A Mechanism Design Model of Firm Dynamics: The Case of Limited Commitment

Hengjie Ai, Dana Kiku, and Rui Li *

November 2012

We present a general equilibrium model with two-sided limited commitment that accounts for the observed heterogeneity in firms' investment, payout and CEO-compensation policies. In the model, shareholders cannot commit to holding negative net present value projects, and managers cannot commit to compensation plans that yield life-time utility lower than their outside options. Firms operate identical constant return to scale technologies with i.i.d. productivity growth. Consistent with the data, the model endogenously generates a power law in firm size and a power law in CEO compensation. We also show that the model is able to quantitatively explain the observed negative relationship between firms' investment rates and size, the positive relationship between firms' size and their dividend and CEO payouts, as well as variation of firms' investment and payout policies across both size and age.

*Ai (hengjie@umn.edu) is affiliated with the Carlson School of Management, University of Minnesota; Kiku (kiku@wharton.upenn.edu) is at the Wharton School, University of Pennsylvania; and Li (li1220@purdue.edu) is at the Economics Department of Purdue University. We would like to thank Zhiguo He, Erzo Luttmer, Adriano Rampini, Vish Vishwanathan, and seminar participants at the 2012 European Summer Symposium of Financial Markets at Gerzensee and the Finance Department of University of Wisconsin at Madison for their helpful comments.

Introduction

It has been documented in the literature that the distribution of firm size and the distribution of CEO compensation obey power law. It is also well known that investment and dividend policies of firms depend significantly on firms' size. Small firms invest dis-proportionally more and pay out less compared with large firms. What is the economic mechanism that accounts for both the observed heterogeneity in firms' policies and the cross-sectional distribution of firm size and CEO pay? As we show, the standard neoclassical model with no contracting frictions is able to explain the power law of firm size. It is, however, inconsistent with other stylized features of the data. To address this issue, we develop a general equilibrium model with heterogenous firms and limited commitment that can jointly account for the cross-sectional distribution of firms' size, investment, CEO compensation and payout policies. We take a mechanism design approach and explore the implications of the constrained efficient allocation subject to limited commitment.

The key elements of our model are: constant return to scale technology, i.i.d. productivity growth, and two-sided limited commitment. We assume that shareholders cannot commit to negative net present value projects, and that managers cannot commit to wage contracts that result in life-time utility lower than their outside option. Under the optimal contract, CEO compensation takes the following simple form: it stays constant most of the time, rises after a sequence of good productivity shocks, and shrinks after a sequence of negative productivity shocks. A series of positive productivity shocks raises the value of the manager's outside option and forces shareholders to raise CEO wage to retain the manager. Thus, manager compensation increases whenever his participation constraint binds. A sequence of negative productivity growth rates lowers the value of the firm. To prevent bankruptcy, manager's wage has to drop whenever the value of the firm approaches zero. Hence, our model generates a positive relationship between CEO pay and firm size observed in the data.

Our model is also able to endogenously generate a power law in firm size and CEO compensation. Given that technology is constant return to scale, a sequence of positive productivity shocks increases the size of a firm unboundedly, which results in a fat tail of firm-size distribution. Because managers' outside option rises with firm size, their compensation under the optimal contract has to rise proportionally. Consequently, the power law in firm size translates into a power law in CEO pay.

We show further that our model predicts an inverse relationship between investment rate and firm size, and a positive relationship between dividend payout and firm size. Small firms in our model are those that have recently experienced a sequence of negative productivity shocks. As the size of a firm shrinks, shareholders' commitment constraint is likely to bind.

A binding constraint destroys complete risk sharing and is welfare reducing. Constrained efficiency requires these firms to increase investment to avoid further downsizing. Conversely, after a series of positive productivity shocks, a firm grows and so does the outside option of the manager. The manager's participation constraint binds whenever the value of her outside option equals the value of the compensation contract. To reduce the likelihood of a binding constraint, it is optimal for large firms to downsize by reducing investment. As a result, small firms in our model invest more and grow faster than large firms. By the same logic, small firms have low dividend yields as they have to spend most of their resources on funding investment. Both implications are consistent with the observed cross-sectional patterns in firms' investment and dividend choices.

We calibrate our model to match standard macroeconomic moments and volatility of output at the firm level and show that it can quantitatively account for the key moments of the joint distribution of firms' size, investment, payout and CEO-compensation policies observed in the data. We also show that, despite its simplicity, our model has rich implications for investment and payout behavior conditional on both firm size and age, and explains a significant amount of the cross-sectional variation in firms' decisions conditional on the two characteristics.

We show that both types of limited commitment, on the shareholder side and on the manager side, are important for understanding empirical relationships among CEO compensation, firms' investment and size. To highlight their importance, we first discuss the standard neoclassical model without contracting frictions. Because managers are risk averse and shareholders are well diversified, the optimal contract in this framework features complete risk sharing and a constant manager compensation. Due to convex adjustment costs, all firms here have the same investment-to-capital ratio and identical expected growth rates. Hence, this is a model where Gibrat (1931)'s law holds and the distribution of firm size obeys power law. However, it also implies a zero correlation between CEO pay and firm size and rules out any dependence of firm growth rate on size. Modeling limited commitment on the shareholder side provides a theory for endogenous bankruptcy and generates an inverse relationship between investment and size. However, as in the frictionless case, risk sharing implies that CEO compensation never rises under the optimal contract, and consequently there is no power law in CEO pay. We demonstrate how our model with two-sided limited commitment improves upon the above models and explains important stylized features of the cross-sectional data.

Our paper builds on the large literature on limited commitment and its implications for firm behavior. Early contributions include Kehoe and Levine (1993), Kocherlakota (1996) and Kiyotaki and Moore (1997). Albuquerque and Hopenhayn (2004) provide a theoretical foundation for limited commitment models of firm dynamics. More recently, Lorenzoni and Walentin (2007) study the implications of limited commitment on the investment-Q relationship. Rampini and Viswanathan (2010, 2012) focus on firms' risk management and capital structure decisions. Lustig, Syverson, and Van Nieuwerburgh (2011) consider a model with one-sided limited commitment and study the link between the inequality of CEO compensation and productivity growth. Our model differs from the above literature in several respects. We use continuous time method to characterize the solution to the optimal contract and the cross-sectional distribution of firms as ordinary differential equations, which allows for sharper analytical results and efficient numerical solutions. We solve the mechanism-design problem with two-sided limited commitment in a general-equilibrium setting. Other models typically focus on limited commitment on the agent side only. In addition, none of above mentioned papers attempts to explain the power law in firm size and CEO compensation and their interaction.[1] More generally, we confront our model with a comprehensive set of cross-sectional characteristics summarized in Section I.

The continuous-time methodology of this paper builds on the fast growing literature of continuous time dynamic contracting, for example, Sannikov (2008), DeMarzo and Sannikov (2006), DeMarzo, Fishman, He, and Wang (2009), He (2009), He (2011), Biais, Mariotti, and Villeneuve (2010). For an excellent survey of this literature, see Biais, Mariotti, Plantin, and Rochet (2004). The solution of the optimal dynamic contract in our paper is based directly on Ai and Li (2012a), who analyze the optimal contract with two-sided limited commitment in a model similar to ours but allow managers to have stochastic differential utility (Duffie and Epstein (1992)).

Our paper is also related to the literature on power law in firm size and CEO compensation. Gabaix (2009) surveys power law in economics and finance. Recent literature on firm dynamics and power law is reviewed in Luttmer (2010). Luttmer (2007) proposes a general equilibrium model where firms' growth rate is i.i.d. and the equilibrium size distribution obeys power law. The neoclassical model without frictions considered in this paper is essentially an interpretation of Luttmer (2007) with neoclassical production technology. Tervio (2003) and Gabaix and Landier (2008) are assortative matching models that link CEO compensation to firm size.[2] Our model provides an alternative, mechanism-design based explanation of the level of CEO pay and its dependence on firm size. Tervio (2003) and Gabaix and Landier (2008) study CEO compensation taking size distribution of firms as given. In our model, both the distribution of firm size and CEO compensation are endogenous outcomes of the optimal dynamic contract. An additional advantage of our

[1] Lustig, Syverson, and Van Nieuwerburgh (2011) is an exception. Their model also produces a power law for the distribution of firm size.

[2] For a survey on the literature of the economics of super stars, see Gabaix and Landier (2008).

dynamic model is that it can be used to study the cross-sectional distribution as well as the life-cycle dynamics of firms' investment, CEO compensation and dividend payout policies.

The rest of the paper is organized as follows. In Section I, we summarize the key stylized features of the joint empirical distribution of firms' size, age, investment, dividend payout and CEO compensation policies. In Section II, we consider a frictionless Arrow-Debreu economy and discuss its implications. We augment the baseline model with limited commitment on the shareholder side in Section III and further extend it to the case of two-sided limited commitment on both the principle and the agent side in Section IV. We demonstrate how these frictions improve upon the basic neoclassical model. Section V evaluates the quantitative implications of our model with two-sided limited commitment against the set of stylized facts documented in Section I. Section VI concludes.

I Stylized Facts

In this section, we summarize some stylized features of firms' investment, payout, and CEO compensation policies and their variation with firms size and age. We will discuss this empirical evidence in greater detail in Section V below. The first five facts describe the distribution of firm size and CEO compensation and reveal the effect of size on firms' policies and firms' survival.

1. Firm size is characterized by a power-law distribution with a slope coefficient close to 1.1. The distribution of CEO compensation is also well approximated by a power law with a somewhat larger slope coefficient of about 1.7.

2. The elasticity of CEO pay with respect to firm size is close to 1/3. The elasticity is larger for firms in the left and right tail of the size distribution, and smaller for medium-sized firms.

3. Small firms have higher investment rates than large firms.[3] The average investment rate in our sample is about 10% and is almost the same for firms in the top ten-percentile of the size distribution. Small firms (those in the bottom decile) have an average investment-to-capital ratio of about 17%.

4. Small firms are much less likely to make dividend and/or interest payments than large firms. In the bottom size decile, on average, only one out of ten firms have non-zero

[3]We define investment rate as a ratio of firm investment in period t to its (gross) stock of capital at the end of $t-1$.

payouts. The fraction of dividend- and/or interest-paying firms increases to more than 80% in the right tail of firm size distribution.

5. Small firms are more likely to become bankrupt than larger firms.

The next set of facts summarizes variation of firm policies across both size and age. The numbers reported below correspond to 3×3 double-sorted portfolios.

6. Controlling for age, firms' investment rate decreases with size, and controlling for size, investment rate decreases in age. Overall, investment-to-capital ratio of young small is almost 4 times higher than that of old large firms.

7. Controlling for age, CEO compensation increases in size, and controlling for size, CEO compensation decreases with age.

8. The ratio of CEO compensation to firm size is decreasing with size and age after controlling for the other characteristic.

9. Controlling for age (size), dividend and payout yields are increasing with firm size (age). The average yield of old large firms is about 5 times higher than that of young small firms.

We use this empirical evidence as guidance in developing our theoretical model. In the next sections, we evaluate the qualitative implications of the frictionless model (Section II), the model with one-sided limited commitment (Section III), and the model with two-sided limited commitment (Section IV) against stylized facts 1-5. In Section V, we provide a formal calibration of our model with two-sided limited commitment and compare its quantitative implications with all nine empirical features. The data description and further discussion of empirical evidence are provided in Section V and the Appendix.

II An Arrow-Debreu Economy

A Setup of the Model

A.1 Preferences

We consider a continuous time infinite horizon economy with two types of agents, shareholders and managers. The representative shareholder is infinitely lived and her preference is represented by a time additive constant relative risk aversion utility:

$$E\left[\int_0^\infty e^{-\beta t}\frac{1}{1-\gamma}\mathbf{C}_t^{1-\gamma}dt\right], \tag{1}$$

where $\beta > 0$ is the time discount rate, and $\gamma > 0$ is the relative risk aversion coefficient. $\mathbf{C}_t$ denotes consumption flow rate of the shareholder at time t. Managers value consumption streams using the same preferences with identical risk aversion and time discount rate.[4,5]

A.2 Production Technology

Production in this economy is processed at a continuum of locations indexed by $j \in \mathcal{J}$, where $\mathcal{J}$ is the set of all possible locations. At location j, general output is produced using capital and labor though a Cobb-Douglas technology:

$$y_{j,t} = K_{j,t}^{\alpha}\left(\mathbf{z}_t N_{j,t}\right)^{1-\alpha},$$

where $y_{j,t}$ denotes the output, $K_{j,t}$ is the amount of capital and $N_{j,t}$ is the amount of labor hired at location j at time t. $\mathbf{z}_t$ is the labor-augmenting productivity. We set $\mathbf{z}_t = \mathbf{z}$ to be constant to save notation, but allow for aggregate productivity growth in our calibration. The representative shareholder owns all the capital and supply one unit of labor inelastically per unit of time. General output can be used for consumption by either the shareholder or the manager. However, only managers have access to the technology that transforms general output into new capital goods.

Labor market is competitive. Let $\mathbf{W}_t$ denote the real wage at time t and $\Pi_{j,t}$ denote the equilibrium payment to capital at location j at time t. Our convention is to use bold face letters to denote aggregate quantities. We have:

$$\Pi_{j,t} = \Pi\left(K_{j,t}\right) = \max_{N_{j,t}}\left\{\mathbf{z}_t K_{j,t}^{\alpha} N_{j,t}^{1-\alpha} - \mathbf{W}_t N_{j,t}\right\}. \tag{2}$$

We call $\Pi(K)$ the operating profit function. Because the technology is constant return to scale, and labor market is competitive, the operating profit function is linear: $\Pi(K) = \mathbf{A}K$, where $\mathbf{A}$ is the economy-wide (equilibrium) marginal product of capital.

The manager hired at location j has access to a technology that accumulates capital

[4] Our model can be easily extended to incorporate the case where shareholders and managers have different time discount rate and/or different risk aversion parameters. We do not entertain these extensions to maintain parsimony in our quantitative exercise.

[5] We refer to the shareholder as she and the manager as he in the rest of the paper.

according to the following law of motion:

$$dK_{j,t} = (I_{j,t} - \delta K_{j,t})\, dt + K_{j,t}\sigma dB_{j,t},$$

where $\delta > 0$ is the instantaneous depreciation rate of capital. The standard Brownian motion, $B_{j,t}$, is i.i.d. across locations and represents productivity shocks to the capital accumulation technology.[6] The term $I_{j,t}$ is investment made at time t in location j. Investing I at a location with total capital stock K costs general output $h\left(\frac{I}{K}\right)K$, where

$$h(i) = 1 + h_0 i^2$$

is a strictly convex adjustment cost function.

