16) in terms of collateral as follows:

**Proposition 2** $V(w'(z'), z') \geq D(k', z'), \quad \forall z' \in \mathbb{Z}$ *if and only if*

$$z'k'^{\alpha} - z'(k'(1-\theta))^{\alpha} + \theta k'(1-\delta) \geq q'(z'), \quad \forall z' \in \mathbb{Z} \tag{17}$$

Although seemingly restrictive, 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.

When we use the collateral form of the enforcement constraint, the Bellman equation can be solved by iteration with an initial guess of a function that is strictly increasing in $w$. We use the Pareto frontier of the first-best problem as the initial guess. In our case, the first best problem is for the shareholders to maximize the total gain from the project in the absence of enforcement problems. For this solution, we let $\beta_C = \beta$. The Pareto frontier is then defined as follows:

$$W(w, z) = \max_{k', q'} \quad w - k' + \beta(1-\phi)q' + \beta(1-\phi)\mathbb{E}W(w', z'), \tag{18}$$

where $w = zk^{\alpha} + k(1-\delta) - q$. Note that here $w'$ does not depend on the future state $z'$, whereas $w'$ does depend on $z'$ in the constrained optimization problem. The reason is that in the constrained problem, $w'$ depends on the state contingent payments to the lender, whereas in the first-best problem, any such payments are undefined.

Define the mapping $T$ in the space of bounded functions.

$$T(V)(w,z) = \max_{k',q(z')} \quad w - k' + \beta_C(1-\phi)\mathbb{E}q(z') + \beta(1-\phi)\mathbb{E}V(w'(z'),z') \tag{19}$$

subject to (15) and (16).

The following proposition establishes the existence of a solution.

**Proposition 3** *$T^n(W)$ converges pointwise to $V$ as $n \to \infty$.*

This proposition is useful because it implies that the solution to the model can then be obtained by iterating on (A.15) with $W(w,z)$ as an initial guess. We describe our numerical solution procedure in Appendix B.

## 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 (15), and let $\beta(1-\phi)\pi(z'|z)\lambda_{z'}$ be the Lagrange multiplier associated with the enforcement constraint (17) at state $z'$, where $\pi(z'|z)$ is the transitional probability from state $z$ to state $z'$. The first order condition for $k'$ is:

$$\begin{aligned} & \beta(1-\phi)\sum_{z'}\pi(z'|z)V_w(w',z')\frac{\partial w'}{\partial k'} \\ + \quad & \beta(1-\phi)\sum_{z'}\pi(z'|z)\lambda_{z'}\left(\alpha z\left(k'^{\alpha-1} - (1-\theta)(k'(1-\theta))^{\alpha-1}\right) + \theta(1-\delta)\right) - \mu = 1. \end{aligned} \tag{20}$$

where $\frac{\partial w'}{\partial k'} = z'\alpha k'^{\alpha-1} + 1 - \delta$. The first term in (20) 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 (20) 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 value of payments to the lender. The first order condition for $q(z')$ for any given value of $z'$ is

$$1 + \mu + \frac{\beta}{\beta_C}(1 + \lambda_{z'})V_w(w', z')\frac{\partial w'}{\partial q'} = 0, \quad \forall z' \in \mathbb{Z} \tag{21}$$

Using the envelope theorem and the condition $\frac{\partial w'}{\partial q'} = -1$, we can rewrite (21) as

$$\frac{1 + \mu}{1 + \mu(w', z')} = \frac{\beta}{\beta_C}(1 + \lambda_{z'}) \quad \forall z' \in \mathbb{Z} \tag{22}$$

First, note that even 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 will always be constrained because they always want to borrow more. Conversely, (22) shows that if $\beta = \beta_C$, the firm does not pay out dividends when the enforcement constraint binds, that is, when the firm is financially constrained. Interestingly, in the case in which $\beta = \beta_C$, the firm can eventually become completely unconstrained as long as $z$ has finite support. To see this result, note that $1 + \mu(w', z') = V_w(w', z')$ is a decreasing function of $w'$, and that $w'$ is an increasing function of $w$. Thus, there exists a cut-off value of $w'$ (dependent on $z'$) such that $\mu(w', z') = 0$. If $z$ has finite support, we can find a value of $w'$ such that $\mu(w', z') = 0$ for all $z'$.

## 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. Thus, to split the payment into these two parts, we define the face value of debt, $f \equiv \beta_C p$. In this case, the interest component can be expressed as:

$$\beta_C p \left( \frac{1}{\beta_C} - 1 \right).$$

With the interest deduction, from the point of view of the firm, the net-of-deduction payment to the intermediary becomes:

$$\beta_C p + \beta_C p \left( \frac{1}{\beta_C} - 1 \right) (1 - \tau_c) = p(1 - \tau_c(1 - \beta_C))$$

We assume that the intermediary continues to receive $p$, so the participation constraint and the recursion that defines $q$ remain unaffected.[3] This assumption is innocuous inasmuch as it does not affect the form of the first-order condition for optimal debt or the model's optimal policy function because the lender's participation constraint never binds.

Now we can rewrite the recursive problem as:

$$\begin{aligned} V(k, q, z) \quad &= \quad \max_{k', q(z')} \Big( (1 - \tau_c) z k^\alpha + k(1 - \delta) - k' - (1 - \tau_c(1 - \beta_C))(q - \beta_C(1 - \phi)\mathbb{E}q(z')) \Big) \\ &+ \quad \beta(1 - \phi)\mathbb{E}V(k', q'(z'), z') \end{aligned} \tag{23}$$

subject to

$$(1 - \tau_c) z k^\alpha + k(1 - \delta) - k' - (1 - \tau_c(1 - \beta_C))(q - \beta_C(1 - \phi)\mathbb{E}q(z')) \geq 0 \tag{24}$$

$$V(k', q'(z'), z') \geq V(k'(1 - \theta), 0, z') \equiv D(k', z'), \quad \forall z' \in \mathbb{Z}, \tag{25}$$

In the model with taxes, we define net wealth as $w \equiv (1 - \tau_c) z k^\alpha + k(1 - \delta) - (1 - \tau_c(1 - \beta_C))q$. As before, the model with taxes can be written in terms of net wealth, and the proofs of Propositions 1–3 proceed with only minor modification.

Of particular interest is the form that the collateral constraint (17) takes when there is taxation:

$$(1 - \tau_c) \left( z' k'^\alpha - z'(k'(1 - \theta))^\alpha \right) + \theta k'(1 - \delta) \geq (1 - \tau_c(1 - \beta_C)) q'(z'), \quad \forall z' \in \mathbb{Z} \tag{26}$$

[3] Jermann and Quadrini (2012) is the only other dynamic contracting model that contains a tax benefit, which they model as in Hennessy and Whited (2005).

Equation (26) shows that the effect of taxes on the collateral/enforcement constraint is ambiguous. The presence of taxation lowers both the left and right hand sides of (26). 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 renders debt less onerous for the firm but it also makes capital less valuable, so the punishment in default is less severe.

The sign of the net effect depends on the relative magnitudes of $\beta_C$, and $\alpha$. For example, as the lender becomes more patient (as $\beta_C$ approaches 1), the effect of the tax deductibility of interest falls because the interest component of the payments to the lender shrinks. Conversely, as the curvature of the production function, $\alpha$, approaches 1, the effect of taxes on the value of capital approaches $(1 - \tau_c)\theta z' k'$, and as $\alpha$ shrinks to zero, the effect of taxes on the value of capital approaches zero. The relative magnitude of these effects is a quantitative question, which requires estimating the model.

One final point of interest in the case with taxation only affects the first-order condition for optimal debt (22) only indirectly because the term $(1 - \tau_c(1 - \beta_C))$ cancels out on both sides of (22). However, taxes do have a material effect on the first order condition, but only inasmuch as they determine the extent to which the enforcement and limited liability constraints bind. This feature of the model stems from our definition of the face value of debt, but it also makes sense in that a lender is not going to change the rate at which it discounts the payments it receives if the firm's tax rate changes.[4] It is also useful in that it separates the revenue consequences of using debt from the discount rate consequences, which are bound together in theories of capital structure based on neoclassical investment models (e.g Hennessy and Whited 2005, 2007).

To understand this latter point, it is useful to consider the case of collateralized, riskless debt. In this case debt in a neoclassical investment model can be thought of as a constant-returns storage technology that earns the after-tax rate risk-free rate. Thus, whenever the

[4] In a model in which taxation affected both lenders and firms, the ratio of the discount factors would be unaffected, and so the first-order condition for optimal debt would remain unchanged.

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, and this deviation results in large effects on optimal debt holdings. We refer to this effect as a discount rate effect.

## 2.7 Policy Functions

We now extend the intuition behind the role of taxes in the model by examining the policy functions from two versions of the model: one with and another without taxes. Figure 1 contains the policy functions for the model, where we set the tax rate to zero and use the parameterization from the full sample estimation in Table 1 below, in which the tax rate is also set to zero. The top panel plots several optimal policies as a function of current net wealth for the case of a high 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 $\mathbb{E}(\alpha z'(k')^{\alpha-1})$ with the user cost, which is given by $\delta + 1/\beta - 1$. Figure 1 is drawn for the median value of the current state.

Three patterns stand out. First, when the firm has low net wealth, the enforcement constraint always binds if debt is contingent on the low state, as seen in the proportionality of debt to capital. 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 high levels of net worth, the binding constraint comes from the lender having a higher discount factor than the firm. However, the enforcement constraint does not always bind when debt is contingent on the medium or high shock. In this case, the firm knows it will have good investment opportunities and will preserve debt capacity accordingly. Second, if the firm has relatively low net wealth, the optimal choice of tomorrow's capital stock is increasing in today's net wealth. This pattern makes sense 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 for the high state falls below the user cost. Third, 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 and external resources to fund optimal capital expenditures.

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. Here, we scale all variables by the capital stock that sets the expected *after-tax* capital stock equal to the user cost. Strikingly, Figures 1 and 2 appear almost identical. Of course, taxation reduces the optimal steady-state capital stock by just one-third, so the firm operates on a much smaller scale. Except for this difference in scale, the policies for a taxed firm are qualitatively identical to those for an untaxed firm.

# 3 Data

Our data are from the 2012 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. As a result of these selection criteria, we obtain a panel data set with 100,149 observations for the time period between 1964 and 2011 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), cash and equivalents as CHE, operating income as OIBDP, equity repurchases as PRSTK, dividends as the sum of common and preferred dividends (DVC + DVP), debt as (DLTT + DLC), depreciation as DP, current assets as AC, and Tobin's $q$ as the ratio of (AT + PRCC_F $\times$ CSHO $-$ TXDB $-$ CEQ ) to AT. Investment, net leverage (DLTT+DLC $-$ CHE), total payout (dividends plus repurchases), and operating profit are expressed as fractions of total assets.