A.3 Entry and Exit of Firms

A unit measure of managers arrive at the economy per unit of time. Upon arrival, a manager is endowed with an outside option that delivers life-time utility $\bar{U}$.[7] Operating a technology at a given location requires managers, who are the only agents that have access to the capital accumulation technology. A manager who chooses to operate a production technology for the shareholder must give up his outside option permanently.

The shareholder offers a contract to the manager upon his arrival. A contract is a plan for investment, managerial compensation, and dividend payout as a function of the entire history of the economy. A firm is a contractual relationship between the manager and the shareholder organized for production at a particular location. We let $V(K,U)$ denote the value of a firm with total initial capital stock K and the manager's promised utility U.[8] Creating a firm of size K requires a total cost of $H(K)$ in terms of current period consumption goods, where $H(\cdot)$ is a strictly increasing and a strictly convex cost function. At every point in time, the shareholder chooses the initial capital stock of the new generation of firms, K^*, and the promised utility to the manager, U^*, optimally to maximize profit:

$$(K^*, U^*) \in \arg\max_{K, U \geq \bar{U}} \{V(K,U) - H(K)\}. \tag{3}$$

[6]We show in the Appendix of the paper that $K_{j,t}$ can be interpreted as the product of location specific productivity and location specific capital. In this case, Brownin motion $B_{j,t}$ can be interpreted as a combination of productivity shocks and capital depreciation shocks.

[7]For simplicity, we do not explicitly specify the technology that delivers the reservation utility. The outside option is never taken under our assumptions.

[8]Because there is no aggregate uncertainty, constrained efficient allocations in the economies considered later in the paper can be achieved by policies that depend only on two state variables (K,U). In the equilibrium implementation of the efficient allocations, firm value depends only on (K,U) without loss of generality.

Managers are also subject to random health shocks that follow a Poisson process with intensity κ. Once hit by a health shock, the manager exits the economy and all capital accumulated by the manager evaporates. Health shocks are i.i.d. across managers.

A.4 Equilibrium

In the economy with perfect commitment on financial contracts considered here, standard welfare theorems apply and the competitive equilibrium implements Pareto efficient allocations. To incorporate cases with limited commitment, we describe a general notion of equilibrium that provides a unified framework for us to discuss the frictionless case, as well as cases with various forms of limited enforcement. In the Appendix, we show that the equilibrium allocation is, in fact, constrained efficient subject to the frictions of limited commitment.

We use $\mathbf{r}$ to denote the equilibrium real interest rate. In our economy, at any point in time t, a new generation of firms are created. Let $C_{j,s}^t$, $I_{j,s}^t$, and $D_{j,s}^t$ denote the managerial compensation, investment, and dividend payout policy, respectively, for generation-t firm at location j at time s. An equilibrium allocation must specify the managerial compensation, investment, and dividend payout policies for firms of all generations at all times, $\left\{\left[\left(\hat{C}_{j,s}^t, \hat{I}_{j,s}^t, \hat{D}_{j,s}^t\right)_{s=t}^{\infty}\right]_{t=0}^{\infty}\right\}_{j\in\mathcal{J}}$. Taking equilibrium interest rate as given, the policy of firm j of generation t, denoted $\left\{\hat{C}_{j,s}^t, \hat{I}_{j,s}^t, \hat{D}_{j,s}^t\right\}_{s=t}^{\infty}$, maximizes the present value of the firm subject to feasibility constraints:

$$\left\{\hat{C}_{j,s}^t, \hat{I}_{j,s}^t, \hat{D}_{j,s}^t\right\}_{s=t}^{\infty} \in \arg\max E_t\left[\int_t^{\tau} e^{-\mathbf{r}(s-t)} D_s ds\right] \tag{4}$$

$$\textit{subject to}: \ \{C_s, I_s, D_s\}_{s=t}^{\infty} \in \Omega\left(K_{j,t}^t, U_{j,t}^t\right), \tag{5}$$

where $\Omega(K,U)$ denotes the set of feasible allocations given initial condition (K,U), and τ is the stopping time at which the manager of firm j is hit by the Poisson health shock.

In the case of perfect commitment, given the initial condition $\left(K_{j,t}^t, U_{j,t}^t\right)$, feasibility requires that $\{C_s, I_s, D_s\}_{s=t}^{\infty}$ satisfy the following resource constraint:

$$C_s + \phi(I_s, K_s) + D_s = \Pi(K_s), \quad \textit{all} \ \ s \geq t, \tag{6}$$

the law of motion of capital:

$$dK_s = K_s\left[\left(\frac{I_s}{K_s} - \delta\right) ds + \sigma dB_{j,s}\right], \quad s \geq t, \quad \textit{and } K_t = K_{j,t}^t, \tag{7}$$

and the "promise keeping" constraint for the entrepreneur at time t:

$$\left\{E\left[\int_t^{\tau} e^{-\beta(s-t)}\left(\beta+\kappa\right)C_s^{1-\gamma}ds\right]\right\}^{\frac{1}{1-\gamma}} \geq U_{j,t}^t, \tag{8}$$

Formally, $\Omega\left(K_{j,t}^t, U_{j,t}^t\right)$ is the set of allocations $\{C_s, I_s, D_s\}_{s=t}^{\infty}$ such that $\{C_s, I_s, D_s\}_{s=t}^{\infty}$ is adapted to the Brownian filtration generated by $\{B_{j,s}\}_{s=t}^{\infty}$, and $\{C_s, I_s, D_s\}_{s=t}^{\infty}$ satisfies conditions (6)-(7).[9] In what follows, we suppress the subscript j to save notation whenever there is no confusion. We use $V(K,U)$ to denote the value function of the optimization problem in Equation (4) subject to feasibility constraints.

A competitive equilibrium must specify the path of interest rates, $\{\mathbf{r}_t\}_{t\geq 0}$, and wages, $\{\mathbf{W}_t\}_{t\geq 0}$, consumption of the representative shareholder, $\{\mathbf{C}_t\}_{t\geq 0}$, consumption, investment, and dividend payout policies for all firms. In general, allocations are history dependent. We focus our attention on the stationary equilibrium where the exit rate of firms equals the entry rate, and the cross-section distribution of firm characteristic is time-invariant.[10] In this case, equilibrium allocations can be achieved by allocation rules (Atkeson and Lucas (1992))) that specify allocations as functions of a pair of state variables (K,U), the total capital stock of the firm and the continuation utility promised to the manager. Below we provide a definition of the equilibrium using allocation rules.[11]

An allocation rule consists of functions,

$$C(K,U),\ I(K,U),\ D(K,U),\ N(K,U),\ G(K,U),$$

that map the state space into the real line. Given the allocation rules, allocations can be constructed using a two-step procedure. First, for each firm of type (K,U), $\{C(K,U),\ I(K,U),\ D(K,U),\ N(K,U)\}$ specify the flow rate of manager's consumption, investment, dividend payout and amount of labor hired in the current instant. Next, the law of motion of the state variables is constructed from the allocation rule using:

$$dK = K\left[\left(\frac{I(K,U)}{K}-\delta\right)dt+\sigma dB\right], \tag{9}$$

and

$$dU = \left[-\frac{\beta+\kappa}{1-\gamma}\left(C^{1-\gamma}U^{\gamma}-U\right)+\frac{1}{2}\gamma\frac{G(K,U)^2}{U}\right]dt+G(K,U)\,dB, \tag{10}$$

[9] Technically, $\{C_s, I_s, D_s\}_{s=t}^{\infty}$ also need to satisfy certain integrability conditions to ensure that the relevant stochastic integrals are well defined.

[10] We prove the existence of such an equilibrium by construction.

[11] There is little need in using the construction of allocation rules in the frictionless economy here. We nevertheless use this formulation to facilitate comparison across economies.

where Equation (10) is the stochastic differential utility representation of the manager's preference (Ai and Li (2012a)).

Formally, the equilibrium consists of interest rate, $\mathbf{r}$, real wage, $\mathbf{W}$, allocation rules, $\{C(K,U),\ I(K,U),\ D(K,U),\ N(K,U),\ G(K,U)\}$, consumption of the representative shareholder, $\mathbf{C}$, and the cross-section distribution of types, $\Phi(K,U)$, such that:[12]

1. Taking interest rates as given, the allocation constructed from the allocation rules described above solves the firm's inter-temporal maximization problem in Equation (4).

2. The initial choice of (K^*, U^*) solves the maximization problem in Equation (3) for all firms.

3. Taking real wages as given, $N_{j,s}^t$ constructed from allocation rules solves the intra-temporal profit maximization problem in Equation (2) for all (j,t) and all $s \geq t$.

4. The representative shareholder chooses consumption, investment in creating new firms, and investment and payout policies in existing firms to maximize utility in Equation (1).

5. Goods market clears:
$$\mathbf{C} + \int \left[C(K,U) + h(I(K,U),K)\right] d\Phi(K,U) + H(K^*) = \int K^\alpha (\mathbf{z}N)^{1-\alpha} d\Phi(K,U). \tag{11}$$

6. Labor market clears:
$$\int N(K,U)\, d\Phi(K,U) = 1.$$

7. The cross sectional distribution of types, $\Phi(K,U)$, is consistent with the law of motion of (K,U) implied by the allocation rules, as in Equations (9) and (10).[13]

B Firm Dynamics and the Cross Section

Because there is no contracting friction, given the initial condition (K,U), the maximization problem in Equation (4) can be solved in two steps. First, choose the optimal investment

[12] Here we conjecture and later on verify that $\mathbf{r}_t$, $\mathbf{W}_t$, $\mathbf{C}_t$ and $\mathbf{m}_t$ are constant in the stationary equilibrium. $\mathbf{W}_t$ and $\mathbf{C}_t$ will be time-dependent but grow at a constant rate in our calibration as we allow for aggregate productivity growth.

[13] Technically, $\Phi(K,U)$ must satisfy a version of the Komogorov forward equation as we show in the Appendix.

policy to maximize the total value of the firm:

$$\max_{\{I_s\}_{s=t}^{\infty}} E_t \left[\int_t^{\tau} e^{-\mathbf{r}(\mathbf{s}-t)} \left[\mathbf{A} K_s - h\left(I_s, K_s\right) \right] ds \right] \tag{12}$$

$$\textit{subject to}: \; dK_s = K_s \left[\left(\frac{I_s}{K_s} - \delta \right) ds + \sigma dB_s \right], \quad s \geq t,$$

$$K_t = K$$

Second, choose a compensation policy to deliver the promised utility U in a way that minimizes cost:

$$\min_{\{C_s\}_{s=t}^{\infty}} E \left[\int_t^{\tau} e^{-\mathbf{r}(s-t)} C_s ds \right] \tag{13}$$

$$\textit{subject to}: \; \left\{ E \left[\int_t^{\tau} e^{-\beta(s-t)} \left(\beta + \kappa\right) C_s^{1-\gamma} ds \right] \right\}^{\frac{1}{1-\gamma}} \geq U.$$

Note that Equation (12) is the standard profit maximization problem with neoclassical technology as in Hayashi (1982). The solution to (13) is also straightforward: risk aversion of the manager and the condition $\mathbf{r} = \beta$ imply that the optimal policy satisfies:

$$C_t = U. \tag{14}$$

It is convenient to denote

$$\hat{r} = \kappa + \mathbf{r} + \delta.$$

The solution to the firm's problem is summarized in the following proposition.

Proposition 1. *The First-Best Case*

Suppose

$$0 < A - \hat{r} < \frac{1}{2} h_0 \hat{r}^2, \tag{15}$$

then the value of a firm with initial capital stock K and promised utility U is given by

$$V\left(K, U\right) = \bar{v} K - \frac{1}{\mathbf{r} + \kappa} U, \tag{16}$$

where the constant $\bar{v} = h'\left(\hat{\imath}\right)$ and $\hat{\imath}$ is the optimal investment-to-capital ratio given by:

$$\hat{\imath} = \arg\max_i \frac{\mathbf{A} - h\left(i\right)}{\hat{r} - i} = \hat{r} - \sqrt{\hat{r}^2 - \frac{2}{h_0}\left(\mathbf{A} - \hat{r}\right)} \in \left(0, \hat{r}\right). \tag{17}$$

Proof. See the Appendix. □

Equation (16) has an intuitive interpretation. The term $\bar{v}K = h'(\hat{\imath})K$ is the firm value in the neoclassical model with capital adjustment cost (for example, Hayashi (1982)), and $\frac{1}{\mathbf{r}+\kappa}U$ is the present value of manager's compensation. Perfect risk sharing implies that managerial consumption is constant (see Equation (14)). Therefore, the present value of managerial compensation is simply given by the Gordon (1959)'s formula.

Note that the value function $V(K,U)$ is strictly decreasing in U; therefore the optimal choice of initial utility promised to the manager in Equation (3) is $\bar{U}$. The optimal choice of the initial capital stock, K^*, is given by:

$$K^* = \arg\max_K \left\{ \bar{v}K - \frac{1}{\mathbf{r}+\kappa}\bar{U} - H(K) \right\}. \tag{18}$$

For a given equilibrium marginal product of capital $\mathbf{A}$, Equation (16) determines firms' value function, and equation (18) determines the initial size of all firms. Equation (17) implies that the investment-to-capital ratio is constant across all firms. As a result, Gibrat's law holds: firm growth rate is i.i.d. and does not depend on size. We assume that when indifferent, managers choose to give up the outside option and work for the firm. In this case, a unit measure of firms will be created per unit of time. We can solve for the cross-section distribution of firm size in closed form as in Luttmer (2007).

Proposition 2. *Power Law of Firm Size*

Given K^* and $\hat{\imath}$, the total measure of firm is $\frac{1}{\kappa}$ and the total amount of capital stock is

$$\mathbf{K} = \frac{K^*}{\kappa + \delta - \hat{\imath}}. \tag{19}$$

Furthermore, the distribution of firm size is given by:

$$\phi(K) = \begin{cases} \frac{1}{\sqrt{\left(\hat{\imath}-\delta-\frac{1}{2}\sigma^2\right)^2+2\kappa\sigma^2}} K^{*-\alpha_2} K^{\alpha_2-1} & K \geq K^* \\ \frac{1}{\sqrt{\left(\hat{\imath}-\delta-\frac{1}{2}\sigma^2\right)^2+2\kappa\sigma^2}} K^{*-\alpha_1} K^{\alpha_1-1} & K < K^*, \end{cases}$$

where $\alpha_1 > \alpha_2$ are the two roots of the quadratic equation

$$\kappa + \left(\hat{\imath} - \delta - \frac{1}{2}\sigma^2\right)\alpha - \frac{1}{2}\alpha^2\sigma^2 = 0.$$

In particular, the right tail of firm size obeys power law with exponent α_2.

Proof. See the Appendix. □

For a given marginal product of capital, $\mathbf{A}$, we can solve for the total capital stock of the economy, $\mathbf{K}$, using Equations (17) and (19). Because total labor supply is normalized to 1, we must have $\mathbf{A} = \alpha \left(\frac{\mathbf{z}}{\mathbf{K}}\right)^{1-\alpha}$, which completely determines the equilibrium.

Several implications of the above model are worth attention. First, the model generates a power law distribution of firm size. The average investment rate in COMPUSTAT data is about 12%. Firm death rate is about 4% per year, and volatility of sale growth is around 40% per year. With $\delta = 9\%$, which implies a total depreciation rate of capital of 13% per year, the implied exponent of the tail slope of the power law is 1.09, which is fairly close to the empirical evidence we presented in first section of the paper.

Second, the model implies a flat investment-size relationship and a flat CEO pay-size relationship. Equation (14) implies that managerial compensation of all firms is identical and equals $\bar{U}$. Equation (17) implies that investment rates of all firms are identical as well. These features of the model are grossly inconsistent with the data.

Third, the model implies an inverse relationship between dividend payout and firm size, qualitatively consistent with the stylized fact 4 documented in Section I of the paper. Because investment rate is constant across firms, $C_t + D_t = \mathbf{A}K_t - h(\hat{\imath})\,K_t$ is proportional to the size of the firm. Because $C_t = \bar{U}$ is a constant, $\frac{D_t}{K_t} = \mathbf{A} - h(\hat{\imath}) - \frac{\bar{U}}{K_t}$ must increase with K_t. As a result, small firms pay less dividends than larger firms. This feature is consistent across all models we study in the paper.

Fourth, these is no endogenous bankruptcy in the model. The death rate of firms is constant and identical across firms of all sizes and ages.

C Normalized Continuation Utility

To facilitate comparison across models with different commitment frictions, it is useful to specify value functions and policy functions in terms of normalized utility. Given policy $\left\{\hat{C}_t, \hat{I}_t, \hat{D}_t\right\}_{t\geq 0}$, define

$$U_t = \left\{E\left[\int_t^\tau e^{-\beta(s-t)}\left(\beta+\kappa\right)\hat{C}_s^{1-\gamma}ds\right]\right\}^{\frac{1}{1-\gamma}}, \tag{20}$$

as the continuation utility of the manager at time t. Let $u_t = \frac{U_t}{K_t}$ denote the normalized utility. In all models considered in the paper, given the equilibrium interest rate, the firm's objective function is linear and the feasibility constraint is homogenous of degree one in the state variable K. As a result, the value function satisfies

$$V(K, U) = v(u) K, \tag{21}$$

and policy functions satisfy

$$C(K, U) = c(u) K; \qquad I(K, U) = i(u) K. \tag{22}$$

for some $v(\cdot)$, $c(\cdot)$, and $i(\cdot)$. We will call $v(\cdot)$ the normalized value function, and $c(\cdot)$ and $i(\cdot)$ the normalized policy functions.

In the first-best case discussed above, $V\left(K, \bar{U}\right) = \bar{v}K - \frac{1}{\mathbf{r}+\kappa}\bar{U}$ and

$$v(u) = \bar{v} - \frac{1}{\mathbf{r}+\kappa} u$$

is linear in u.

Figure 1 plots the normalized value function of the firm. As shown in the figure, the normalized value function, $v(u)$, is linear with a negative slope $\frac{1}{\mathbf{r}+\kappa}$. Note that the continuation utility promised to the manager, $U_t = \bar{U}$, is constant due to perfect risk sharing. Therefore, as the size of the firm grows larger, $K \to \infty$, the normalized utility $u = \frac{\bar{U}}{K} \to 0$, and $v(u) \to \bar{v}$. In this case, the present value of managerial compensation as a fraction of the total value of the firm converges to zero. The ratio of the total value of the firm to the total capital stock converges to the average Q in neoclassical models:

$$\lim_{K\to\infty} \frac{V\left(K, \bar{U}\right)}{K} = \bar{v} = h'(\hat{\imath}).$$

Alternatively, a sequence of negative shocks moves K_t towards zero and $u_t = \frac{\bar{U}}{K_t}$ increases without bound. At $u^* = (\mathbf{r}+\kappa)\bar{v}$, firm value becomes zero. A further decrease in K_t moves the firm value into the negative region: $\bar{v} - \frac{1}{\mathbf{r}+\kappa}u < 0$. Intuitively, optimal risk sharing implies that the compensation to manager must be constant, $\bar{U}$. A sequence of negative shocks lowers the cash flow of the firm. The value of the firm becomes negative when the present value of cash flow is lower than the present value of future compensation promised to the manager. We view this as another counter-factual implication of the model. Below, we first consider the case in which the shareholder cannot commit to compensation plans that yield negative firm value at any point in time. As we show, in this case, firm value can never be negative, which provides a micro-foundation for bankruptcy and limited liability.

III One-Sided Limited Commitment

In this section, we consider the case where the shareholder cannot commit to negative net present value projects. In this case, in addition to Equations (6)-(8), feasibility also requires policy $\{C_s, I_s, D_s\}_{s=t}^{\infty}$ to satisfy

$$E_u\left[\int_u^{\tau} e^{-\mathbf{r}(s-u)} D_s ds\right] \geq 0 \quad for\ all\ u \geq t. \tag{23}$$

The firm's maximization problem in this case differs from that in Equation (4) because of the constraint in Equation (23). That is, the shareholder is no longer allowed to choose from all forms of compensation contracts. Those contracts that render firm value negative in some future states is no longer implementable due to the lack of commitment technology on the shareholder side. Because the lack of commitment restricts the set of feasible contracts, everything else being equal, the value of the firm will be lower than that in the frictionless economy.

As in the frictionless case, the value function and policy functions satisfy the homogeneity properties given in Equations (21) and (22). Given the equilibrium marginal product of capital, $\mathbf{A}$, the normalized value function, $v(u)$, can be characterized as the solution to an ordinary differential equation, which can be found in the Appendix. The properties of the value function and policy functions are characterized by the following proposition.

Proposition 3. *One-Sided Limited Commitment*

1. *The normalized value function $v(u)$ is strictly decreasing and strictly concave with a bounded domain, $(0, u_{MAX}]$.*
2. *Under the optimal contract, the normalized utility u moves to the interior with probability one on the right boundary, u_{MAX}.*
3. *Under the optimal contract, u is decreasing in productivity shocks.*
4. *Managerial compensation, $c(u_t)K_t$ is constant as long as $u_t < u_{MAX}$. In addition, $\lim_{u\to 0} c(u) = u$.*
5. *The optimal investment rate, $i(u)$, is a strictly increasing function of u. Also, $\lim_{u\to 0} i(u) = \hat{\imath}$, where $\hat{\imath}$ is the optimal investment level in the friction-less case.*
6. *$\lim_{t\to\infty} u_t = 0$ with probability one. $\lim_{u\to 0} v(u) = \bar{v}$, where $\bar{v}$ is given in Equation (16).*

Proof. Ai and Li (2012a) □

Parts 1 and 2 of the above proposition imply that the support of the normalized utility in the one-sided limited commitment case is bounded: $(0, u_{MAX}]$. Whenever u_t hits u_{MAX} from the left, it will come back to the interior with probability one. Note that u_{MAX} is the maximimum amount of (normalized) utility that can be delivered to the manager without rendering the value of the firm negative. From the social planner's point of view, risk sharing is strictly welfare improving. Therefore, efficiency precludes negative firm values, which would result in the shareholder's abandoning the project and terminating the risk sharing contract. In what follows, we will call the maximum normalized utility under the optimal contract, u_{MAX}, the bankruptcy point. At u_{MAX}, because u_t cannot increase further, a negative shock that lowers K_t must be associated with a one-to-one drop in U_t. Figure 2 plots the normalized value function, $v(u)$, for the one-sided limited enforcement case (dashed line) and that for the frictionless case (dash-dotted line) assuming the same marginal product of capital, $\mathbf{A}$.[14] Note that firm value in the one-sided limited commitment case, in general, is lower than that in the frictionless case because of imperfect risk sharing, especially when u is close to u_{MAX}, where the value of the firm hits zero and risk sharing is poor.

Note that $u_t = \frac{U_t}{K_t}$ depends both on the promised utility U_t and the size of the firm. We can intuitively think of u as a measure of the manager's equity share in the firm. A higher u implies that a larger fraction of firm's cash flow will be used to compensate the manager to deliver the promised utility. In the frictionless economy, optimal risk sharing implies that $U_t = \bar{U}$ for all t; therefore changes in u_t are completely due to changes in the size of the firm. In the case of one-sided limited commitment, complete risk sharing is no longer feasible, and U_t increases with K_t in general. Part 3 of the above proposition implies that the optimal contract in the one-sided limited commitment case nevertheless preserves some basic features of the first best case, namely, continuation utility is less sensitive to productivity shocks than firm size. A positive productivity shock increases K_t and U_t at the same time, but U_t increases less than proportionally so that the net effect is that u_t decreases. If we interpret u_t as the manager's equity share in the firm, then our model implies manager's equity share is inversely related to firm size. As positive productivity shock increases firm size and lowers manager's equity share at the same time. This implication holds for all models we consider and is qualitatively consistent with the empirical evidence discussed in Section I.

Part 4 of the above proposition implies that manager compensation is constant whenever the bankruptcy constraint is not binding. In Figure 3, we plot the sample path of a firm with

[14] Note that our comparison between the first best case and the case with one-sided limited commitment here is a partial equilibrium one. In general equilibrium, fixing the preference and technology parameters of the model, adding one-sided limited commitment will result in an endogenous change in the steady-state capital stock of the economy and therefore a different marginal product of capital.

u close to the bankruptcy point, u_{MAX}. The top panel in Figure 3 is the trajectory of the log size of the firm, $\ln K_t$, and the second panel is the path of the normalized utility, or the manager's equity share in the firm, u_t. The third panel is the corresponding realizations of the value of the firm, $V(K_t, U_t)$, and the bottom panel shows the log managerial compensation, $\ln C_t$. At time 0, the firm starts from the interior of the normalized utility space, $u_0 < u_{MAX}$. A sequence of negative productivity shocks from time 0 to time 2 lowers the capital stock of the firm (top panel). For $t < 1$, $u_t < u_{MAX}$ is in the interior (second panel). In this region, firm value is strictly positive (third panel) and managerial compensation is constant (bottom panel). At $t = 1$, u_t hits the boundary u_{MAX} and cannot increase further despite subsequent negative productivity shocks. For $t \in (1, 2)$, the firm continues to receive a sequence of negative productivity shocks and the total capital stock of the firm shrinks (top panel); however, u_t stay at u_{MAX}, as shown in the second panel of Figure 3. In this case, the firm value remains at zero and do not cross over the negative region due to a reduction in managerial compensation: managerial compensation keeps decreasing until the firm starts experiencing positive productivity shocks at time $t = 2$. From time $t = 2$ to $t = 3$, the firm experiences a sequence of positive productivity shocks followed by a sequence of negative productivity shocks. As a result, firm value bounces back to the positive region and decreases afterwards (third panel). Because the normalized utility u_t stays in the interior before $t = 3$ (second panel), managerial consumption stays constant (bottom panel), although at a lower level than C^*. At time $t = 3$ the size of the firm hits its previous running minimum, and u_t reaches u_{MAX} again. As before, firm value stays at zero, and managerial consumption keep decreasing, until the firm starts to receive positive productivity shocks for the next time.

Let $C^* = C(K^*, U^*)$ be the managerial compensation for a new entrant firm. The above analysis implies the C_t will stay at C^* until the firm hits the bankruptcy constraint, in which case C_t drops below C^*. As a results, in the stationary equilibrium, managers of firms who have not hit the bankruptcy point will stay at C^* and managers of firms who have experienced bankruptcy will be below C^*. No manager's compensation is above C^*.

The point u_{MAX} can be interpreted as the bankruptcy state of the firm. As shown in Ai and Li (2012a), the equilibrium allocation can be implemented by the following compensation contract. The contract promises a constant wage to the manager, C^* in the example in Figure 3. At the same time, the shareholder is given a default option. The default option allows the shareholder to reset the wage contract at a lower level. However, exercise of the default option also triggers bankruptcy, in which case the shareholder is no longer entitled to any cash flow from the firm. In the case of bankruptcy, the asset of the firm is liquidated: an independent trustee sells the asset of the firm on a competitive market and pays off the manager's wage at the lower reset rate. Our model therefore provides a microfoundation for bankruptcy through optimal mechanism design. Note that firms that are close to the bankruptcy point,

u_{MAX}, are those that experienced a sequence of negative productivity shocks. As a result, our model with one-sided limited commitment implies that small firms are more likely to become bankrupt, qualitatively consistent with empirical evidence discussed in Section. In fact, this is a feature shared by both the current model and the model with two-sided limited commitment, which we study in the next section.

Part 5 of Proposition 3 implies that investment is an increasing function of manager's equity share, u. We plot the investment-to-capital ratio, $i(u) = \frac{I(K,U)}{K}$ as a function of the manager's normalized utility, u, in Figure 4. Note that investment rate is a constant in the frictionless economy but increases in u in the case of one-sided limited commitment. The intuition for this result is that as u increases, the manager's equity share becomes larger, and the firm gets closer to the bankruptcy point, u_{MAX}. u_{MAX} is associated with inefficient risk sharing, and therefore it is in the interest of both parties to avoid it. High investment increases the size of the firm, lowers the manager's equity share and pushes the firm away from the bankruptcy point, u_{MAX}. As we note in part 3) of the proposition, firms close to $u = 0$ are large firms who experienced a sequence of positive productivity shocks and firms close to $u = u_{MAX}$ are small because of negative productivity shocks. As a result, our model implies that small firms's investment rate is higher than that of large firms: small firms are riskier from the manager's perspective, and optimal risk sharing requires higher investment rate and faster growth. This is another feature that is qualitatively consistent with the empirical evidence (stylized fact 3 stated in Section I) and is shared by models with one-sided as well as two-sided limited commitment.

As the size of firm increases, $K_t \to \infty$ and $u_t = \frac{U_t}{K_t} \to 0$. The optimal investment rate converges to the first best level. In this case, the probability of bankruptcy is small and both investment and compensation policy converge to the first best case.

Part 6 of the proposition implies that the firm will eventually grow out of the constraint in the long-run and converge to the frictionless case. On average, investment is higher than depreciation and the size of firms, K_t grows. In fact, conditioning on survival, $K_t \to \infty$ with probability one. By part 3) of the proposition, under the optimal contract, U_t increases at a lower rate than K_t. As $K_t \to \infty$ and $u_t = \frac{U_t}{K_t} \to 0$, the optimal policies converge to those in the first best case.