# 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. We then estimate the following parameters using simulated method of moments: the depreciation rate, $\delta$; the production function curvature, $\alpha$; the probability of exit, $\phi$, 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

$$\ln z_{t+1} = \rho \ln z_t + \varepsilon_{t+1}. \tag{27}$$

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.

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 $V(w,z)/k$.

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. Global identification of a simulated moments estimator obtains when the expected value of the difference between the simulated moments and the data moments equal zero if and only if the structural parameters equal their true values. A sufficient condition for identification is a one-to-one mapping between the structural parameters and a subset of the data moments of the same dimension. Because our model does not yield such a closed-form mapping, we take care to 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 the means and variances of all of the policy variables in our model. We also employ several serial correlation moments.

We now describe and rationalize the 12 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 asset pricing models, which cannot reconcile the observed high volatility of asset prices. Finally, we also include the serial correlation of operating profits and two correlations that link the real and financial sides of the firm: the correlation between leverage and investment and the correlation between leverage and 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 the parameter $\phi$ comes mainly from the mean of Tobin's $q$, which falls as $\phi$ rises because the discount rate used to determine equity value rises. 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 for 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, these effects occur largely at the transition from a zero difference to a positive difference. In addition, few of the other moments are affected by this parameter, with the exception of some of the serial correlation parameters. 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 over half of the simulated moments are statistically significantly different from their real-data counterparts, but only a few are *economically* different. This good fit is remarkable, given that the model has few parameters relative to the number of moments we match. Both versions of the model do a good job of matching the means of net leverage and investment. However, the model with taxes does slightly better at matching average profits, Tobin's $q$, and distributions. Both models struggle more with standard deviations. Although we find an insignificant difference between the actual and simulated standard deviations of investment and dividends, both models underestimate the standard deviation of operating income, and especially of leverage. 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.

This discussion reveals our second main result, which 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 cases. The intuition is that while taxes have important effects on profits, they do nothing to the relative difference in discount rates between lenders and firms, which determines whether the enforcement constraint binds for mature firms. Taxes do have the potential to affect the enforcement constraint, so we use our results to assess these affects quantitatively. We calculate the steady state capital stock, given our parameter estimates, and plug this result, along with the relevant parameters into the enforcement constraint: (26). We find that the addition of taxes actually *tightens* the collateral constraint, but only by approximately 1%, which is a tiny amount.

Panel B shows that our estimates of fundamental contracting frictions are significantly different from zero. We estimate that the average firm loses about 15% of its assets as a deadweight cost when it repudiates a contract. These estimates are in line with the direct estimates in Andrade and Kaplan (1998). We also find that the starting capital stock in both models is quite low, with the firm starting life with less than 5% of its eventual steady state capital stock. Interestingly, although we find positive estimates for the difference between the lenders' and firms' discount rates, the estimates are not significant in either model. 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 other three technological parameters, $\alpha$, $\rho$, and $\sigma_z$, are largely in line with the estimates from this previous work.

## 5.1 Large versus Small Firms

We now turn to the issue of cross-sectional heterogeneity, which is important because SMM estimates the parameters of an average firm, and the concept of an average firm is ill defined in a large heterogeneous sample. In addition, it is useful to determine whether the model is powerful enough to match moments from different subsamples of firms. Along this line, we first split the sample by firm size, with the cutoff point being the sample median of real book assets, which are calculated using the producer price index, with 2002 as the base year. The results are in Table 2, which is organized the same way as Table 1.

Panel A shows that for both groups, the model is able to reconcile several of the means. In particular, it matches both the high net leverage of the old firms and the much lower net leverage of the young firms. As in the case of the full sample, the model struggles more with second moments. The parameter estimates in Panel B are also of interest, with the smaller firms having a higher estimate for the variance of the $z$ shock, but a slightly lower shock serial correlation. More striking are the estimates of the diversion parameter, $\theta$. For the small firms, this parameter lies close to their actual average leverage, which implies that the firm is usually hugging the collateral constraint tightly. In contrast, for the large firms, the estimate of $\theta$ is 0.37, which is substantially higher than the large firms' average leverage of 0.27. This result implies that the average large firm is conserving debt capacity in the highest states of the world. The reason for this important difference between the large and small firms is not sharp differences in the shock processes, but the difference in the parameter, $\phi$. As discussed above, this parameter is identified by the mean of Tobin's $q$. However, its level has another important implication for the model because a higher value of $\phi$ implies that simulated firms "turn over" more within an industry. Thus, the industry contains more young firms, all of which face a binding collateral constraint.

## 5.2 Industry Estimation

To take a finer cut at the data, we also estimate the model on the six two-digit industries for which we have the most data: Oil and Gas (13), Food (20), Chemicals (28), Machinery (35), Electronics (36), and Instrumentation (38). The results are in Figure 3, which plots the moments, and Table 3, which reports the parameter estimates. In Figure 3 we see that the model can match the wide variation across industries in net leverage, from Oil and Gas with with net leverage of 0.23 to Instrumentation with net leverage of 0.04. Interestingly, the low-leverage industries have the lowest asset tangibility, as measured by the ratio of net property, plant, and equipment to total asset, and the highest leverage industry, Oil and Gas, has the highest measure of asset tangibility. 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.[5] Accompanying the high asset tangibility of the oil and gas industry is extremely high capital expenditures, which our model comes close to matching. Indeed, the model does an excellent job of matching the wide variation in investment, payout, Tobin's $q$ across industries.

Table 3 presents the parameter estimates accompanying Figure 3. One result is of particular interest here. The diversion parameter, $\theta$, lies near average leverage for all industries except for Oil and Gas and Electronics, in which $\theta$ exceeds average leverage by 40–50%. The clear reason is the higher level of of uncertainty in these industries, which gives the firms incentives to preserve debt capacity.

[5] See also the models in Rampini and Viswanathan (2010, 2013), as well as the empirical evidence in Erickson, Jiang, and Whited (2013).

# 6 Counterfactuals

We now examine what would happen to optimal financing and investment 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 taxes. We then investigate results of changing two key parameters: $\theta$ and $\tau_c$, which are the proportion of the capital stock lost to the firm if it repudiates a contract and the corporate tax rate.

The results from these exercises are in Figures 4 and 6, which contain the counterfactuals for leverage, investment, and dividends, respectively. To constrict each of these 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 4. In the top panel we see that leverage is largely unresponsive to taxes. The intuition is as follows. Taxes affect optimal choices via their effects on the firm's budget constraint and on the enforcement constraint. They do not affect the rate at which lenders discount the payments they receive from firms. The large effect of taxes seen in leverage models based on neoclassical investment models (Hennessy and Whited 2005, 2007) operates largely via the effect of taxes on the rate of return on debt relative to the firm's discount rate. Here, in contrast, we have separated out the revenue consequences of taxes from the discount rate consequences. The revenue consequences for old firms are minimal. The dividend nonnegativity constraint does not bind, and the quantitative effects of taxes on the enforcement constraint are tiny. Although the dividend nonnegativity constraint does bind for young firms, because the enforcement constraint binds for them as well, taxes again have little effect on their leverage.

Figure 4 contains two 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, lenders have little incentive to extend the firm much credit. Equivalently, a higher level of $\theta$ relaxes the collateral constraint. Second, the level of leverage does increase with the difference in discount factors, but only for mature firms. The young firms always face a binding collateral constraint, so the difference in discount factors does not affect their optimal policies. However, for mature firms, as the lender's patience approaches that of the firm, it becomes less costly for the firm to conserve debt capacity contingent on high productivity shocks. It thus engages in more of this risk management behavior and the average level of leverage falls.

Figure 5 contains our counterfactuals for investment. We see in the top panel that taxes have a profound negative effect on investment, but only for young firms. The intuition is that the young firms are growing at a rate much faster than the depreciation rate, so the negative effect of taxes on the marginal product of capital affects them a great deal. In contrast, the old firms are just replacing depreciated capital, so the effect of taxes on the marginal product of capital has little effect on their optimal investment choices. In the bottom panel, we see that increasing $\theta$ (i.e. relaxing the enforcement constraint) leads to higher investment for the young firms because they can grow at a higher rate if they can borrow more.

Figure 6 contains our counterfactuals for dividends, which are exactly the opposite of those for investment. In the top panel, we find a strong positive effect of taxes on dividends because the lower productivity of capital means that the dividend nonnegativity constraint is less likely to bind.[6] Of course, this effect is absent from in the older firms that already pay dividends, so we see almost no effect. In the bottom panel, we see a small positive effect of relaxing the enforcement constraint on dividends for the young firms because they grow to a point at which they are unconstrained more quickly.

[6]This sharp response is an artifact of our exclusion of dividend taxes in the model. When we do include dividend taxes, this effect becomes strongly muted.

One concern with our analysis is that we have modeled the effects of taxes in such a way that the lender's discount rate remains unchanged in the face of taxes. However, Rampini and Viswanathan (2013) suggest that one reason for the difference between lenders' and borrows' discount factors might well be taxes. On the other hand, if tax rates are going to affect discount rates, it is hard to imagine that they would not have similar effects on lenders' and borrowers' discount rates. Another 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.

# 7 Conclusion

We have sought to deepen our understanding of whether and how corporate taxes affect capital structure. We have therefore estimated 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 exogenously.

Our main findings are twofold. 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. The literal intuition is that taxing the firm does not affect the rate at which the lender discounts the principal and interest payments from the firms. More generally, it is hard to imagine that taxation would change the *relative* discount rates of lenders and firms. So the strong discount rate effects that operate in other partial-equilibrium capital structure models are absent here. Further, taxes have almost no effect on the constraint that allows the contract between the lender and the firm to be self-enforcing. In order for the firm to want to repay its debt, it must be better off repaying than repudiating the contract. The firm will want to repay only if the benefits from repayment are larger than the benefits from repudiating the contract, shedding debt, and starting over with fewer resources. The interest tax deduction makes shedding debt less attractive but it also makes the resources lost in default less valuable. Our estimation results show that these two effects quantitatively cancel each other out.

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.

Estimating an optimal contracting model has given us a new perspective on the relative importance of leverage versus agency issues for capital structure. 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.

**Proof of Proposition 1** We employ the technique in Abreu et al. (1990) to rewrite the sequential problem as a recursive problem.

Let $x = \{k_{j+1}, \{q_{j+1}(z)\}_{z\in\mathbb{Z}}\}_{j=0}^{\infty}$ denote a sequence of capital stocks and state-contingent debt holdings. For simplicity, we set $t = 0$. The initial state is $s_0 = (w_0, z_0)$, where $w_0 = z_0 k_0^{\alpha} + (1-\delta)k_0$, and $k_0$ is initial capital stock. The original problem of an entrant is stated as follows:

$$\max_x U(x) = \mathbb{E}_0 \sum_{j=0}^{\infty} \{(\beta(1-\phi))^j (z_j k_j^{\alpha} + k_j(1-\delta) - k_{j+1} - q + \beta_C(1-\phi)\mathbb{E}_{j+1} q_{j+1}(z)\} \tag{A.1}$$

subject to

$$z_j k_j^{\alpha} + k_j(1-\delta) - k_{j+1} - q + \beta_C(1-\phi)\mathbb{E}_{j+1} q_{j+1}(z) \geq 0 \tag{A.2}$$
$$U(x; z^{\tau}) \geq D(k_j; z^{\tau}), \quad \forall z^{\tau}, \tau > 0 \tag{A.3}$$
$$\text{given} \ \ k_0, z_0 \ \ \text{and} \ \ w_0. \tag{A.4}$$

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

The original problem can be written using net wealth as the state variable as in (14)–(16) in the main text. The logic of this formulation is similar to the previous work in Abreu et al. (1990), Atkeson (1991), and Bai and Zhang (2010). For ease of reference we reproduce these equations here:

$$V(w, z) = \max_{k', q'(z')} \quad w - k' + \beta_C(1-\phi)\mathbb{E}q'(z') + \beta(1-\phi)\mathbb{E}V(w'(z'), z') \tag{A.5}$$

subject to

$$w - k' + \beta_C(1-\phi)\mathbb{E}q'(z') \geq 0 \tag{A.6}$$
$$V(w'(z'), z') \geq V(\hat{w}(z'), z'), \quad \forall z' \in \mathbb{Z}, \tag{A.7}$$

where $w'(z') = z'k'^{\alpha} + k'(1-\delta) - q'(z')$ and $\hat{w}(z') = z'(k'(1-\theta))^{\alpha} + k'(1-\theta)(1-\delta)$.

We now demonstrate that the problem in (A.1)–(A.4) can be restated as (A.5)–(A.7). For the simplicity and without loss of generality, we assume $k' \in [\underline{k}, \bar{k}]$ and $q'(z) \in [\underline{q}, \bar{q}]$. Let $\bar{w}$ and $\underline{w}$ be the corresponding boundaries for net wealth.

Next, define the domain $S = \mathbb{Z} \times [\underline{w}, \bar{w}]$, and define $\tilde{V} : S \to R$ as an entrant's profit possibility correspondence. Here we first demonstrate this correspondence can be defined in a recursive way by adopting the notions of admissibility, self-generation, and factorization of Abreu et al. (1990). Then we show that the original problem characterizing the optimal contract can be rewritten as a functional equation defined in (A.5)-(A.7).

For any $s \in S$, the profit possibility correspondence is defined as:

$$\tilde{V}(s) = \{U(x) | x \ \ \text{satisfies constraints}(A.2) - (A.4) \ \ \text{given} \ \ k_0 \ \text{and} \ \ s = (w_0, a_0)\}$$

We assume that the correspondence is non-empty under some parameterizations.[7] Let $G : S \to R$ be a correspondence on the domain $S$. Assume that $G(s)$ is nonempty valued and uniformly

[7] In the literature where the punishment of repudiating is going back to autarky, this assumption is not necessary as the payoff in autarky is in the correspondence. But in most of cases, the outcome of interest is the non-autarky case. Thus, one must assume that another solution exists, for example, as in Kocherlakota (1996).

bounded for all $s \in S$. Denote $g = (k', \{q'(z)\}_{z\in\mathbb{Z}})$ as the vector of control variables. A function $U^c : S \to R$ is defined as a continuation value function with respect to $G$ if it is a selection from $G$, $U^c(s) \in G(s)$ for all $s$.

**Definition 1** *The pair $(g, U^c)$ of control variables and a continuation value function is* admissible with respect to $G$ at $s$ *if it satisfies*

$$w - k' + \beta_C(1-\phi)\mathbb{E}q(z') \geq 0; \tag{A.8}$$

$$U^c(w'(z'), z') \geq U^c(\hat{w}(z'), z'), \quad \forall z' \in \mathbb{Z}. \tag{A.9}$$

Denote $E(g, U^c)(s)$ as the payoff to the shareholders generated by a pair $(g, U^c)$.

$$E(g, U^c)(s) = w - k' + \beta_C(1-\phi)\mathbb{E}(q'(z')|z) + \beta\mathbb{E}(U^c(s')|z)$$

Denote $B(G)(s)$ as the set of payoffs that can be generated by the pairs $(g, U^c)$ admissible with respect to $G$ at $s$.

$$B(G)(s) = \{E(g, U^c)(s)|(g, U^c)\text{admissible with respect to G at S.}\}$$

**Definition 2** *The correspondence $G$ is self-generating if for all $s \in S$, $G(s) \subseteq B(G)(s)$.*

**Lemma 1 (Self-generation)** *If $G$ is self-generating, then for all $s \in S$, $B(G)(s) \subseteq \tilde{V}(s)$.*

**Proof.** Pick any $v \in B(G)(s)$, we want to show that $v \in \tilde{V}(s)$. This is done by constructing an allocation $x(v)$ such that $U(x(v)) = v$. We then prove the allocation satisfies (A.2)-(A.4).

1) Pick any $v \in B(G)(s_0)$, $s_0 \in S$ and satisfies (A.4). Thus, we can find an admissible pair $(g(s_0), U^c(s_0))$ such that $E(g(s_0), U^c(s_0)) = v$. Let $x(s_0, v) = g(s_0)$ and $w(s^1, v) = U^c(s_0)(s_1)$. $s^1 = \{s_0, s_1\}$ and $s_1 = (w_1, z_1)$. $w_1$ is calculated from $(k', \{q'(z)\})$. Since $U^c \in G$ and G is self-generating, $w(s^1, v) \subseteq G(s_1) \subseteq B(G)(s_1)$. Thus, there exists an admissible pair $(g(s_1), U^c(s_1))$ such that $E(g(s_1, U^c(s_1))) = w(s^1, v)$. Let $x(s^1, v) = g(s_1)$ and $w(s^2, v) = U^c(s_1)(s_2)$. By repeating this process, we can construct the allocation $x(v) = \{x(s^t, v)\}_{j=0}^{\infty}$.

2) We then need to show that $U(x(v)) = v$. According to the above formulation, $E(g(s_0), U^c(s_0)) = v$, and we then have

$$v = w_0 - k' + \beta_C(1-\phi)\mathbb{E}q'(z') + \beta\mathbb{E}U^c(s_0)(s_1). \tag{A.10}$$

From the sequential problem, we know

$$U(x(v)) = w_0 - k' + \beta_C(1-\phi)\mathbb{E}q'(z') + \beta\mathbb{E}U(x(v|s^1)). \tag{A.11}$$

where $x(v|s^1) = \{x(s^t, v)\}_{j=1}^{\infty}$. Subtracting (A.10) minus (A.11), we have

$$v - U(x(v)) = \beta\mathbb{E}(U^c(s_0)(s_1) - U(x(u|s^1)))$$

Thus

$$v - U(x(v)) \leq \beta \sup_{v_1 \in B(G)(s_1)} |U^c(s_0(s_1)) - U(x(v|s^1))|$$

As $v \in G(s_0)$ is arbitrarily picked, we have

$$\sup_{v \in B(G)(s_0)} |v - U(x(v))| \leq \beta \sup_{v_1 \in B(G)(s_1)} |U^c(s_0(s_1)) - U(x(v|s^1))|$$

Since $\beta < 1$, $B(G)(s)$ is uniformly bounded for any $s \in S$ and $U$ is also bounded, $v$ then equals $U(x(v))$ for any $v \in B(G)(s)$. Because $x(v)$ is constructed to satisfy condition (A.2)-(A.4), $v \in \tilde{V}(s)$. ■

**Lemma 2 (Factorization)** *$\tilde{V}(s) = B(\tilde{V})(s)$ for all $s \in S$.*

**Proof.** We need to show that $\tilde{V}(s)$ is self-generating. Pick $v \in \tilde{V}(s)$ for any $s \in S$. Suppose that $v$ is generated by an allocation $x(v)$ satisfying (A.2)-(A.4). We can construct an admissible pair $(g(s), U^c(s))$ such that $g(s) = x_0(v)$ and $U^c(s)(s_1) = U(x(v|s^1))$. Thus, $E(g(s), U^c(s)) = v$ and obviously the pair is admissible. Therefore, $\tilde{V}(s) \subseteq B(\tilde{V})(s)$. Using Lemma 1, we know $B(\tilde{V})(s) \subseteq \tilde{V}(s)$. As a result, $\tilde{V}(s) = B(\tilde{V})(s)$. ■

After showing that the payoffs to the shareholders can the defined recursively as the sum of the current dividend payout and the expected payoff from the next period onwards, we restate the value of the optimal contract for the shareholders as a function $U^{opt}(s)$,

$$U^{opt} = \sup_{v \in \tilde{V}(s)} v.$$

According to Lemmas 1 and 2, $U^{opt}(s)$ can be rewritten as follows:

$$(P): \quad U^{opt}(s) = \sup_{(g, U^c)} w_0 - k' + \beta_C(1-\phi)\mathbb{E}q(z') + \beta \mathbb{E}U^c(s)(s'), \tag{A.12}$$

where $(g, U^c)$ is admissible with respect to $\tilde{V}$ at s. Next we prove the the maximum $U^{opt}(s)$ exists since $\tilde{V}(s)$ is compact.

**Lemma 3** *If $G$ has a compact graph, then $B(G)$ has a compact graph.*

**Proof.** 1) Boundedness. As $G$ has a compact graph, $U^c(s)$ is then bounded. Because the state variables $s$ and the control variables $g(s)$ are all bounded, $B(G)$ has a bounded graph.