The last part of Proposition 3 has strong implications for the cross-section distribution of firms. First, small firms on average invest at a higher rate than large firms, because they are typically closer to the bankruptcy point. Second, investment policy and, therefore, the growth rate of large firms converge to those in the first best case. In particular, although expected growth rates of large firms are smaller than of small firms, they remain strictly positive. This feature of the model produces a power law distribution in the right tail similar to that in the first best case. Third, CEO compensation of most firms are identical, in particular, there is no power law is CEO pay. Note that CEO compensation only changes in the bankruptcy state. Because firms on average grow, most of them do not go through bankruptcy, and their CEO compensation is constant over time.

We plot the counter-cumulative distribution function of CEO compensation for the one-sided limited commitment case (dashed line) in Figure 5, where we use the calibrated parameter values in Section V. In the same figure, we also plot the empirical complementary cumulative distribution function for all CEOs with available data in 1996 (dotted line). The horizontal axis is log-equally-spaced CEO compensation. We scale the CEO compensation in the model so that the median CEO compensation in the model matches the median CEO compensation in the data. The vertical axis is the rank of CEO compensation (log equally spaced). We normalize rank by the total number of firms in the model and in the data, respectively, so that the vertical axis has the interpretation of a probability.[15]

Note that the right tail of the empirical complementary cumulative distribution is well approximated by a power law. We highlight the top 300 highest paid CEO in the right tail of the distribution with plus signs. We also plot the estimated power law for the right tail (dark dotted line) using the estimate discussed below. The right tail of the CEO compensation produced by the one-sided limited commitment model is trivial: the top 82% of the highest paid CEO have identical compensation level C^*. As a result, the elasticity of CEO pay with respect to firm size is very small in the one-sided limited commitment case, of about 0.06 under the calibrated parameters.

To summarize, one-sided limited commitment improves on the frictionless model and generates several additional features that are qualitatively consistent with the stylized facts we document in Section I, for example, the inverse relationship between investment rate and firm size and the positive relationship between CEO compensation and firm size. Importantly, it provides a theory for endogenous bankruptcy and is consistent with the fact that small firms become bankrupt more often than large firms. However, there is no power law in CEO compensation, and the elasticity of CEO pay with respect to firm size is close to zero. We now turn to the model with two-sided limited enforcement.

IV Two-Sided Limited Commitment

In this section, we introduce an additional friction into our model. Following Kehoe and Levine (1993), Kiyotaki and Moore (1997), and Albuquerque and Hopenhayn (2004), we

[15]Plotted this way, a linear counter-cumulative distribution function is the defining characteristic of a power law distribution.

assume that the manager has an option to default and cannot commit to compensation contracts that yield life-time utility lower than that provided by the default option. Upon default, the manager can retain a fraction θ of the capital stock and hire labor on a competitive market to produce output. However, he is forever excluded from the credit market. That is, he can only consume the operating profit from capital stock he possesses after the default, but cannot enter into any intertemporal risk sharing contract. Due to homogeneity of the utility function, the utility that the manager receives by taking the default option is of the form $u_{MIN}K_t$ for some parameter u_{MIN}, which is a function of θ. The expression for u_{MIN} is given in the Appendix.

In this case, limited commitment on the manager side further restricts the set of feasible allocations. In addition to Equations (6)-(8), and (23), feasibility also requires that the continuation utility provided by the policy $\{C_s, I_s, D_s\}_{s=t}^{\infty}$ is higher than that associated with the default option at all times and in all states of the world:

$$\left\{E_u\left[\int_u^{\tau} e^{-\beta(s-u)}\left(\beta+\kappa\right)C_s^{1-\gamma}ds\right]\right\}^{\frac{1}{1-\gamma}} \geq u_{MIN}K_u \quad for \quad u \geq t. \tag{24}$$

In what follows, we will call both Equations (23) and (24) commitment constraints. To distinguish the commitment constraint on the shareholder side and that on the manager side, we will call Equation (23) the bankruptcy constraint and Equation (24) the participation constraint.

While the limited commitment constraint on the shareholder side affects mainly the properties of the optimal compensation contract for small firms that are close to bankruptcy, the impact of limited commitment on the manager side primarily changes the optimal contract for large firms, where the value of managers' outside option is high. As the size of the firm grows, the right hand side of the inequality (24) increases. To discourage the manager from default, compensation must rise. The limited commitment on manager side, therefore, creates a mechanism where CEO compensation increases with firm size, and potentially allows our model to generate a power law in CEO compensation. The properties of the optimal compensation contract are discussed in the following proposition. Again, we use the homogeneity property of the value function, Equation (21), and that of the policy functions, Equation (22), and focus on the normalized value function and policy functions.

Proposition 4. *Two-Sided Limited Commitment*

1. *The normalized value function $v(u)$ is strictly decreasing and strictly concave.*

2. *Under the optimal contract, the normalized utility u moves to the interior with probability one on the boundaries, u_{MIN} and u_{MAX}.*

3. Under the optimal contract, u is decreasing in productivity shocks.

4. Managerial compensation, $c(u_t) K_t$ is constant as long as $u \in (u_{MIN}, u_{MAX})$.

5. The optimal investment rate, $i(u)$, is a strictly increasing function of u.

Proof. See Ai and Li (2012a) □

Note that the normalized utility under the optimal contract must stay in the interval $[u_{MIN}, u_{MAX}]$. In Figure 6, we plot the value function for the frictionless model (dash-dotted line), that for the case of one-sided commitment (dashed line), and that for the case of two sided limited commitment (solid line). All value functions are computed assuming the same marginal product of capital, $\mathbf{A}$. As we add more contracting frictions, the value of the firm becomes lower. Note that limited commitment on the manager side prevents the normalized continuation utility from approaching zero as in the frictionless case and the case with one-sided limited commitment. u_t must stay to the left of u_{MIN}. Also the maximum level of normalized utility u_{MAX} is lower than that in the case of one-sided limited commitment. The lack of commitment on the manager side reduces efficiency and limits the set of feasible continuation utilities that can be supported under the optimal contract.

Part 4 of the proposition implies that the optimal compensation contract inherits some properties from the one sided limited commitment case. In particular, managerial compensation is constant whenever none of the commitment constraints binds. Compensation increases only when the participation constraint binds, and drops only when the bankruptcy constraint binds. A sequence of positive productivity shocks raises the value of the outside option of the manager, and eventually results in a pay raise as u_t hits u_{MIN}. A sequence of negative productivity shocks pushes the value of the firm toward zero, and eventually results in a pay reduction at $u_t = u_{MAX}$ because of the bankruptcy constraint.

The properties of the optimal compensation contract on the right boundary, u_{MAX}, is the same as those in the case of one sided limited commitment. We plot a sample path of a firm with u_0 close to the left boundary, u_{MIN}, in Figure 7. The top panel in Figure 7 is the realization of the log size of the firm, $\ln K_t$. The second panel is the path of the normalized utility, or the manager's equity share in the firm, u_t. The third panel is the trajectory of the normalized value of the firm, $v(u_t)$, and the bottom panel is that of the log managerial compensation, $\ln C_t$. At time 0, the firm starts from the interior of the normalized utility space, $u_{MIN} < u_0 < u_{MAX}$. A sequence of positive productivity shocks from time 0 to 0.5 increases the capital stock of the firm (top panel). For $t < 0.5$, $u_t > u_{MIN}$ is in the interior (second panel) and manager's consumption is constant (bottom panel). During this period, both the size of the firm, K_t and the size-normalized firm value, $v(u_t)$, increases.

At time 0.5, the normalized continuation utility, u_t, reaches the left boundary, u_{MIN}, and the participation constraint binds. Further realizations of positive productivity shocks from $t = 0.5$ to $t = 1$ translate directly into increases in managerial compensation (bottom panel), but the normalized continuation utility (second panel), and the normalized firm value (third panel) remain constant. At time $t = 1$, the firm starts to experience a sequence of negative productivity shocks. As a result, the size of the firm shrinks, and the normalize utility $u_t = \frac{U_t}{K_t}$ increases because risk sharing implies that the continuation utility U_t is less sensitive to shocks than K_t (part 3 of Proposition 4). During the period $t \in (1, 3)$, u_t stays in the interior of $[u_{MIN}, u_{MAX}]$ and manager's consumption stays constant. At time $t = 2$, the firm starts to receive a sequence of positive productivity shocks. During this period, u_t stays in the interior of its domain until the size the firm, K_t, reaches its previous running maximum at $t = 5$, in which case, the participation constraint starts to bind again, and manager's consumption increases as a result (bottom panel).

We plot the optimal investment policy, $i(u)$, for the first best case (dash-dotted line), that for the case with one-sided limited commitment (dashed line) and that for the case of two-sided limited commitment (solid line) in Figure 8. Note that investment rate in both the one-sided commitment and the two sided commitment cases is an increasing function of u. However, in the one-sided limited commitment case, as $u \to 0$, $i(u) \to \hat{\imath}$, which is the optimal investment level in the frictionless economy. In the two-sided limited commitment case, on the other hand, u is bounded by the manager's outside option, u_{MIN}. Moreover, as u decreases, investment rate falls below the first best level, $\hat{\imath}$. It is optimal to invest at a level lower than $\hat{\imath}$ because u_{MAX} is associated with inefficient risk sharing and therefore is welfare reducing. As the manager's equity share in the firm, u_t, is low close to u_{MIN}, it is optimal to reduce $u_t = \frac{U_t}{K_t}$ in the future to lower the probability of reaching the inefficient risk sharing point at u_{MIN}. Reducing investment and therefore the size of the firm tends to move u_t to the interior and away from the binding participation constraint.

We illustrate the dynamics of the state variable, u_t, under the optimal contract in Figure 9. The top panel of Figure 9 is the expected change, or the drift coefficient of $\ln u$ as a function of u. The bottom panel is volatility, or the diffusion coefficient of $\ln u$ as a function of u. The point $\bar{u}$ represents the average of normalized continuation utility across all firms in the stationary distribution. Note that under the optimal policy, the drift of u reaches its maximum at u_{MIN}, and stays positive in the region close to u_{MAX}, indicating a tendency for u to return to it steady state mean, $\bar{u}$, when manager's equity share is small. Similarly, the drift of u achieves its minimum at u_{MIN} and remains negative in regions close to u_{MIN}. This implies that a tendency to converge to $\bar{u}$ when manager's equity share in the firm is large, and a binding bankruptcy constraint is likely. The pattern of the drift of $\ln u$ reveals that under the optimal policy, u is a mean reverting process, and firms tend to accumulate in the area close to $\bar{u}$ in the long run. Note also that mean reversion is stronger in the region close to u_{MAX} compared to that in areas close to u_{MIN}. The reason is that under the first best investment policy, firms on average grow and $\ln u$ has a strictly negative drift. This effect reinforces mean reversion when u is close to u_{MAX}, and offsets the positive drift for small values of u. These features of the dynamics of the state variable have important implications for the dependence of firm policies on age.

The diffusion coefficient of $\ln u$ on the Brownian motion dB_t is negative. Note that $u_t = \frac{U_t}{K_t}$ and the Brownian productivity shock dB_t directly affects K_t. Therefore the diffusion coefficient being negative is the consequence of risk sharing: a positive productivity shock raises K_t and U_t simultaneously, but U_t increases by less than proportionally, and as a result, the ratio u_t responds negatively to dB_t. By comparison, perfect risk sharing implies a diffusion coefficient of -1 and no risk sharing at all corresponds to a diffusion coefficient of 0. Note that diffusion of u_t tends to zero on the boundaries u_{MIN} and u_{MAX}, indicating that risk sharing is poor when constraints are binding and that normalized utility returns to the interior with probability one on the boundaries.

The behavior of large firms in the economy with two sided limited commitment is quite different from that in the case of one-sided limited commitment. In the case of one sided limited commitment, firms on average grow, and as K_t gets larger, $u_t = \frac{U_t}{K_t} \to 0$. Conditioning on survival, firms grow and converge to the region where u_t is close to zero. They have similar investment policies, and risk sharing is perfect, unless, with extremely small probability, they are driven to the bankruptcy point after a long series of negative productivity shocks. Perfect risk sharing implies that manager compensation is constant and identical in most of the firms in the economy, as shown in Figure 5. In the model with two-sided limited commitment, managerial consumption cannot stay constant with the size the firm growing indefinitely — it has to increase whenever the participation constraint binds. Therefore managerial compensation has to increase unboundedly as the size of the firm does so. As we show in Section V of the paper, this feature of our model allows the power law of firm size to translate into the power law of CEO compensation. In addition, as u_t approaches u_{MIN}, investment rate falls, and this creates a tendency for the normalized utility to revert back to the interior. As a result, u_t does not converge in the long run, and this feature of the model creates heterogeneity in firms' investment and payout policies, allowing our model to match many features of the cross-sectional distribution as we show below.

V Quantitative Results

A Aggregation and General Equilibrium

For a given equilibrium marginal product of capital, $\mathbf{A}$, Sections II, III and IV describe the solution to the optimal contracting problem for a typical firm. General equilibrium requires that the marginal product of capital is consistent with the total capital stock of the economy and productivity $\mathbf{z}$. In this section, we briefly describe a procedure how to solve for the equilibrium marginal product of capital from the optimal policies and the market clearing condition.

In general, the market clearing condition in Equation (11) requires solving a two dimensional distribution $\Phi(u, K)$. Following Ai (2012), and Ai and Li (2012b) we explore homogeneity of the decision rules and define the one dimensional measure

$$m(u) = \int K d\Phi(u, K).$$

As shown in Ai (2012), the market clearing condition in Equation (11) can be expressed using the one-dimensional measure m. In addition, $m(u)$ obeys a version of the Komogorov forward equation that can be solved efficiently numerically. We present the forward equation in the Appendix. We follow Ai (2012) and express the market clearing condition as

$$\mathbf{C} + H(K^*) + \int_{u_{MIN}}^{u_{MAX}} [c(u) + i(u)]\, m(u)\, du = \int_{u_{MIN}}^{u_{MAX}} \mathbf{z} N(u)^{1-\alpha} m(u)\, du. \tag{25}$$

We use the following iterative procedure to solve for the general equilibrium of our model.

1. Step 1: Starting from an initial guess of the marginal production of capital $\mathbf{A}$, we solve for the optimal contract and allocation rules. Numerically, we use the Markov chain approximation method (Kushner and Dupuis (2001)) described in the Appendix to solve the optimal control problem.

2. Step 2: After obtaining the policy functions $c(u), i(u), g(u)$, we use the forward equation given in the Appendix to construct measure $m(u)$.

3. Step 3: We use measure m and Equation (25) to calculate the total capital stock in the steady state for the given $\mathbf{A}$.

4. Step 4: We verify that $\mathbf{A}$ is the marginal product of capital. If $\mathbf{A} > (<) \alpha z \mathbf{K}^{\alpha-1}$, we choose a smaller (larger) $\mathbf{A}$ and resolve the contracting problem by repeating the above steps. We iterate on this procedure until convergence, that is, until $\mathbf{A} = \alpha z \mathbf{K}^{\alpha-1}$.