2) Closedness. Let $\{v_n, s_n\}_{n=1}^{\infty}$ be a sequence in the graph of $B(G)$ converging to $(v, s)$. Thus, we can find a sequence $\{g_n, U_n^{c*}\}_{n=1}^{\infty}$ in which $(g_n, U_n^{c*})$ is admissible with respect to G at S. $U_n^{c*} = \{U_n^c(w_n(z), z)\}_{z \in \mathbb{Z}}$. $w_n(z)$ is the net wealth calculated from $g_n$. This sequence belongs to a compact set, thus we can find a convergent subsequence with a limit $(g, U^{c*})$. By the property of the limit, we know $(g, U^{c*})$ satisfies conditions (A.6) and (A.7). By the continuation of $E$ operator, we have $E(g, U^{c*}) = v$. Thus, $(v, s)$ is in the graph of $B(G)$, that is, $B(G)$ has a closed graph. ■

**Lemma 4** *If graph($G_1$) $\subseteq$ graph ($G_2$), then graph($B(G_1)$) $\subseteq$ graph($B(G_2)$).*

**Proof.** Because the domains of $G_1$ and $G_2$ itself are subsets of the domain of $G_2$, the constraint set defined by $B(G_1)$ is thus contained in that defined by $B(G_2)$. ■

**Lemma 5** *$\tilde{V}$ has a compact graph.*

**Proof.** $\tilde{V}$ has a bounded graph because its domain is bounded below and above, so we can easily construct an upper bound. For the closedness, we want to show that the graph of V equals to the closure of its graph. Define the correspondence $V_1$ such that graph($V_1$)=closure(graph($\tilde{V}$)). According to the definition of closure, graph($V_1$)$\subseteq$ graph($V_1$). By Lemma 4, graph(B($\tilde{V}$))$\subseteq$ graph(B($V_1$)). By Lemmas 1 and 2, graph(B($\tilde{V}$))=graph($\tilde{V}$). By Lemma 3, graph(B($V_1$)) is closed because $V_1$ has a compact graph. As graph($V_1$) is the smallest closed set containing graph($B(V_1)$), graph($V_1$)$\subseteq$graph$B(V_1)$. Thus, $V_1$ is self-generating and graph($V_1$) $\subseteq$ graph($\tilde{V}$) according to Lemma 2. Thus, graph($V_1$) = graph($\tilde{V}$) and $\tilde{V}$ has a closed graph. ■

By Lemma 5, an optimal contract exists for each $s$. Next we will discuss conditions under which $\tilde{V}(s)$ is continuous in the state variables. Because $\mathbb{Z}$ has finite support, we can work with the vector $U^d(s) \equiv (U^c(s)(w_1, z_1), U^c(s)(w_2, z_2), \dots, U^c(s)(w_N, z_N))$, where $N = \#\mathbb{Z}$, and $w_n = z_n k'^\alpha + (1-\delta)k' - q'(z')$, $n \in 1, \dots, N$. Similarly, we can define the diversion vector $U^D(s) \equiv (U^c(s)(\hat{w}_1, z_1) U^c(s)(\hat{w}_2, z_2), \dots, U^c(s)(\hat{w}_N, z_N))$, where $\hat{w}_n = z_n(k'(1-\theta))^\alpha + (1-\delta)k'(1-\theta)$, $n \in 1, \dots, N$. By rewriting the $\mathbb{E}$ operator defining the payoffs to the shareholders in terms of the combinations of the continuation function and control variables, $U^D$ and $U^d$, the $\mathbb{E}$ operator is then continuous in all their arguments.

**Lemma 6** *$\tilde{V}$ is a continuous correspondence and $U^{opt}$ is a continuous function.*

**Proof.**

By Lemma 5, $\tilde{V}$ has a compact graph thus $\tilde{V}$ is upper hemicontinuous. Next we prove the lower hemicontinuity. Pick $v \in \tilde{V}(s)$ and a pair $(g, U^d)$. $E(g, U^d) = v$. For any sequence $s_n \to s$, we need to find an $M$ and a sequence $\{v_n\}_{n=M}^\infty$ such that $v_n \to y$ and $v_n \in B(\tilde{V})(s_n)$ for $n \geq M$. The construction is as follows. Take sequences $k_n \to k'$ and $q_n'(z) \to q'(z)$. For any $\epsilon > 0$, due to the continuity of the payoff function E, we can find $\delta_1 > 0$ and $\delta_2 > 0$ and a $M_1$ such that $|k_n - k'| < \delta_1$ and $|q_n'(z) - q'(z)| < \delta_2$ for any $z$ and $w - k_n' + \beta_C(1-\phi)\mathbb{E}q_n'(z) > 0$, we have $|E(g_n, U^d(s)) - E(g, U^d(s))| < \epsilon/2$ when $n \geq M_1$. On the other hand, we can find an $\delta_3$ and a $M_2$ such that $|s_n - s| < \delta_3$, let $U^d(s_n) = U^d(s)$, we have $|E(g_n, U^d)(s_n) - E(g_n, U^d)(s)| < \epsilon/2$ when $n \geq M_2$. By the triangle inequality, we have $|E(g_n, U^d)(s_n) - E(g, U^d)(s)| < \epsilon$ when $n \geq M = \max(M_1, M_2)$. By construction, we know $(g_n(s_n), U_n^d(s_n))$ is an admissible pair of $\tilde{V}$ at $s_n$ when $n \geq M$. Let $v_n = E(g_n, U^d(s_n))$, we know $v_n \to v$ when $n \geq M$. Thus $\tilde{V}$ is lower hemicontinous. Thus, $\tilde{V}$ is a continuous correspondence. By the maximum theorem, $U^{opt}$ is then a continuous function. ■

**Lemma 7** *$U^{opt}(w, z)$ increases strictly with $w$ for any $z \in \mathbb{Z}$.*

**Proof.** Pick $w_1 > w_2$ and any $z \in \mathbb{Z}$. The constraint set of $U^{opt}(w_1, z)$ contains that of $U^{opt}(w_2, z)$ as the non-negative dividend constraint is looser with a larger $w_1$. Thus the objective function $U^{opt}(w_1, z) > U^{opt}(w_2, z)$. ■

**Assumption 1** *For any $(w, z) \in S$, the set $\{(k', \{q'(z')\}) | w - k' + \beta_C(1-\phi)\mathbb{E}q'(z') \geq 0; \quad z'k'^\alpha + k'(1-\delta) - q'(z') \geq z'(k'(1-\theta))^\alpha + k'(1-\theta)(1-\delta), \forall z' \in \mathbb{Z}.\}$ is non-empty.*

By Assumption 1, there exists at least one admissible pair $(g, U^{opt})$ such that $E(g, U^{opt})(s) \in \tilde{V}(s)$. With this assumption, we show that the optimal contract can be characterized in a recursive way.

**Lemma 8** *The continuation function $\hat{U}^c$ solving the program (P) from (A.12) equals to $U^{opt}$.*

**Proof.** The proof proceeds by contradiction. Suppose $(g, \hat{U}^c)$ solving the program (P) with $\hat{U}^c(s)(w', z') < U^{opt}(w', z')$ for some $s$ and $(w', z')$. $g = (k', \{q'(z')\})$. We next find a $\tilde{w}$ such that $U^{opt}(\tilde{w}, z') = \hat{U}^c(s)(w', z')$. Under Assumption 1, there always exists some $U^{opt}(w^0, z')$ such that $\hat{U}^c(s)(w', z') \geq U^{opt}(w^0, z')$. Otherwise, we can construct a new function with $U^c(s)(w', z') = U^{opt}(w^0, z')$. This new function generates a strictly higher payoff thus contradicts the fact of the optimal $\hat{U}^c$. Thus, we can apply the intermediate value theorem on the continuous function $U^{opt}$, and $\tilde{w}$ is then well-defined. As $U^{opt}(w') > U^{opt}(\tilde{w})$, $w' > \tilde{w}$. By fixing $\tilde{k} = k'$, we have $q'(z') < \tilde{q}(z')$. Let $\tilde{q}(z) = q'(z)$ for $z \neq z'$. The pair $((\tilde{k}, \{\tilde{q}(z)\}), U^{opt})$ is then admissible and generates a strictly higher payoff that $(g, \hat{U}^c)$. That contradicts with the assumption that $\hat{U}^c$ solves the program (P). ■

**Proof of Proposition 2** We want to prove that the enforcement constraint (A.7) can take the form of a collateral constraint.

$$z'k'^{\alpha} + k'(1-\delta) - q'(z') \geq z'(k'(1-\theta))^{\alpha} + k'(1-\theta)(1-\delta), \quad \forall z' \in \mathbb{Z}. \tag{A.13}$$

According to Lemma 7 and 8, $V$ strictly increases with $w$. If the control variables satisfy (A.13), (A.7) also holds because $V$ increases with $w$ given $z'$. On the other side, if (A.7) holds with some set of control variables, (A.13) must be satisfied, that is, $w' \geq \hat{w}$. Suppose not, if $w' < \hat{w}$, then $V(w', z') < V(\hat{w}, z')$ as $V$ strictly increases with the net wealth $w$, which is a contradiction. ■

**Proof of Proposition 3** Now we show that the functional equation defined in (A.5)-(A.7) has a solution. The first best problem is defined as

$$W(w, z) = \max_{k', q'} \quad w - k' + \beta(1-\phi)q' + \beta(1-\phi)\mathbb{E}W(w', z'), \tag{A.14}$$

where $w = zk^{\alpha} + k(1-\delta) - q$. Applying the contraction mapping theorem and the theorem of maximum, the first best problem has a unique solution, $W$, which is continuous and strictly increases with the current net wealth w. Define the mapping $T$ in the space of bounded and continuous functions.

$$T(V)(w, z) = \max_{k', q(z')} \quad w - k' + \beta_C(1-\phi)\mathbb{E}q(z') + \beta(1-\phi)\mathbb{E}V(w'(z'), z') \tag{A.15}$$

subject to constraints (A.6) and (A.7).

Using $W$ as an initial guess, the sequence of functions $\{T^n(W)\}_{n=0}^{\infty}$ converges to a solution of the mapping $T$. Define the operator $L$ as follows:

$$L(V) = \max_{k', q(z')} \quad w - k' + \beta_C(1-\phi)\mathbb{E}q(z') + \beta(1-\phi)\mathbb{E}V(w'(z'), z') \tag{A.16}$$

subject to constraints (A.6) and (A.13). With the standard proof of the contraction mapping theorem and the maximum theorem, the $L$ operator has a unique fixed point $V^*$, which is continuous and strictly increases with $w$. By using $W$ as an initial guess, we construct a sequence $\{L^n(W)\}_{n=0}^{\infty}$ convergent to $V^*$. As $W$ is a continuous function and strictly increases with $w$, every function in $\{T^n(W)\}_{n=0}^{\infty}$ increases strictly with $w$. We thus have $L^n(W) = T^n(W), \forall n$. $\{T^n(W)\}_{n=0}^{\infty}$ then converges to $V^*$. Next we prove $V^*$ is a fixed point of $T$ operator. Because $T(V^*) = L(V^*)$ and $L(V^*) = V^*$, we have $T(V^*) = V^*$. That is, $V^*$ is a fixed point of the T operator. ■

# Appendix B

This appendix summarizes the numerical methods used to solve the model in the main text. The basic procedure follows Proposition 3 by using value function iteration. As shown in Appendix A, the model has a unique fixed point if a solution exists under a specific parameterization, so it is important to check the existence of a solution.