B Choice of Parameter Values

We calibrate our model to match standard macroeconomic moments and evaluate its performance along the dimensions stated in Section I. We focus on the model with two-sided limited commitment, which we call the benchmark model hereafter.

We set the shareholder's and the manager's discount rate at 0.02, which together with a two percent annual growth rate of productivity $\mathbf{z}_t$ implies a risk-free interest rate of 4% per year. We calibrate the risk-free interest rate at 4% because our model does not have aggregate uncertainty and therefore risk premium. We set the risk-free rate in our model to equal to the average of the equity return and bond return in the US data. We choose risk aversion of 2.

In terms of the technology parameters, we set $\sigma = 36\%$, this allows our model to match the average volatility of annual sale growth rate of firms in our sample. We set $\kappa = 4\%$, which implies a 4% annual rate of firm death, consistent with the same moment in the COMPUSTAT data set we use. We choose the capital depreciation rate $\delta = 0.09$, along with $\kappa = 4\%$, this implies a 13% total depreciation rate of physical capital, as typical in neoclassical macro models. Finally, we choose $\phi = 5$ as the adjustment cost parameter. Our calibration implies an aggregate investment-output ratio of 23%, which roughly matches the corresponding moment in the US data in the post war period.

We set $\theta = 1$ to maintain the parsimony of the model. This specification corresponds to the specification of the autarky as the outside option as in Kehoe and Levine (1993).

After numerically solving the model, we simulate it out for 300 years, and discard the first 100 years data. Our model is continuous time, and all quantities are aggregated at the annual level. We have two million firms in the cross-section. This allows us to obtain very tight estimates of various quantities of interest within our model.

C Power Law in Firm Size and CEO Compensation

We first evaluate the predictions of our benchmark model on the power law of firm size and the power law of CEO compensation.

Following Luttmer (2007) and Gabaix (2009), we use the following definition and parametrization of power law. A distribution of a random variable X obeys power law if it has density of the form

$$f(x) = k\zeta x^{-(1+\zeta)},$$

for some constants, k, $\zeta > 0$. The complementary cumulative distribution function of X is

$$P(X > x) = kx^{-\zeta}.$$

The parameter ζ is called the power law exponent of X.

Our estimates of the power law parameter for firm size and CEO compensation are presented in Table 1. We rely on the number of employees to measure firm size, and estimate the slope of the right tail using data on the top 300 firms in a given year. We find strong evidence of power low in both firms size and CEO compensation, with an average power-law coefficient of 1.1 and 1.7 for size and CE pay, respectively. Our empirical evidence on size distribution is consistent with Luttmer (2007), who reports a power-law exponent of 1.07 using U.S. Census data with firm size also measured by the total number of employees, and Gabaix and Landier (2008), who report estimates close to one using firms' market value.

Our calibrated model implies a power-law distribution of firm size with an exponent of 1.08, which is close to the average estimate in the data. The log-log plot of the right tail of the distribution of firm size implied by our model and observed in the data is presented in Figure 10. The horizontal axis in the figure represents firm size and the vertical axis shows the complementary cumulative distribution function (both are equally-spaced on the log scale). We show the right tail of size distribution for three representative years in our sample – 1996, 2000 and 2006. As the figure shows, the slope of the right tail of size distribution implied by the model matches well the slope observed in the data.

The power law coefficient of the right tail of CEO compensation generated by our model is almost identical to that of firm size, i.e., about 1.08. Hence, our model somewhat overstates the fat tail of the distribution of CEO compensation relative to the data. We illustrate the fit of the model in Figure 11 by presenting the complementary cumulative distribution function of CEO pay, side by size, in the data and in the model. Similar to the previous picture, we define the right tail by the 300 highest paid CEOs and show three years of data. As the figure shows, the right tail of the empirical and model-implied distribution of CEO compensation is fairly liner on a log-log scale and conforms well to power law.

D CEO Compensation Conditional on Size and Age

It has been documented in the literature that CEO compensation increases with firm size. Table 2 shows the elasticity of CEO pay with respect to firm size for several alternative measures of firm size: the number of employees, market capitalization, capital stock and book value of assets. The average elasticity is estimated at about 0.32 and the point estimates are largely similar across different measures of size. Others report similar (of around 1/3) estimates (see Gabaix (2009) for a survey of the literature).

The average elasticity of CEO compensation to firm size across simulations of our model is 0.24 with a mean R^2 of 0.3. Although our model understates somewhat the elasticity relative to the data, the model-implied magnitude is quite reasonable given that limited enforcement is the only mechanism that generates a positive relationship between size and CEO compensation in the model. We expect a richer model, especially the one that also incorporates moral hazard, to produce a higher elasticity and improve on this dimension of the data.

The elasticity of CEO compensation with respect to firm size is not homogenous across different size groups. Figure 12 shows the elasticity of CEO compensation to firm size conditional on size in the data (Panel A) and in the model (Panel B). We consider 5 size-sorted portfolios and use different measures of size: market capitalization (stars), the number of employees (squares), capital stocks (circles) and book value of assets (triangles). As the figure shows, the elasticity is V-shaped for all measures of size. That is, CEO compensation is more sensitive to firm size for firms in the left and right tails of size distribution. This feature is, in fact, the signature characteristic of the two-sided limited commitment model. In the model, CEO compensation changes with firm size either because the bankruptcy constraint is binding, or because the participation constraint is binding. Because small firms tend to be in region close to the bankruptcy point, and firms with a binding participation constraint are typically large firms, CEO compensation is more sensitive to firm size for very small and very large firms. The model-implied effect of size on CEO pay–size relationship is illustrated bottom panel of Figure 12. As the plot shows, the elasticity in the model is close to zero for medium sized firms, and close to one for firms in the top and bottom size quintile.

The level of CEO compensation depends not only firm size, but also on age. In Table 5, we report the median level of log CEO compensation for size and age sorted portfolios. We construct portfolios my sorting firms into three size groups, and dividing each size bin into three age-sorted portfolios. In this and related tables, size is measured by the number of firm employees. The table shows that controlling for age, the level of CEO compensation increases in size, and conditioning on size, the CEO pay decreases with age. With some exceptions, the model is able to account for the observed cross-sectional variation of CEO compensation.

In the model, the level of CEO compensation increases in firm size due to limited commitment constraints. This pattern remains even after controlling for firm age. Conditioning on being in the same age group, large firms are firms that have experienced a sequence of positive productivity shocks. High productivity shocks raise the value of the manager's outside option and send the firm to the region, where participation constraint is likely to be binding. As a result, the manager gets a larger share of firms' earnings. Similarly, small firms are those that have experienced a series of negative productivity shocks and are typically close to the bankruptcy region. It is, therefore, optimal to reduce CEO wage and redirect firm resources towards investment to avoid further downsizing of the firm. In this region, risk sharing is poor and manager bears a lot of idiosyncratic risk because of a looming binding constraint.

Conditional on size, CEO compensation in the model decreases in age for small firms, is flat for medium sized firms, and is slightly increasing in age for large firms. To understand this pattern, it is important to note that CEO compensation in the model is history dependent and that firms, on average, grow.[16] Consider two small firms of different age: young and old. To stay as small as the young firm, the old one must have received a lot of negative productivity shocks in the past that have offset its trend growth. This means that the bankruptcy constraint is likely to have bound for many more times for the old firm relative to the young one. As a result, CEO compensation in small-old firms is typically lower than that of small-young firms. The situation is reversed for the large size group. Unconditionally, older firms tend to be much larger due to the deterministic trend in growth. Among large firms, younger ones have not had enough time to move to the participation constraint as often an older ones. Consequently, for large firms, CEO compensation is increasing in age.

Panel A of Table 3 shows variation of CEO pay-to-capital ratio across firms of different size and age. We find that controlling for age, compensation ratio shrinks with firm size. Consistent with the data, the model-implied ratio of CEO pay to capital decreases with size for all age groups (see Panel B of the table). This is the implication of (imperfect) risk sharing under the optimal contract. Risk sharing reduces sensitivity of CEO compensation to productivity shocks. Thus, as firms become large, CEO compensation as a fraction of firm size declines.

Table 3 further shows that, empirically, conditioning on size, CEO pay-to-capital ratio decreases with age. Similarly, compensation ratio in our model declines with age for small and medium-sized firms, and is relatively flat across large firms. For small firms, age is slightly negatively correlated with size because, as discussed above, small older firms happen to hit the bankruptcy constraint more times than small younger firms. As a result, in the cross-section of small firms, the pattern in CEO-to-capital ratio across age mimics and even magnifies the corresponding pattern in CEO pay. As firm size increases, age becomes positively correlated with size due to productivity trend growth. This effect is particularly important for medium-size firms. For them, the level of CEO pay varies little with age, and the cross-sectional pattern in CEO compensation-to-capital is driven mostly by size. Therefore, for medium-sized firms, compensation ratio features a decline in age. For the

[16]Under our calibration, the average annual investment rate is about 15%. Given a depreciate rate of 9%, this implies a growth rate of roughly 6% for an average firm.

large cohort, a positive correlation between size and age becomes even stronger. An increase in size across age almost entirely offsets an increase in the level of CEO compensation, making compensation ratio similar for young and old firms.

E Investment Conditional on Firm Size and Age

Panel A of Figure 13 shows that, in the data, small firms invest at a higher rate than large firms. The figure presents investment-to-capital ratios for 25 portfolios sorted on size. The average investment rate of the smallest percentile is about 18%, while the top portfolio has an average investment-to-capital ratio of only 10%. This evidence is consistent with empirical findings of the literature on firm growth and size. Evans (1987) and Hall (1987) document a negative relationship between firm growth and firm size. In our model, firms' expected growth is a strictly increasing function of investment rate, while realized growth rates depend on both investment rate and productivity shocks. Therefore, differences in investment rates across firms translate directly into differences in average growth rates. One advantage of comparing investment rates rather than growth rates in the model and the data is that the measurement of investment rates is not subject to selection biases induced by negative productivity shocks. That is, observed realized growth rates of small firms may be high (even if expected growth rates are independent of firm size) because small firms hit by negative productivity shocks are more likely to drop out of the sample than larger firms are.

The investment-size relationship implied by our model is plotted in the bottom panel of Figure 13. As discussed above, small firms in the model invest much more than large firms because managers' equity share is higher in small firms and investment policy $i(u)$ is increasing in u.

Table 5 characterizes the empirical distribution of investment rates across both size and age. Controlling for age, investment rate in the data declines with size. Similarly, conditioning on size, average investment-to-capital ratio decreases with age. Quantitatively, per each unit of capital, small young firms invest almost four times more relative to large old firms. This evidence is consistent with finding in Davis, Haltiwanger, and Schuh (1996) and Cooley and Quadrini (2001), who show that, controlling for the other characteristic, firm growth is negatively correlated with size and age.

As Panel B of Table 5 shows, the quantitative implications of our model for the cross-sectional distribution of investment rates across firm size and age are largely consistent with the pattern documented in the data, except for firms in the large size portfolio. The model-implied investment rates of large firms are generally independent of age. Heterogeneity in firms' investment policy in the same age group is informative about the basic properties of the optimal contract in our model. Note that firms' realized growth has two components: a (locally) deterministic component that is due to investment, and a stochastic component that is due to productivity shocks. Firms in the same age group are similar in terms of the deterministic growth component, and they differ in size mainly due to different realizations of unexpected productivity shocks. Controlling for age, large firms are those that have received a sequence of positive productivity shocks and are close to the participation constraint, u_{MIN}. As a result, investment rate decreases with size for all age cohorts. The size-driven dispersion in investment rates in our model matches almost exactly the pattern in the data.

Because investment rate is a function of the normalized utility, u, the dependence of investment rate on age within different size groups is due to the mean reversion of the state variable. For small firms, younger firms are closer to the right boundary u_{MAX} (see Figure 9). Older firms tend to be closer to the middle ($\bar{u}$) because of strong mean reversion of the normalized utility at the boundary. Hence, for small firms, investment rate is decreasing with age. In the large size group, younger firms on average are closer to the left boundary u_{MIN}, where investment rate is low. Older firms due to mean-reversion tend to be in the interior and, thus, have somewhat higher investment rates. Because mean-reversion is much stronger at the right boundary u_{MAX} than at u_{MIN}, the decrease in investment rate with age is much stronger for small and medium-sized firms relative to the increase across large size portfolios. In fact, as Table 5 shows, in our simulations, for large firms, investment rate virtually does not vary with age.

F Payout Policy Conditional on Firm Size and Age

It has been well documented in the literature that small firms are less likely to pay out dividends compared with large firms. Panel A of Figure 14 confirms this evidence by plotting the fraction of dividend- and/or interest-paying firms across 25 size sorted portfolios. As the figure shows, the vast majority of small firms makes neither type of payments. In fact, less than 10% of firms in the bottom size percentile pay dividends or make any interest payments. In contrast, nine out of ten firms in the top size portfolio pay either dividend or interest, or both. The fraction of non-zero paying firms increases monotonically across size-sorted portfolios. The corresponding model output is presented in Panel B. While the model-implied shape is somewhat different, overall, the model captures well the cross-sectional increase observed in the data. Small firms in our model invest at a higher rate and tend not to pay out dividends. In contrast, large firms invest less and pay out more frequently (most of the time).

Similarly, the model is able to account for a strong positive correlation between dividend yields and firms' size: in the data, the average dividend yield of large firms is about five times higher than that of small firms. In the model, dividend yields are also monotonically increasing in size. In fact, the net payout of small firms is negative as they try to raise additional funds to accelerate investment and growth. We also find that the model is generally consistent with the observed heterogeneity in dividend yields across double-sorted portfolios. In the data, conditional on size, older firms pay out more than younger ones; and the increase in dividend yields across size persists after controlling for firms' age. Except for the very large firms that show no dependence on age, the model also implies an increase in dividend yields with size and age, controlling for the other characteristic.

VI Conclusion

We present a mechanism design model of firm dynamics. We start with a friction-less model with Arrow-Debreu contracts and illustrate how different forms of limited commitment on compensation contracts help explaining a wide range of empirical regularities in firms' investment, CEO compensation and dividend payout policies. We show that a simple model with two-sided limited commitment is consistent with key cross-sectional characteristics of firms' behavior.

Our goal is to build on the recent developments in continuous time contracting theory to develop a quantitative framework for firms using a mechanism design approach. Closing the model in general equilibrium allows us to use empirical evidence from the cross-section to discipline our dynamic model. Our model has predictions on both the time-series and the cross-sectional distribution of firms' decision that could be confronted with the data. We view limited commitment as the first step in building contracting frictions into dynamic general equilibrium models with heterogeneous firms. There are several aspects of our model that require improvement. At the moment, the model overstates the fat tail of CEO compensation, and it predicts zero pay-performance sensitivity for mid-sized firms. We believe that other frictions such as moral hazard and adverse selection could potentially help better align predictions of our model with the data. These are promising directions for future research.

References

Ai, Henhjie, 2012, A Note on Aggregation in Economies with Homogenous Decision Rules, *Working paper, University of Minnesota.*

Ai, Hengjie, and Rui Li, 2012a, Limited Enforcement, Recursive Preference, and Implications on Executive Compensation, *Working paper, University of Minnesota.*

Ai, Hengjie, and Rui Li, 2012b, Moral Hazard, Investment and Firm Dynamics, *Working paper, University of Minnesota.*

Albuquerque, Rui, and Hugo Hopenhayn, 2004, Optimal Lending Contracts and Firm Dynamics, *Review of Economic Studies* 71, 285–315.

Atkeson, Andrew, and Robert Lucas, 1992, On Efficient Distribution with Private Information, *Review of Economic Studies* 59, 427–453.

Biais, Bruno, Thomas Mariotti, Cuillaume Plantin, and Jean-Charles Rochet, 2004, Dynamic Security Design, *Working Paper, IDEI.*

Biais, Bruno, Thomas Mariotti, and Jean-Charles Rochet Stéphane Villeneuve, 2010, Large Risks, Limited Liability, and Dynamic Moral Hazard, *Econometrica* 78, 73–118.

Cooley, Thomas, and Vincenzo Quadrini, 2001, Financial Markets and Firm Dynamics, *American Economic Review* 91, 1286–1310.

Davis, Steven, John Haltiwanger, and Scott Schuh, 1996, *Davis.* (Cambridge: MIT press).

DeMarzo, Peter, Michael Fishman, Zhiguo He, and Neng Wang, 2009, Dynamic Agency and q Theory of Investment, *Journal of Finance, Forthcoming.*

DeMarzo, Peter, 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 Larry Epstein, 1992, Stochastic Differential Utility, *Econometrica* 60, 353–394.

Evans, David, 1987, The Relationship between Firm Growth, Size, and Age: Estimates for 100 Manufacturing Industries, *Journal of Industrial Economics* 35, 567–581.

Gabaix, Xavier, 2009, Power Laws in Economics and Finance, *Annual Review of Economics* 1, 255–293.

Gabaix, Xavier, and Augustin Landier, 2008, Why Has CEO Pay Increased So Much?, *Quarterly Journal of Economics* 123, 49–100.

Gibrat, Robert, 1931, *Les Inegalites Economiques.* (Paris, France: Librairie du Recueil Sirey, 1931).

Hall, Bronwyn, 1987, The reslationship between firm size and firm growth in the US manufacturing sector, *The journal of industrial economics* 35, 583–606.

Hayashi, Fumio, 1982, Tobin's Marginal q and Average q: A Neoclassical Interpretation, *Econometrica* 50, 213–224.

He, Zhiguo, 2009, Optimal Executive Compensation when Firm Size Follows Geometric Brownian Motion, *Review of Financial Studies* 22, 859–892.

He, Zhiguo, 2011, A Model of Dynamic Compensation and Capital Structure, *Journal of Financial Economics* 100, 351–366.

Kehoe, Timothy, and David Levine, 1993, Debt-Constrained Asset Markets, *Review of Economic Studies* 60, 865–888.

Kiyotaki, Nobuhiro, and John Moore, 1997, Credit Cycles, *Journal of Political Economy* 105, 211–284.

Kocherlakota, Narayana, 1996, Implications of Efficient Risk Sharing without Commitment, *Review of Economic Studies* 63, 595–609.

Kushner, Harold, and Paul Dupuis, 2001, *Numerical methods for stochastic control problems in continuous time.* (Springer).

Lorenzoni, Guido, and Karl Walentin, 2007, Financial Frictions, Investment and Tobin's q, *Working paper, MIT.*

Lustig, Hanno, Chad Syverson, and Stijn Van Nieuwerburgh, 2011, Technological Change and the Growing Inequality in Managerial Compensation, *Journal of Financial Economics.*

Luttmer, Erzo, 2007, Selection, Growth, and the Size Distribution of Firms, *Quarterly Journal of Economics* 122, 1103–1144.

Luttmer, Erzo, 2010, Models of growth and firm Heterogeneity, *Annual Review of Economics* 2, 547–576.

Luttmer, Erzo, 2012, Technology diffusion and growth, *Journal of Economic Theory, Forthcoming* 147, 602–622.

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

Rampini, Adriano, and S. Viswanathan, 2012, Collateral and capital structure, *Working paper, Duke University.*

Sannikov, Yuliy, 2008, A Continuous-Time Version of the Principal-Agent Problem, *Review of Economic Studies* 75, 957–984.

Tervio, Marko, 2003, The Difference That CEOs Make: An Assignment Model Approach, *Working paper, University of California, Berkeley.*

Appendix

A Solution to the Optimal Contracts

A.1 Proof of Proposition 1

By (14), the continuation utility of the manager is his reservation utility and does not vary across time. Therefore, now we set $U = 0$ and the argument U can be omitted from the value function. The HJB differential equation characterizing the value function is

$$\begin{aligned}(\mathbf{r}+\kappa) V(K) &= \max_i \Big\{ (A - h(i)) K \\ &\quad + K (i-\delta) V'(K) + \frac{1}{2} K^2 \sigma^2 V''(K) \Big\}\end{aligned} \tag{26}$$

By (12) the maximization problem is homogenous of degree one with respect to K. So $V(K) = \bar{v}K$ with $\bar{v} \in \mathbb{R}$ and the HJB (26) becomes to

$$(\mathbf{r}+\kappa)\, \bar{v} = \max_i \{A - h(i) + (i-\delta)\, \bar{v}\}, \tag{27}$$

which implies

$$\bar{v} = \max_i \frac{A - h(i)}{\hat{r} - i}.$$

The objective right hand side of (27) is a quadratic function of i with the second order term coefficient being negative. Therefore the first order condition is sufficient for maximization and the equation above implies

$$A - h(\hat{\imath}) - h'(\hat{\imath})(\hat{r} - \hat{\imath}) = 0$$

and

$$A - \hat{\imath} - \frac{1}{2} h_0 \hat{\imath}^2 - (1 + h_0 \hat{\imath})(\hat{r} - \hat{\imath}) = 0.$$

Therefore $\hat{\imath}$ satisfies the quadratic equation

$$\frac{1}{2} h_0 \hat{\imath}^2 - h_0 \hat{r} \hat{\imath} + (A - \hat{r}) = 0$$

which has two positive roots

$$\hat{r} \pm \sqrt{\hat{r}^2 - \frac{2}{h_0}(A - \hat{r})}.$$

Note that, as h_0 converges to ∞, i converges to 0. Therefore we have

$$\hat{\imath} = \hat{r} - \sqrt{\hat{r}^2 - \frac{2}{h_0}(A - \hat{r})}.$$

and assumption $0 < A - \hat{r} < \frac{1}{2} h_0 \hat{r}^2$ implies that $\hat{\imath} \in (0, \hat{r})$.

Suppose the manager's reservation utility is $U > 0$. Then (13) and (14) imply that the expected present value of the manager's total compensation is

$$E\left[\int_0^\tau e^{-\mathbf{r}t} U dt\right] = \frac{1}{\mathbf{r}+\kappa} U$$

and $V(K,U) = \bar{v}K - \frac{1}{\mathbf{r}+\kappa}U$. We have the desired results.

A.2 Proof of Proposition 2

Suppose in the steady state generated by the first best allocation rule the total capital stock is $\mathbf{K}$. Then over each unit of time $(\kappa+\delta)\,\mathbf{K}$ units of capital evaporate, $\hat{\imath}\mathbf{K}$ units of capital are built in existing firms and K^* units of new capital are created. Therefore

$$(\kappa+\delta)\,\mathbf{K} = \hat{\imath}\mathbf{K} + K^*$$

and we have (19).

Under the first best allocation rule, the law of motion of capital stock of each firm is

$$dK_t = K_t\left[(-\delta+\hat{\imath})\,dt + \sigma dB_t\right]$$

with death rate being κ. In steady state, 1 unit of firms dies and 1 unit of new firms are established with initial size K^*. Therefore, the stationary distribution of firms size $\phi(K)$ satisfies the following Kolmogrov forward equation.

$$0 = -\frac{\partial}{\partial K}\left[(-\delta+\hat{\imath})\,K\phi(K)\right] + \frac{\partial^2}{(\partial K)^2}\left[\frac{(\sigma K)^2}{2}\phi(K)\right] - \kappa\phi(K).$$

Elementary solutions of this differential equation are of the form $\phi(K) = CK^{-\alpha-1}$ for some constant C and α satisfies the following quadratic equation.[17]

$$\alpha\,(-\delta+\hat{\imath}) + \frac{\delta^2}{2}\alpha\,(\alpha-1) - \alpha = 0$$

or

$$\kappa + \left(\hat{\imath} - \delta - \frac{1}{2}\delta^2\right)\alpha - \frac{1}{2}\alpha^2\sigma^2 = 0$$

which has a positive root

$$\alpha_1 = \frac{-\left(\hat{\imath}-\delta-\frac{1}{2}\delta^2\right) + \sqrt{\left(\hat{\imath}-\delta-\frac{1}{2}\delta^2\right)^2 + 2\alpha^2\sigma^2}}{2\kappa}$$

[17]See Luttmer (2012) and Gabaix (2009).

and a negative root

$$\alpha_2 = \frac{-\left(\hat{\imath} - \delta - \frac{1}{2}\delta^2\right) - \sqrt{\left(\hat{\imath} - \delta - \frac{1}{2}\delta^2\right)^2 + 2\alpha^2\sigma^2}}{2\kappa}.$$

Then $C = \frac{-\alpha_1\alpha_2}{(\alpha_1-\alpha_2)K^*}$ and

$$\phi(K) = \begin{cases} C\left(\frac{K}{K^*}^{-\alpha_2-1}\right), & \text{if } K < K^*; \\ C\left(\frac{K}{K^*}^{-\alpha_1-1}\right), & \text{if } K > K * . \end{cases}$$

Therefore, we have the desired result.

A.3 Characterization of the Normalized Value Function $v(u)$

Let $V(K,U)$ be the value function generated by the optimal contracts. By (9) and (10), it satisfies the following HJB differential equation.

$$\begin{aligned} 0 \;=\; & \max_{C,I,G}\Big\{ (AK - C - h(I,K)) - (\beta+\kappa)V(K,U) + V_K(K,U)K\left(\frac{I}{K} - \delta\right) \\ & + V_U(K,U)\left[-\frac{\beta+\kappa}{1-\gamma}\left(C^{1-\gamma}U^{\gamma} - U\right) + \frac{1}{2}\gamma\frac{G^2}{U}\right] \\ & + \frac{1}{2}V_{KK}(K,U)K^2\sigma^2 + \frac{1}{2}V_{UU}(K,U)G^2 + V_{KU}(K,U)K\sigma G\Big\} \end{aligned} \tag{28}$$

By homogeneity of degree one and normalization of the continuation utility

$$V(K,U) = KV\left(1, \frac{U}{K}\right) = Kv(u).$$

Therefore

$$\begin{aligned} V_K((K,U) &= v(u) - uv'(u) \\ V_U((K,U) &= v'(u) \\ V_{KK}((K,U) &= \frac{1}{K}u^2 v''(u) \\ V_{UU}((K,U) &= \frac{1}{K}v''(u) \\ V_{KU}((K,U) &= -\frac{1}{K}uv''(u) \end{aligned}$$

Define $g = \frac{G}{U^{1-\gamma}}$ and plug the derivatives of $V(K,U)$ into (28), then we have the HJB that $v(u)$

satisfies.

$$
\begin{aligned}
0 \;=\; & \max_{c,i,g}\Bigg\{ (A-c-h(i)) - (\beta+\kappa)\,v(u) + \big(v(u)-uv'(u)\big)\,(i-\delta) \\
& +\frac{1}{1-\gamma} v'(u)u\,(\beta+\kappa)\left(1-\left(\frac{c}{u}\right)^{1-\gamma}\right) + \frac{1}{2}u^2v''(u)\sigma^2 \\
& +\frac{1}{2}\left[\gamma uv'(u)+u^2v''(u)\right] g^2\sigma^2 - u^2v''(u)g\sigma^2 \Bigg\}. \qquad (29)
\end{aligned}
$$

Note that the optimal g maximizes a quadratic function with the coefficient of the second order term being negative on the right hand side of (29). Therefore the first order condition is sufficient and we have the optimal g satisfies

$$
g(u) = \frac{uv''(u)}{\gamma v'(u) + uv''(u)}.
$$

A.4 The Lower bound of the Manager's Normalized Continuation Utility, u_{MIN}, in Two-Sided Limited Commitment Case

Let K_t^D be the capital stock at t that the manager has after his default at τ_D. Then he solves the following problem:

$$
\max_i E\left[\int_{\tau_D}^{\tau} e^{-\beta(t-\tau_D)}\,(\beta+\kappa)\,\frac{\left((A-h(i_t))\,K_t^D\right)^{1-\gamma}}{1-\gamma}dt\right]
$$

subject to

$$
dK_t^D = K_t^D\left[(i_t-\delta)+\sigma dB_t\right].
$$

Let $W\left(K^D\right)$ be the associated value function of this problem and then it satisfies the following HJB differential equation

$$
\begin{aligned}
0 \;=\; & \max_i\Bigg\{(\beta+\kappa)\,\frac{\left((A-h(i))\,K^D\right)^{1-\gamma}}{1-\gamma} - (\beta+\kappa)\,W(K^D) \\
& +W'\left(K^D\right)K^D\,(i-\delta) + \frac{1}{2}W''\left(K^D\right)\left(K^D\right)^2\sigma^2\Bigg\}
\end{aligned}
$$

Note that the objective function of the manager's maximization problem is homogenous of degree $1-\gamma$ with respect to the capital stock at each point of time. So $W\left(K^D\right) = u^D\left(K^D\right)^{1-\gamma}$ for some constant u^D and then $W'\left(K^D\right) = (1-\gamma)\,u^D\left(K^D\right)^{-\gamma}$, $W''(K^D) = -\gamma\,(1-\gamma)\,u^D\left(K^D\right)^{-\gamma-1}$. Then (30) becomes

$$
0 = \max_i\left\{(\beta+\kappa)\,\frac{(A-h(i))^{1-\gamma}}{1-\gamma}\frac{1}{u^D} - (\beta+\kappa) + (1-\gamma)\,(i-\delta) - \frac{1}{2}\gamma\,(1-\gamma)\,\sigma^2\right\}. \qquad (30)
$$