The main idea is to search for a set of feasible boundaries of net wealth $w$, the capital stock $k$, and debt $q$ such that the model has a well-defined solution, given a specific parameterization. If we can find these boundaries, the model has a solution numerically. We use the enumeration method to search for these boundaries. In particular, the upper bound of $k$, $\bar{k}$, can be derived from the first-best solution. By fixing the lower bound of $q$, $\underline{q}$, 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 $\underline{w}$ and $\underline{k}$. Thus, we only need to search for the feasible values of $\underline{w}$ and $\underline{k}$ in addition to $\underline{q}$. The details of this computational method are as follows:

(a) We first define the sets of possible values of $\underline{w}$ and $\underline{k}$. As $w$ and $k$ are both strictly positive[8], we set the lower bounds of $\underline{w}$ and $\underline{k}$ to be small positive numbers. The upper bounds of $\underline{w}$ and $\underline{k}$ can be derived from the first-best solution. Then, the possible values of $\underline{w}$ and $\underline{k}$ are divided into two vectors each with 30 equal-spaced grid points. As a starting point, we let $\underline{q} = 0$.

(b) Next we search over the two vectors for $\underline{w}$ and $\underline{k}$ in an ascending order. Give one set of boundaries, the state space of $w$ has 32 equal-spaced grids. The control space of $k$ has 100 equally spaced grids. The control space of $q(z')$ has 80 equally spaced grids. 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 11 times. The value function iteration method is described in detail below. During this process, there are two criteria to quit the iteration process and proceed to the next possible boundaries: First, the constraint set is empty for at least one pair of grids $(w, z)$; Second, at least one of the policy functions hits the lower bound of $k$ or the upper bound of $q(z')$. The search process will then go back to step (b) and repeat this process for the next set of boundaries. If the iteration survives after 11 times, we proceed to step (d).

(d) If steps (b) through (c) do not produce a feasible set of boundaries, the procedure for finding boundaries then stops. If steps (b) through (c) produces one feasible pair of $\underline{w}$ and $\underline{k}$ but the lower bound of $q$ is hit by the policy function for optimal debt, $\underline{q}$ needs to be updated and the procedure goes back to step (a). Otherwise, we say the procedure find a feasible set of boundaries.

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

[8]See Lemma 6 in Rampini and Viswanathan (2013).

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 $k$, $i = 1, \ldots, n$, $t = 1, \ldots, T$, 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

$$\frac{1}{nT}\sum_{i=1}^{n}\left(\sum_{t=1}^{T}\phi_{it}\right)\left(\sum_{t=1}^{T}\phi_{it}\right)'$$

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

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

# References


Abreu, Dilip, David Pearce, and Ennio Stacchetti, 1990, Toward a theory of discounted repeated games with imperfect monitoring, *Econometrica* 58, 1041–1063.

Albuquerque, Rui, and Hugo A. Hopenhayn, 2004, Optimal lending contracts and firm dynamics, *Review of Economic Studies* 71, 285–315.

Andrade, Gregor, and Steven N. Kaplan, 1998, How costly is financial (not economic) distress? evidence from highly leveraged transactions that became distressed, *Journal of Finance* 53, 1443–1493.

Atkeson, Andrew, 1991, International lending with moral hazard and risk of repudiation, *Econometrica* 1069–1089.

Bai, Yan, and Jing Zhang, 2010, Solving the Feldstein–Horioka puzzle with financial frictions, *Econometrica* 78, 603–632.

Biais, Bruno, Thomas Mariotti, Jean-Charles Rochet, and Stéphane Villeneuve, 2010, Large risks, limited liability, and dynamic moral hazard, *Econometrica* 78, 73–118.

Cooley, Thomas, Ramon Marimon, and Vincenzo Quadrini, 2004, Aggregate consequences of limited contract enforceability, *Journal of Political Economy* 112, 817–847.

DeAngelo, Harry, Linda DeAngelo, and Toni M. Whited, 2011, Capital structure dynamics and transitory debt, *Journal of Financial Economics* 99, 235–261.

DeMarzo, Peter, Mike Fishman, Zhiguo He, and Neng Wang, 2012, Dynamic agency and the $q$ theory of investment, *Journal of Finance* 67, 2295–2340.

DeMarzo, Peter M., and Yuliy Sannikov, 2006, Optimal security design and dynamic capital structure in a continuous-time agency model, *Journal of Finance* 61, 2681–2724.

Duffie, Darrell, and Kenneth J Singleton, 1993, Simulated moments estimation of markov models of asset prices, *Econometrica* 61, 929–52.

Erickson, Timothy, Colin Jiang, and Toni M. Whited, 2013, Minimum distance estimation of the errors-in-variables model using linear cumulant equations, with two corporate finance applications, Manuscript, University of Rochester.

Erickson, Timothy, and Toni M. Whited, 2002, Two-step GMM estimation of the errors-in-variables model using high-order moments, *Econometric Theory* 18, 776–799.

Graham, John R., 2007, Taxes and corporate finance, in B. Espen Eckbo, ed., *Handbook of Empirical Corporate Finance*, 59–33 (Elsevier).

Graham, John R., Mark T. Leary, and Michael R. Roberts, 2013, A century of capital structure: The leveraging of corporate america, Manuscript, Duke University.

Han, Chirok, and Peter C. B. Phillips, 2010, GMM estimation for dynamic panels with fixed effects and strong instruments at unity, *Econometric Theory* 26, 119–151.

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.

Hennessy, Christopher A., and Toni M. Whited, 2005, Debt dynamics, *Journal of Finance* 60, 1129–1165.

Hennessy, Christopher A., and Toni M. Whited, 2007, How costly is external financing? Evidence from a structural estimation, *Journal of Finance* 62, 1705–1745.

Ingram, Beth F., and Bong-Soo Lee, 1991, Simulation and estimation of time series models, *Journal of Econometrics* 47, 197–205.

Jermann, Urban, and Vincenzo Quadrini, 2012, Macroeconomic effects of financial shocks, *American Economic Review* 102, 238–271.

Jermann, Urban J., and Vincenzo Quadrini, 2007, Stock market boom and the productivity gains of the 1990s, *Journal of Monetary Economics* 54, 413–432.

Khan, Aubhik, and Julia K Thomas, 2008, Idiosyncratic shocks and the role of nonconvexities in plant and aggregate investment dynamics, *Econometrica* 76, 395–436.

Kocherlakota, Narayana, 1996, Implications of efficient risk sharing without commitment, *Review of Economic Studies* 636, 596–610.

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

McKeon, Stephen, 2013, Firm-initiated versus investor-initiated equity issues, Manuscript, University of Oregon.

Morellec, Erwan, Boris Nikolov, and Norman Schürhoff, 2012, Corporate governance and capital structure dynamics, *Journal of Finance* 67, 803–848.

Nikolov, Boris, and Lukas Schmid, 2012, Testing dynamic agency theory via structural estimation, Manuscript, University of Rochester.

Pérez-González, Francisco, Frederic Panier, and Pablo Villanueva, 2012, Capital structure and taxes: What happens when you (also) subsidize equity?, Manuscript, Stanford University.

Rampini, Adriano A., and S. Viswanathan, 2010, Collateral, risk management, and the distribution of debt capacity, *Journal of Finance* 65, 2293–2322.

Rampini, Adriano A., and S. Viswanathan, 2013, Collateral and capital structure, *Journal 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, *The Review of Economic Studies* 54, 599–617.

Tauchen, George, and Robert Hussey, 1991, Quadrature-based methods for obtaining approximate solutions to nonlinear asset pricing models, *Econometrica* 59, 371–396.

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

Thomas, Jonathan, and Tim Worrall, 1994, Foreign direct investment and the risk of expropriation, *The Review of Economic Studies* 61, 81–108.

## Table 1: Simulated Moments Estimation

Calculations are based on a sample of nonfinancial firms from the annual 2012 COMPUSTAT industrial files. The sample period is from 1964 to 2011. 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. $\phi$ is the exogenous exit rate. $\theta$ 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

| | | No Taxes | | Taxes | |
|---|---|---|---|---|---|
| | Actual | Simulated | T-Statistic | Simulated | T-Statistic |
| Average net debt | 0.150 | 0.145 | 1.899 | 0.147 | 1.603 |
| Std. Dev. of net debt | 0.145 | 0.052 | 3.075 | 0.039 | 3.229 |
| Average investment | 0.104 | 0.123 | -6.629 | 0.104 | 0.009 |
| Std. Dev. of investment | 0.060 | 0.140 | -3.145 | 0.125 | -2.277 |
| Average profits | 0.155 | 0.137 | 2.988 | 0.148 | 2.158 |
| Std. Dev. of profits | 0.076 | 0.032 | 3.338 | 0.040 | 2.103 |
| Serial correlation profits | 0.710 | 0.695 | 0.077 | 0.722 | -0.052 |
| Average Tobin's $q$ | 2.453 | 2.358 | 1.422 | 2.260 | 0.552 |
| Average distributions | 0.028 | 0.023 | 3.801 | 0.024 | 1.664 |
| Std. Dev. of distributions | 0.032 | 0.078 | -3.039 | 0.072 | -6.068 |
| Corr. investment and leverage | -0.152 | -0.515 | 4.893 | -0.442 | 1.797 |
| Corr. profits and leverage | -0.154 | -0.097 | -0.339 | -0.080 | -0.376 |

B. Parameter estimates

| | $\delta$ | $\alpha$ | $\rho_z$ | $\sigma_z$ | $\phi$ | $\theta$ | $k_0/k^*$ | $\beta_C - \beta$ |
|---|---|---|---|---|---|---|---|---|
| No Tax | | | | | | | | |
| | 0.060 | 0.802 | 0.705 | 0.189 | 0.005 | 0.143 | 0.049 | 0.028 |
| | (0.008) | (0.031) | (0.113) | (0.037) | (0.001) | (0.013) | (0.014) | (0.022) |
| Taxes | | | | | | | | |
| | 0.035 | 0.775 | 0.712 | 0.203 | 0.010 | 0.148 | 0.036 | 0.027 |
| | (0.009) | (0.001) | (0.302) | (0.065) | (0.006) | (0.014) | (0.017) | (0.018) |