The first order condition of the optimal default investment rate i^D implies

$$(\beta+\kappa)\left(A-h(i^D)\right)^{1-\gamma}\frac{1}{u^D}=(1-\gamma)\,\frac{A-h(i^D)}{h'(i^D)}.$$

By plugging in the first order condition into (30), we have

$$0=\frac{A-h(i^D)}{h'(i^D)}+(1-\gamma)i^D-\left[(\beta+\kappa)+(1-\gamma)\,\delta+\frac{1}{2}\gamma\,(1-\gamma)\,\sigma^2\right]. \tag{31}$$

Define $\hat{\beta}\equiv(\beta+\kappa)+(1-\gamma)\,\delta+\frac{1}{2}\gamma\,(1-\gamma)\,\sigma^2$ and (31) can be written as

$$h_0\left(\frac{1}{2}-\gamma\right)\left(i^D\right)^2+\left(h_0\left(1-\gamma-\hat{\beta}\right)-\gamma\right)i^D+\left(A-\hat{\beta}\right)=0$$

which is a quadratic function with two roots

$$\frac{-\left[h_0\left(1-\gamma-\hat{\beta}\right)-\gamma\right]\pm\sqrt{\left(h_0\left(1-\gamma-\hat{\beta}\right)-\gamma\right)^2-4h_0\left(\frac{1}{2}-\gamma\right)\left(A-\hat{\beta}\right)}}{2h_0\left(\frac{1}{2}-\gamma\right)}.$$

Since, as h_0 converges to 0, the default investment is well defined[18] and then i^D is equal to the root with "+" sign. Therefore,

$$u^D=\frac{\beta+\kappa}{1-\gamma}\left(A-h\left(i^D\right)\right).$$

Note that the manager has capital θK when default and normalization of u implies that

$$u_{MIN}=\theta\left[A-h\left(i^D\right)\right]^{\frac{1}{1-\gamma}}.$$

B The Cross-sectional Distribution of Firms

Let $i(u)$, $c(u)$ be the normalized policy functions indicating the investment-to-capital ratio and compensation-to-capital ratio respectively, and let $g(u)\equiv\frac{G(K,U)}{U^{1-\gamma}}$. According to the definition of the normalized continuation utility, we have its law of motion as follows.

$$du_t=u_t\left[\mu(u_t)dt+\varrho(u_t)dB_t\right]$$

with

$$\begin{aligned}\mu(u)&=\frac{\beta+\kappa}{1-\gamma}\left(1-\left(\frac{c(u)}{u}\right)^{1-\gamma}\right)-(i(u)-\delta)+\left(\frac{1}{2}\gamma g^2(u)-g(u)+1\right)\sigma^2\\ \varrho(u)&=(g(u)-1)\,\sigma\end{aligned}$$

[18] One can easily check that $i^D=\frac{1}{\gamma}\left(A-\hat{\beta}\right)$ if there is no adjustment cost.

for $u \in [u_{MIN}, u_{MAX}]$.

If $\kappa+\delta-i(u) > 0$ for all $u \in [u_{MIN}, u_{MAX}]$, then $m(u)$ satisfies the following Kolmogrov forward equation.

$$\begin{aligned} 0 &= -(\kappa+\delta-i(u))\, m(u) - (1-\gamma)\frac{\partial}{\partial u}\left\{m(u)u\left[\mu(u) + (\varrho(u)-1)\,\sigma^2\right]\right\} \\ &\quad + \frac{1}{2}(1-\gamma)^2\frac{\partial^2}{\partial u^2}\left[m(u)u^2\varrho(u)^2\right]. \end{aligned}$$

Table 1
Estimates of the Power Law Exponent of Firm Size and CEO Compensation

Year	Size	CEO Comp
1992	1.06	1.68
1993	1.13	1.73
1994	1.20	1.77
1995	1.18	1.63
1996	1.18	1.61
1997	1.15	1.42
1998	1.17	1.58
1999	1.15	1.32
2000	1.12	1.22
2001	1.06	1.35
2002	1.05	1.53
2003	1.05	1.60
2004	1.04	1.69
2005	1.10	1.63
2006	1.10	1.76
2007	1.09	1.81
2008	1.10	1.93
2009	1.08	2.05
2010	1.07	2.23
2011	1.09	2.05

Table 1 presents the power-law exponent of firm size and CEO compensation estimated using the largest 300 firms in a given year. Size in the data is measured by the number of firm employees.

Table 2

Elasticity of CEO Compensation with respect to Firm Size

	Size Measures			
	Employees	Market Cap	Capital	Assets
Elasticity	0.31	0.33	0.26	0.37
SE	0.02	0.02	0.02	0.02
R^2	0.20	0.27	0.21	0.23

Table 2 reports estimates of elasticity of CEO compensation with respect to firm size. We consider several measures of firm size: the number of employees (“Employees”), market capitalization (“ Market Cap”), total capital stock (“Capital”), and book value of assets (“Assets”).

Table 3

Distribution of CEO Compensation across Size and Age

Panel A: Data

	Age		
Size	Young	Median	Old
Small	1.14	1.25	0.97
Median	1.75	1.57	1.45
Large	2.05	2.08	1.91

Panel B: Model

	Age		
Size	Young	Median	Old
Small	0.39	-0.04	-0.49
Median	0.70	0.63	0.77
Large	3.09	3.96	5.15

Table 3 presents the distribution of CEO compensation across firms of different size and age. Panel A reports median log CEO compensation in the data, Panel B presents model-implied statistics. Size in the data is measured by the number of firm employees.

Table 4

Distribution of CEO Pay-to-Capital across Size and Age

Panel A: Data

	Age		
Size	Young	Median	Old
Small	0.349	0.297	0.204
Median	0.081	0.039	0.013
Large	0.017	0.013	0.005

Panel B: Model

	Age		
Size	Young	Median	Old
Small	0.168	0.131	0.102
Median	0.057	0.019	0.008
Large	0.001	0.002	0.002

Table 4 presents the distribution of CEO compensation as a fraction of total capital stock across firms of different size and age. Panel A reports the average CEO compensation-to-capital ratio in the data, Panel B presents model-implied statistics. Size in the data is measured by the number of firm employees.

Table 5

Distribution of I/K across Size and Age

Panel A: Data

	Age		
Size	Young	Median	Old
Small	0.38	0.28	0.24
Median	0.24	0.18	0.13
Large	0.18	0.15	0.10

Panel B: Model

	Age		
Size	Young	Median	Old
Small	0.34	0.25	0.21
Median	0.16	0.13	0.13
Large	0.11	0.12	0.12

Table 5 presents the distribution of investment-to-capital ratio across firms of different size and age. Panel A reports average I/K in the data, Panel B presents model-implied statistics. Size in the data is measured by the number of firm employees.

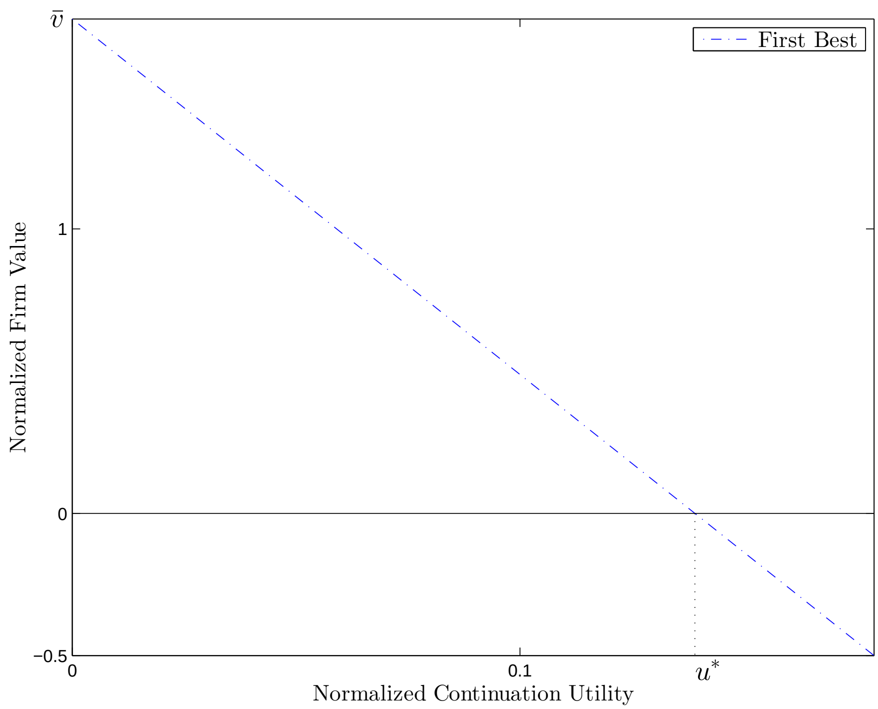


Figure 1. Normalized Value Function for the First Best Case

Figure 1 is the normalized value function for the first best case. $\bar{v}$ is the marginal/average Q in the Hayashi (1982) model. u^* is the point where firm value is zero.

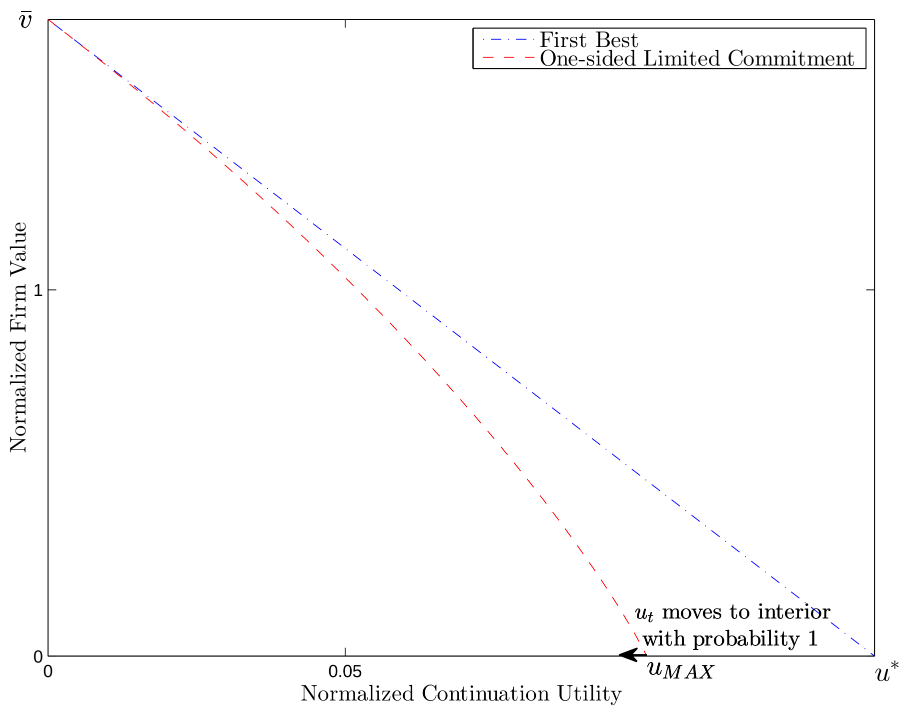


Figure 2. Normalized Value Function: One-Sided Limited Commitment

Figure 2 plots the value function for the case of one-sided limited commitment (dashed line), and that for the case of first best (dash-dotted line). At the bankruptcy point, u_{MAX}, normalized continuation utility returns to the interior with probability one.

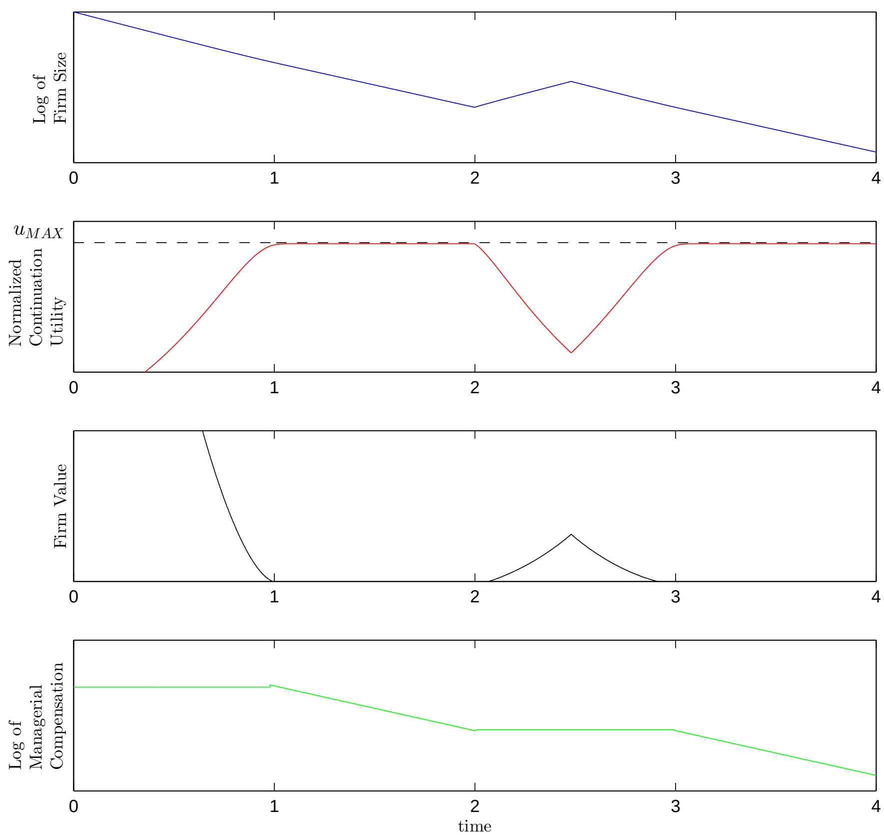


Figure 3. Sample Path of CEO Compensation: One-Sided Limited Commitment

Figure 3 plots the sample paths of firm size (top panel), normalized continuation utility (second panel), firm value (third panel), and log CEO pay (bottom panel) in the neighborhood of the bankruptcy point after a sequence of productivity shocks from time 0 to time 4.