## Table 2: Simulated Moments Estimation: Small versus Large Firms

Calculations are based on a sample of nonfinancial firms from the annual 2012 COMPUSTAT industrial files. The sample period is from 1964 to 2011. 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 the model with corporate taxation, and the sample is split into firms below and above the median real, 2002, asset size. 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. $\phi$ is the exogenous exit rate. $\theta$ 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

| | Small | | | Large | | |
|---|---|---|---|---|---|---|
| | Simulated | Actual | T-Statistic | Simulated | Actual | T-Statistic |
| Average net debt | 0.121 | 0.122 | -0.232 | 0.269 | 0.262 | 1.214 |
| Std. Dev. of net debt | 0.145 | 0.018 | 3.794 | 0.130 | 0.057 | 2.276 |
| Average investment | 0.105 | 0.108 | 0.701 | 0.103 | 0.092 | 1.897 |
| Std. Dev. of investment | 0.059 | 0.113 | -9.850 | 0.056 | 0.085 | -2.158 |
| Average profits | 0.143 | 0.138 | 1.452 | 0.167 | 0.162 | 0.938 |
| Std. Dev. of profits | 0.079 | 0.035 | 4.057 | 0.067 | 0.042 | 2.980 |
| Serial correlation profits | 0.726 | 0.678 | 0.246 | 0.686 | 0.595 | 0.234 |
| Average Tobin's $q$ | 2.039 | 1.987 | 0.543 | 2.844 | 2.290 | 3.786 |
| Average distributions | 0.022 | 0.027 | -1.129 | 0.033 | 0.037 | -1.613 |
| Std. Dev. of distributions | 0.030 | 0.076 | -5.284 | 0.034 | 0.040 | -2.509 |
| Corr. investment and leverage | -0.163 | -0.209 | 1.247 | -0.164 | -0.143 | -1.787 |
| Corr. profits and leverage | -0.105 | -0.196 | 0.666 | -0.209 | -0.363 | 0.828 |

B. Parameter estimates

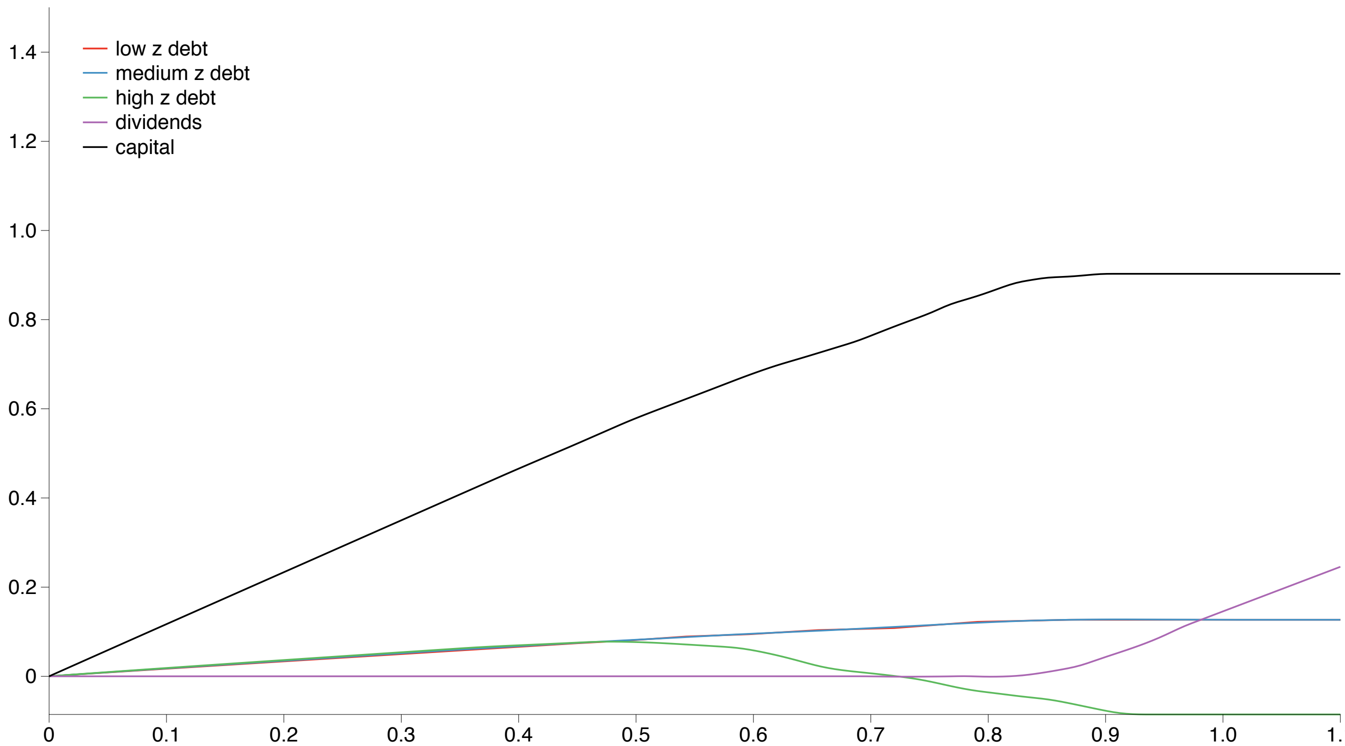
| | $\delta$ | $\alpha$ | $\rho_z$ | $\sigma_z$ | $\phi$ | $\theta$ | $k_0/k^*$ | $\beta_C - \beta$ |
|---|---|---|---|---|---|---|---|---|
| Small | | | | | | | | |
| | 0.031 | 0.736 | 0.790 | 0.243 | 0.039 | 0.121 | 0.056 | 0.027 |
| | (0.007) | (0.222) | (0.041) | (0.050) | (0.012) | (0.019) | (0.014) | (0.019) |
| Large | | | | | | | | |
| | 0.037 | 0.698 | 0.828 | 0.195 | 0.001 | 0.365 | 0.056 | 0.019 |
| | (0.006) | (0.039) | (0.170) | (0.092) | (0.000) | (0.033) | (0.071) | (0.048) |

## Table 3: Industry Parameter Estimates

Calculations are based on a sample of nonfinancial firms from the annual 2012 COMPUSTAT industrial files. The sample period is from 1964 to 2011. 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. $\phi$ is the exogenous exit rate. $\theta$ 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. Clustered standard errors are in parentheses.

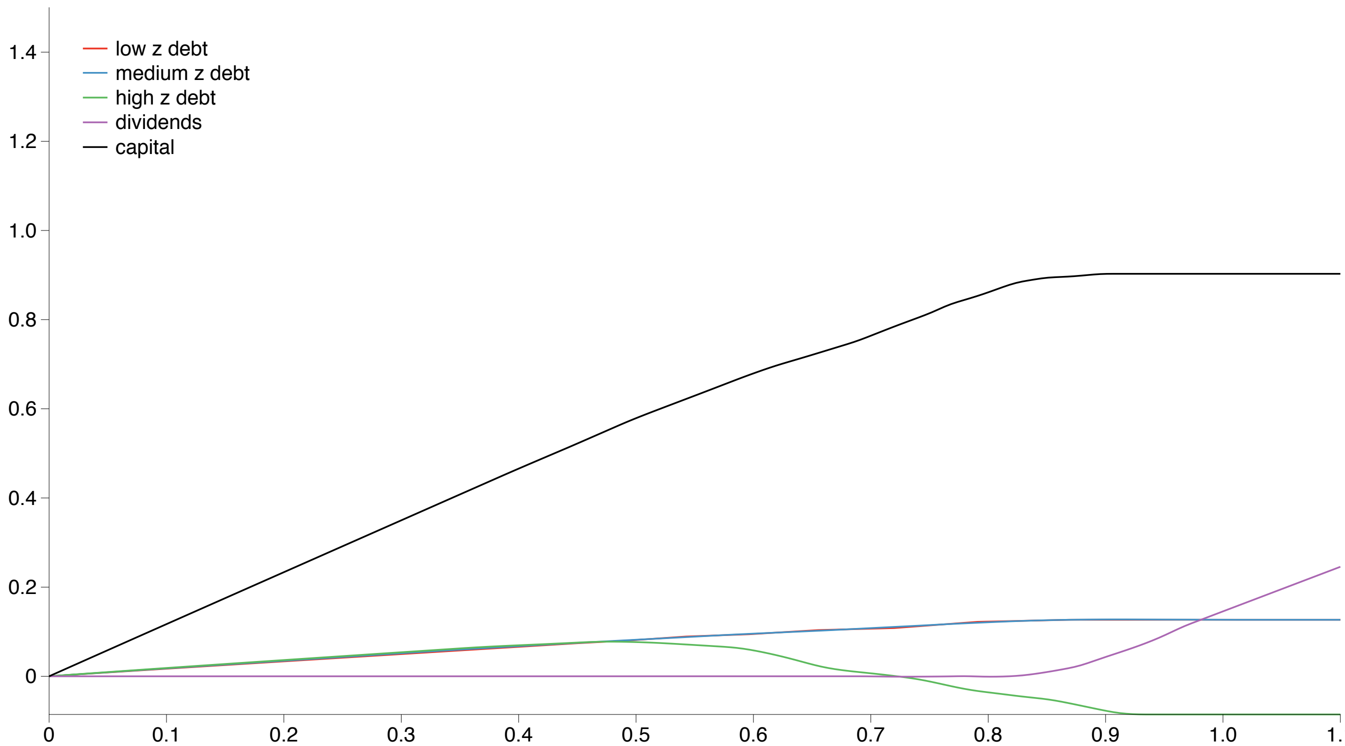
| | $\delta$ | $\alpha$ | $\rho_z$ | $\sigma_z$ | $\phi$ | $\theta$ | $k_0/k^*$ | $\beta_C - \beta$ |
|---|---|---|---|---|---|---|---|---|
| Oil and Gas (13) | | | | | | | | |
| | 0.041 | 0.743 | 0.632 | 0.278 | 0.020 | 0.277 | 0.166 | 0.018 |
| | (0.038) | (0.006) | (0.041) | (0.090) | (0.018) | (0.072) | (0.190) | (0.006) |
| Food (20) | | | | | | | | |
| | 0.046 | 0.550 | 0.622 | 0.093 | 0.017 | 0.164 | 0.190 | 0.008 |
| | (0.024) | (0.096) | (0.066) | (0.019) | (0.005) | (0.043) | (0.079) | (0.016) |
| Chemicals (28) | | | | | | | | |
| | 0.026 | 0.529 | 0.789 | 0.176 | 0.001 | 0.119 | 0.127 | 0.034 |
| | (0.014) | (0.155) | (0.053) | (0.036) | (0.001) | (0.058) | (0.073) | (0.081) |
| Machinery (35) | | | | | | | | |
| | 0.053 | 0.645 | 0.668 | 0.133 | 0.002 | 0.104 | 0.111 | 0.024 |
| | (0.012) | (0.072) | (0.002) | (0.081) | (0.001) | (0.028) | (0.054) | (0.041) |
| Electronics (36) | | | | | | | | |
| | 0.062 | 0.868 | 0.703 | 0.312 | 0.012 | 0.064 | 0.031 | 0.035 |
| | (0.040) | (0.051) | (0.040) | (0.061) | (0.009) | (0.029) | (0.005) | (0.134) |
| Instrumentation (38) | | | | | | | | |
| | 0.028 | 0.681 | 0.747 | 0.219 | 0.013 | 0.094 | 0.034 | 0.021 |
| | (0.016) | (0.124) | (0.163) | (0.040) | (0.003) | (0.050) | (0.005) | (0.055) |

Figure 1: Policy Functions: No Tax Model



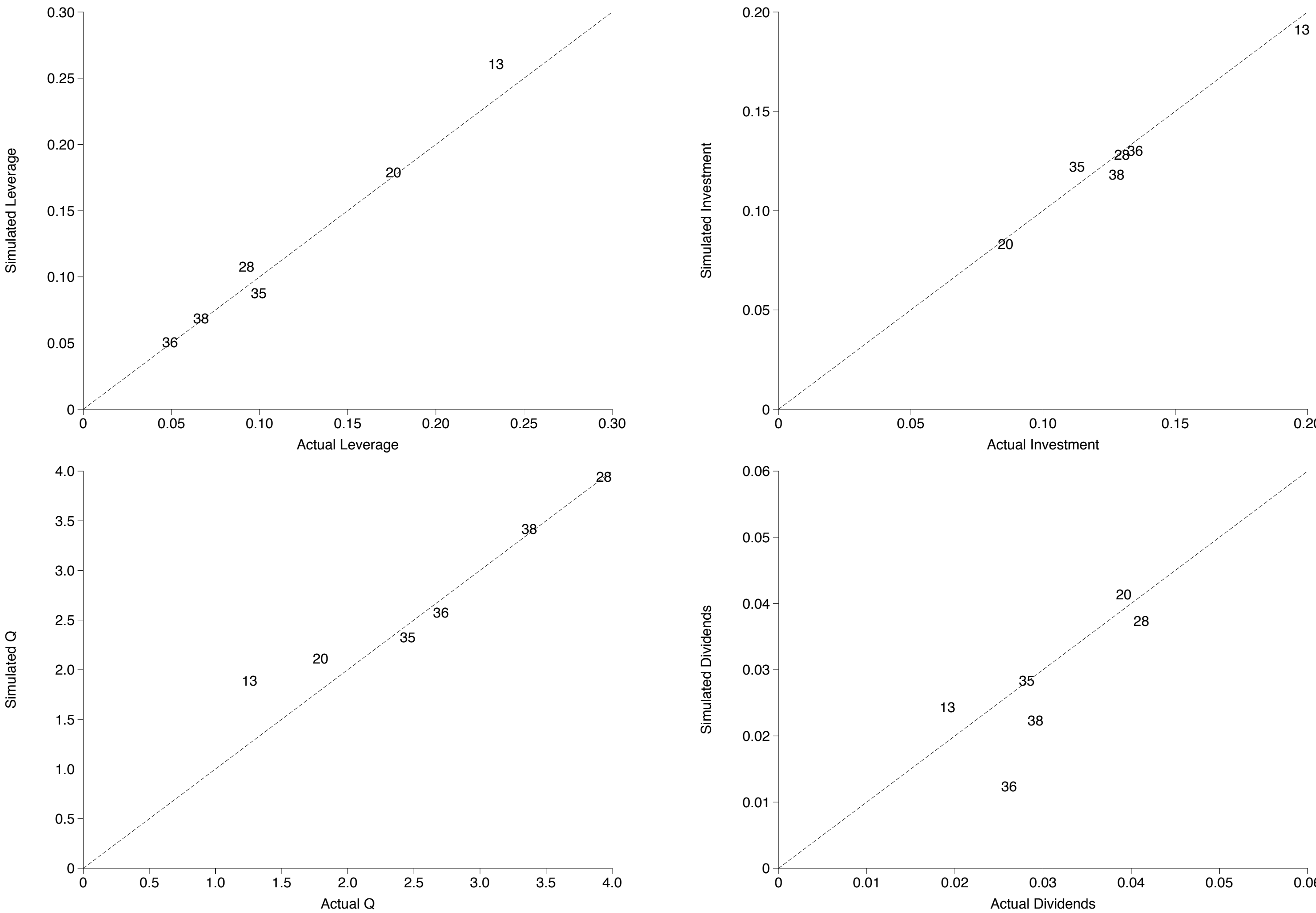

This figure depicts the policy functions for the model of Section 2, with the corporate tax rate set to zero and with all other parameters from the estimation of the no-tax model on the full sample. All of the variables are scaled by the unconstrained steady-state capital stock. The horizontal axis contains net wealth. On the left axis are dividends and debt contingent on a low future state, a medium future state, and a high future state. On the right axis is the capital stock.

Figure 2: Policy Functions: Tax Model



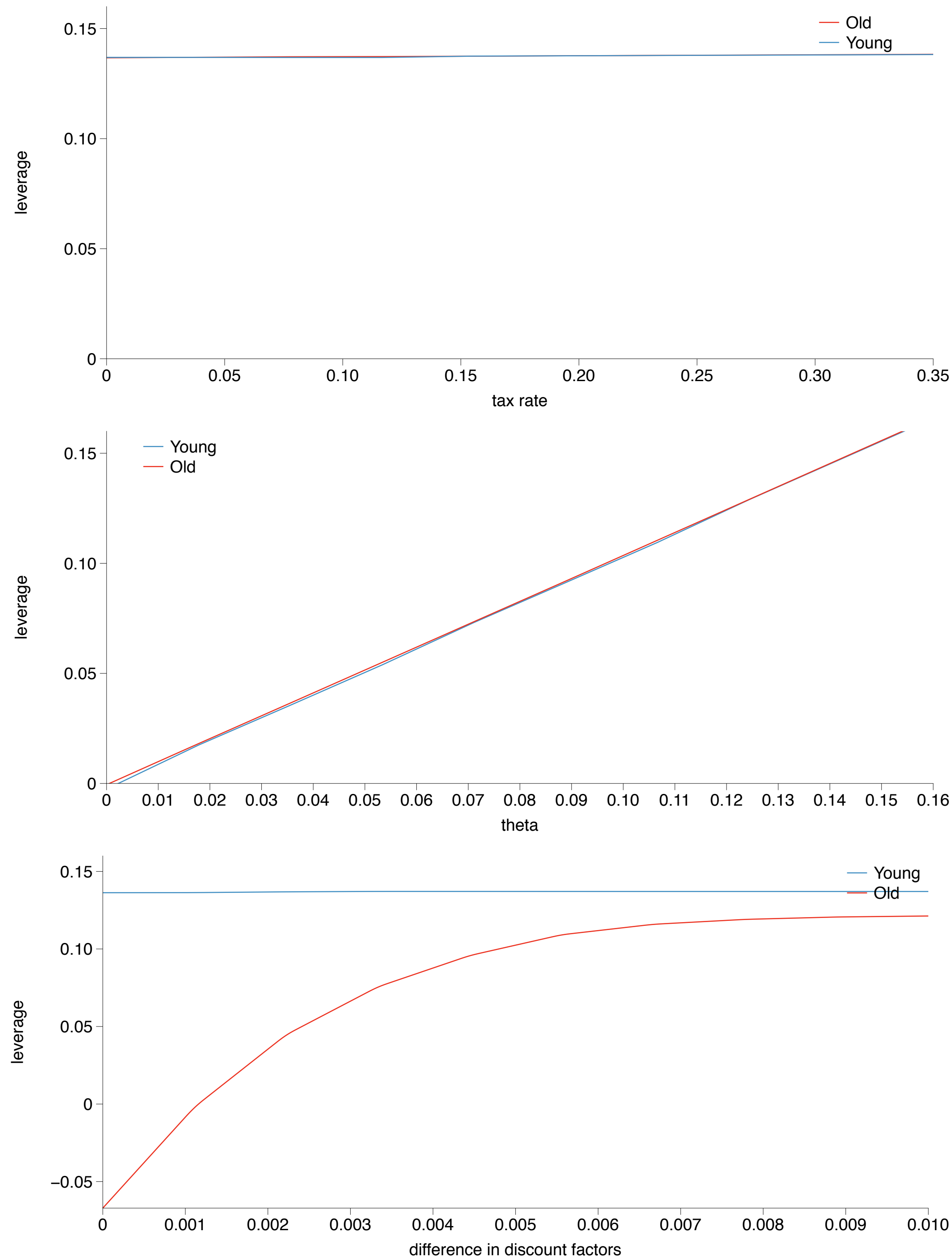

This figure depicts the policy functions for the model of Section 2, with the corporate tax rate set to 0.2 and with all other parameters from the estimation of the *no-tax* model on the full sample. We use the no-tax model parameters for comparability with Figure 1. All of the variables are scaled by the unconstrained steady-state capital stock. The horizontal axis contains net wealth. On the left axis are dividends and debt contingent on a low future state, a medium future state, and a high future state. On the right axis is the capital stock.

Figure 3: Moments from Industry Estimations



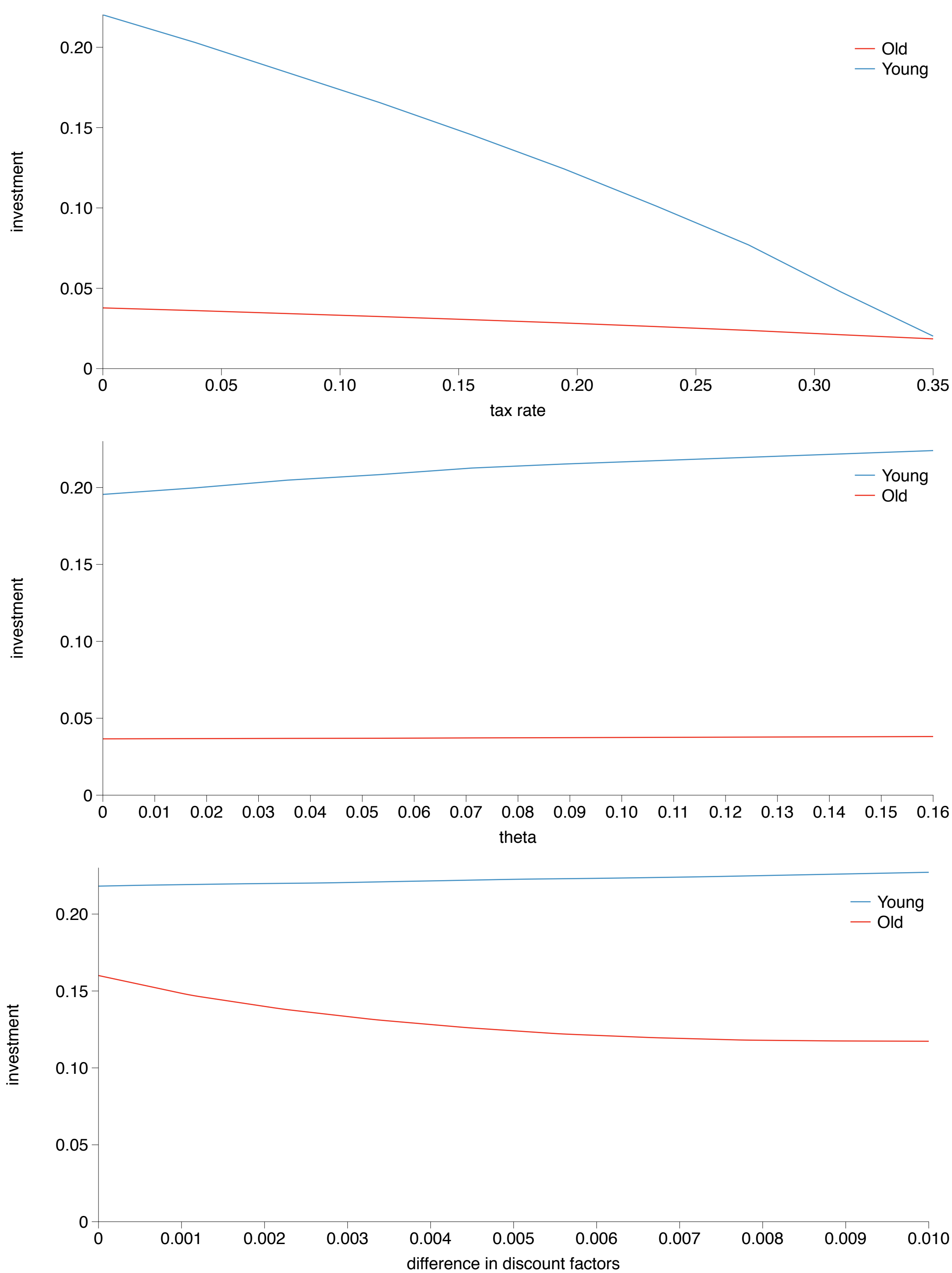

Calculations are based on a sample of nonfinancial firms from the annual 2012 COMPUSTAT industrial files. The sample period is from 1964 to 2011. The sample is split into six industry groups. SIC13 is oil and gas extraction; SIC20 is food products; SIC28 is chemicals and allied products; SIC35 is machinery and computer equipment; SIC36 is electronic and electrical equipment; and SIC38 is measuring instruments; SIC50 is wholesale trade; and SIC73 is business services. 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.

Figure 4: Counterfactuals: Leverage



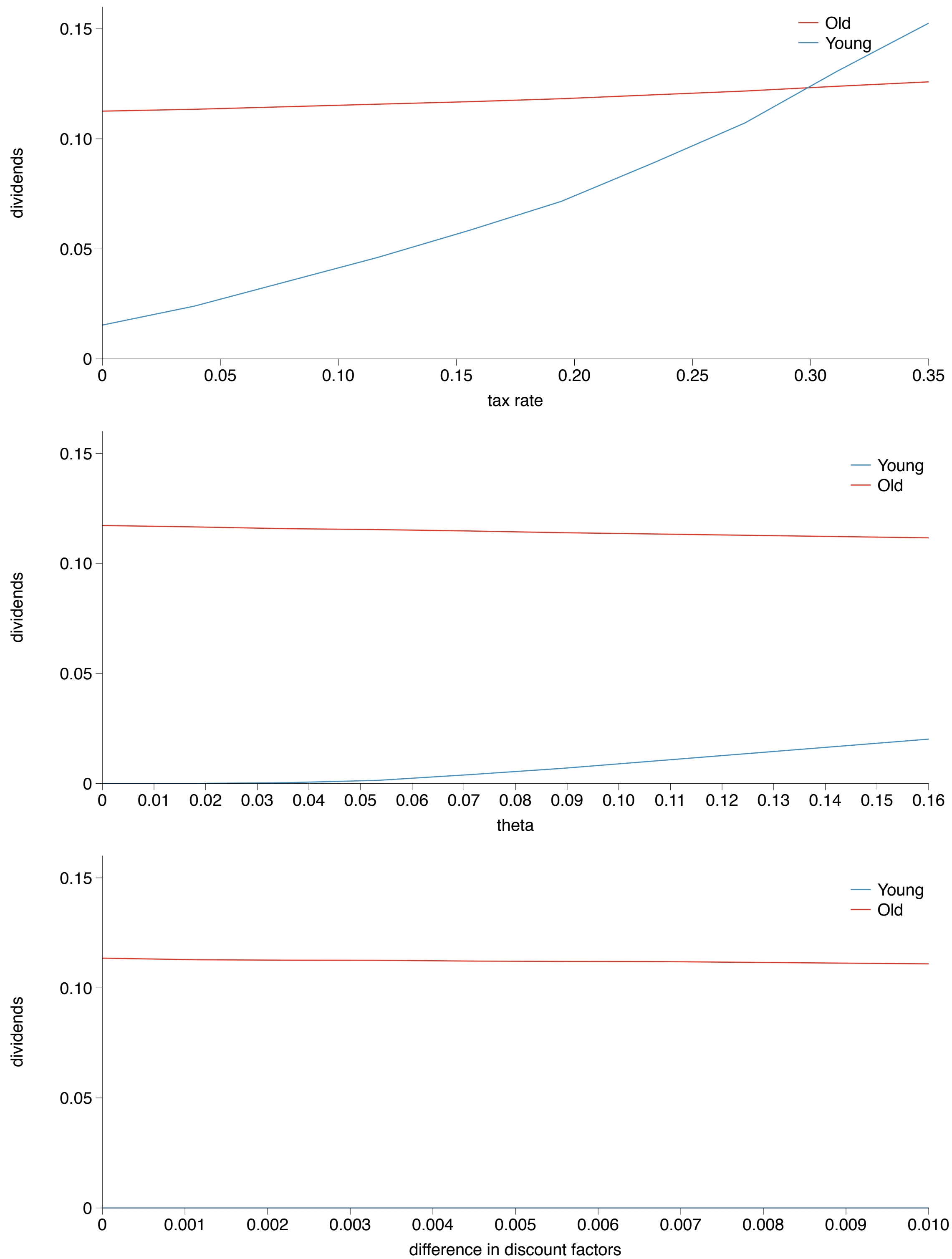

This figure is constructed as follows. We pick a grids for $\theta$, the fraction of the capital stock lost to the firm when it repudiates a contract, $\tau_c$, the corporate tax rate, and $\beta_C - \beta$, the difference in the discount factors between borrows and lenders. We then solve the model for each of the different parameter values, simulate the model for 50 time periods, and then plot average leverage. We use a parameterization that is the average of the parameter estimates 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 5: Counterfactuals: Investment



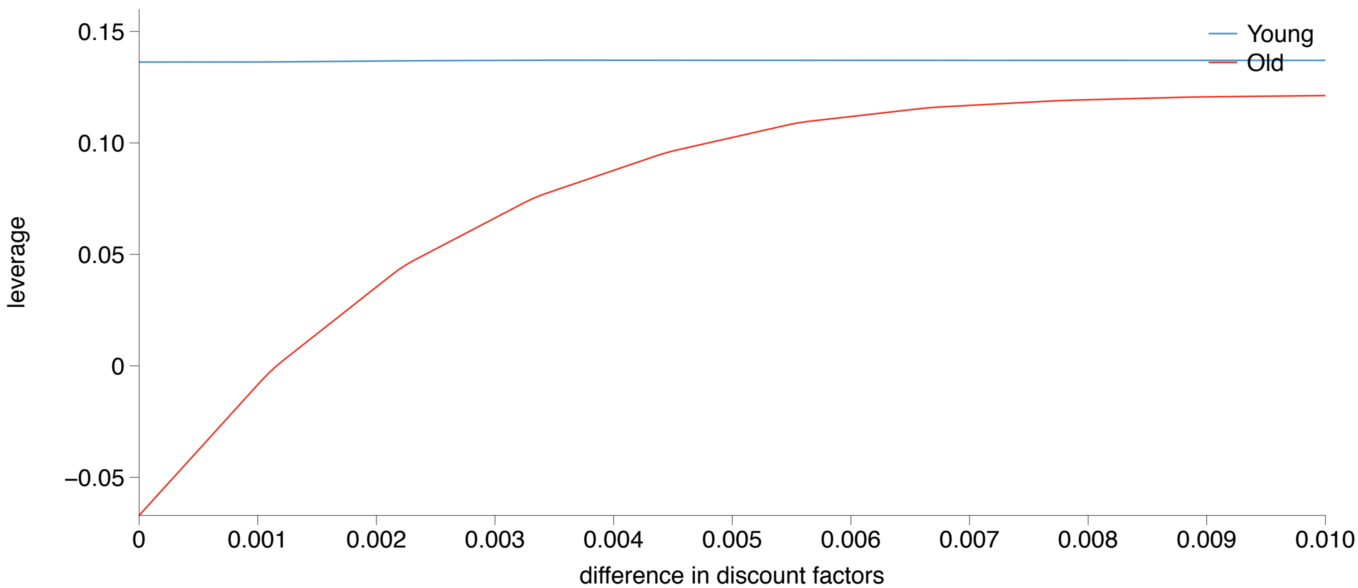

This figure is constructed as follows. We pick a grids for $\theta$, the fraction of the capital stock lost to the firm when it repudiates a contract, $\tau_c$, the corporate tax rate, and $\beta_C - \beta$, the difference in the discount factors between borrows and lenders. We then solve the model for each of the different parameter values, simulate the model for 50 time periods, and then plot average investment. We use a parameterization that is the average of the parameter estimates 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: Counterfactuals: Dividends




This figure is constructed as follows. We pick a grids 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 dividends. We use a parameterization that is the average of the parameter estimates 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 7: Counterfactuals: Difference in Discount Factors




This figure is constructed as follows. We pick a grid for $\beta_C - \beta$, the difference between the lender's and the firm's discount factors. 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 that is the average of the parameter estimates from Table 1. "Young" firms are those in the first 10 periods of life, and "old" firms are those in periods 11 to 50.