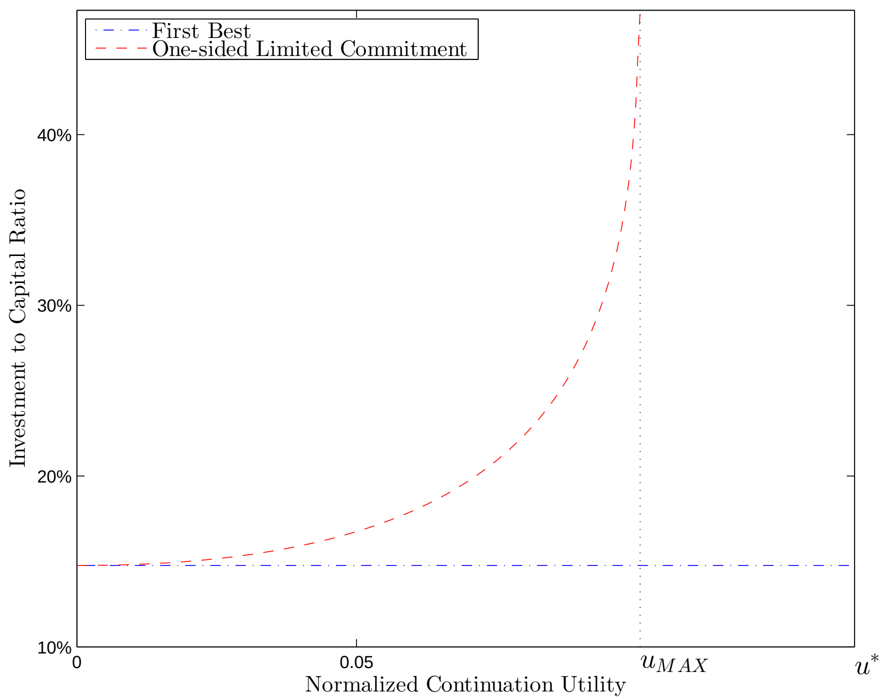


Figure 4. Investment Policy: One-Sided Limited Commitment

Figure 4 plots the optimal investment rate ($\frac{I}{K}$) as a function of the state variable u. Investment rate is monotonically increasing in normalized continuation utility, u.

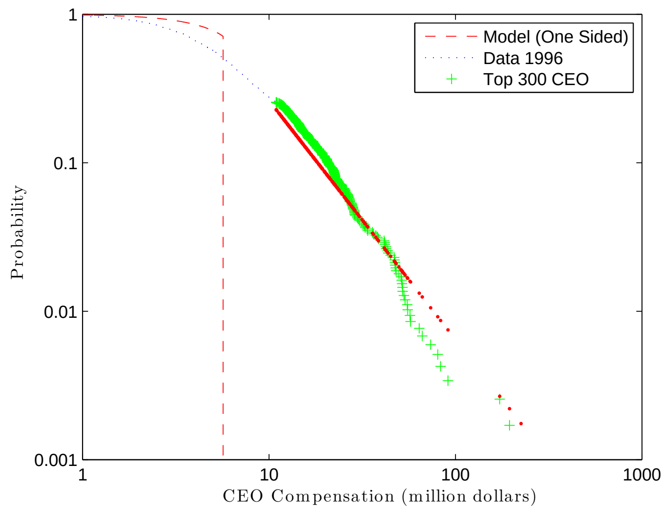


Figure 5. Distribution of CEO Compensation: One-Sided Limited Commitment

Figure 5 plots the counter-cumulative distribution function of CEO compensation in the model with one-sided limited commitment (dashed line) and that in the data (dotted line). Both axis are log-equally spaced. The right tail of the empirical counter-cumulative distribution is marked with "+". Thick dotted line represents the estimated slope of the power law in the data.

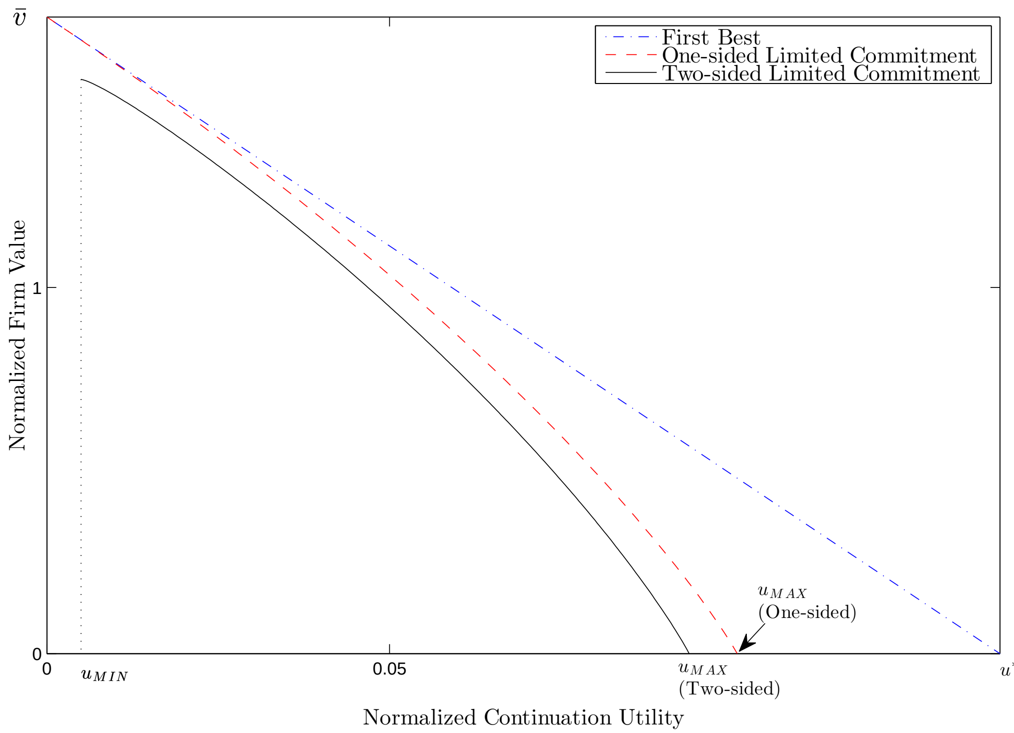


Figure 6. Normalized Value Function: Two-Sided Limited Commitment

Figure 6 plots the value function for the case of two-sided limited commitment (solid line), that for the case of one-sided limited commitment (dashed line), and that for the case of first best (dash-dotted line). Normalized continuation utility returns to the interior with probability one at both boundaries, u_{MIN} and u_{MAX}. The set of equilibrium payoffs in the two-sided case is a strict subset of that in the case of one-sided limited commitment.

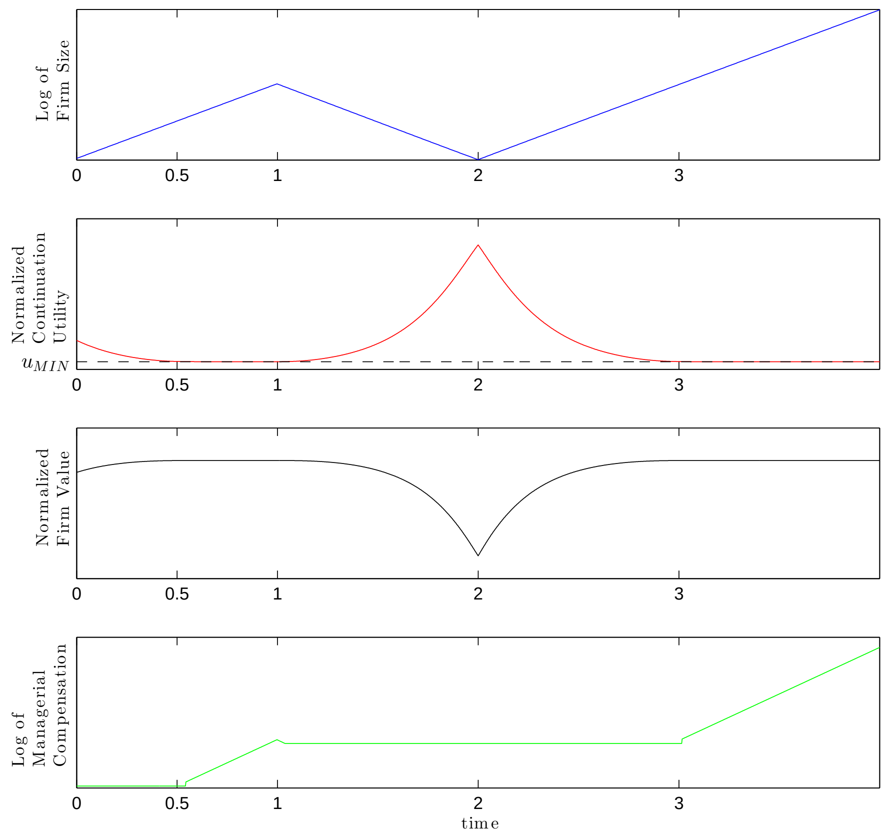


Figure 7. Sample Path of CEO Compensation: Two-Sided Limited Commitment

Figure 7 plots the sample paths of firm size (top panel), normalized continuation utility (second panel), firm value (third panel), and log CEO pay (bottom panel) in the neighborhood of the binding participation constraint, u_{MIN}, after a sequence of productivity shocks from time 0 to time 4.

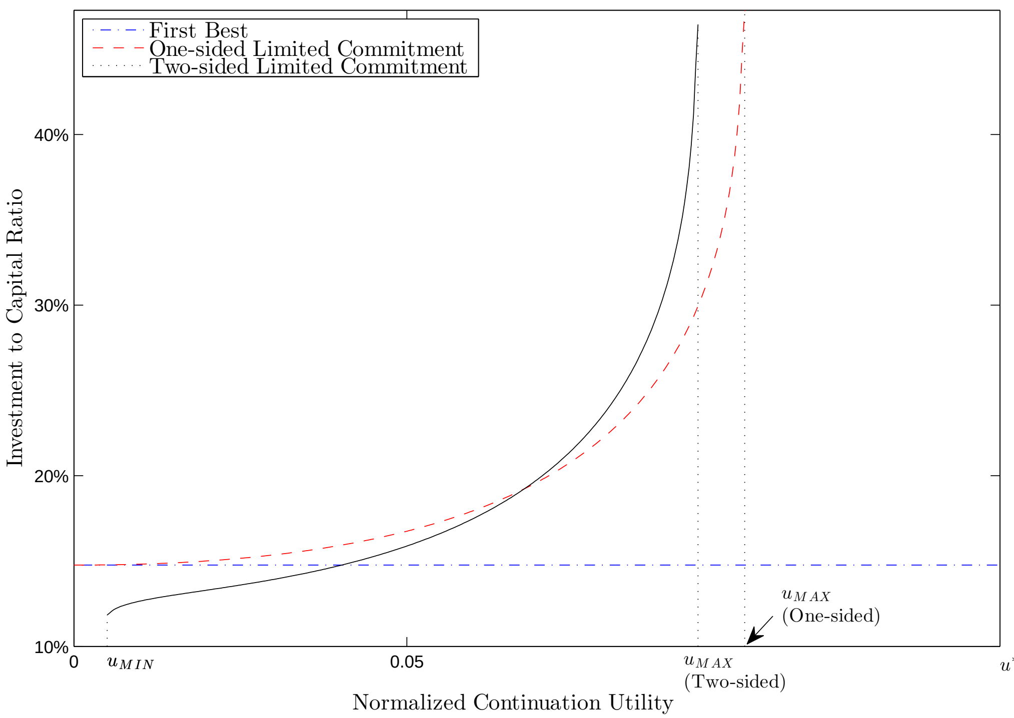


Figure 8. Investment Policy: Two-Sided Limited Commitment

Figure 8 plots the optimal investment rate ($\frac{I}{K}$) as a function of the state variable u. Investment rate is monotonically increasing in normalized continuation utility, u.

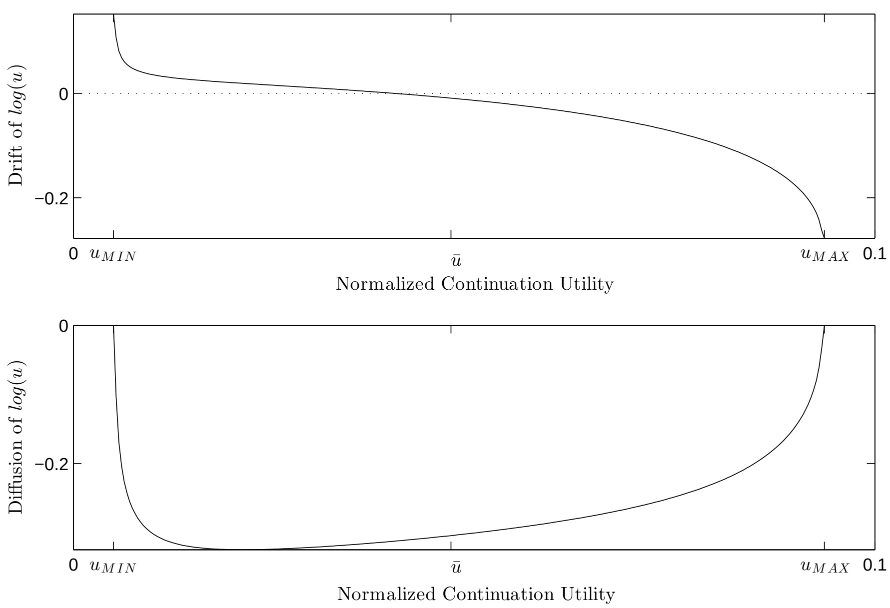


Figure 9. Dynamics of Normalized Continuation Utility: Two-Sided Limited Commitment

Figure 9 illustrates the dynamics of the normalized continuation utility, u, for the case of two-sided limited commitment. The drift of u is strictly positive on the left, monotonically decreasing, and strictly negative on the right. The diffusion is zero on the boundaries and strictly negative in the interior.

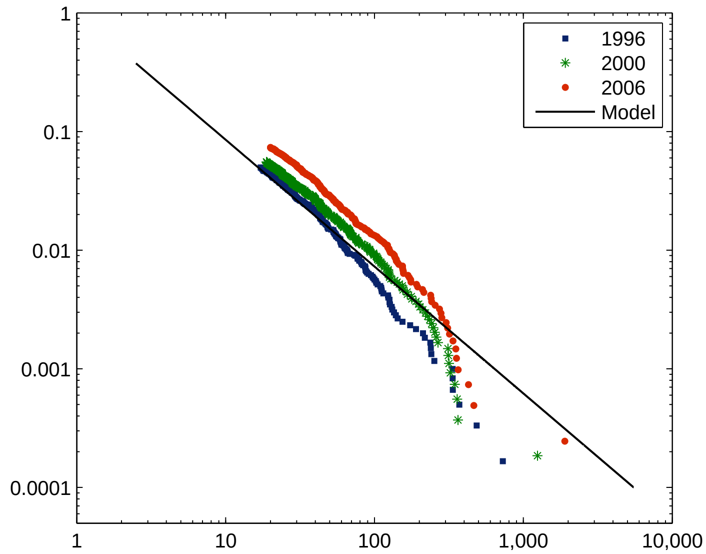


Figure 10. The Right Tail of Size Distribution

Figure 10 plots the right tail of size distribution in the data (for 1996, 2000 and 2006) and the slope implied by the model. Size is measured by the number of firm employees (in thousands).

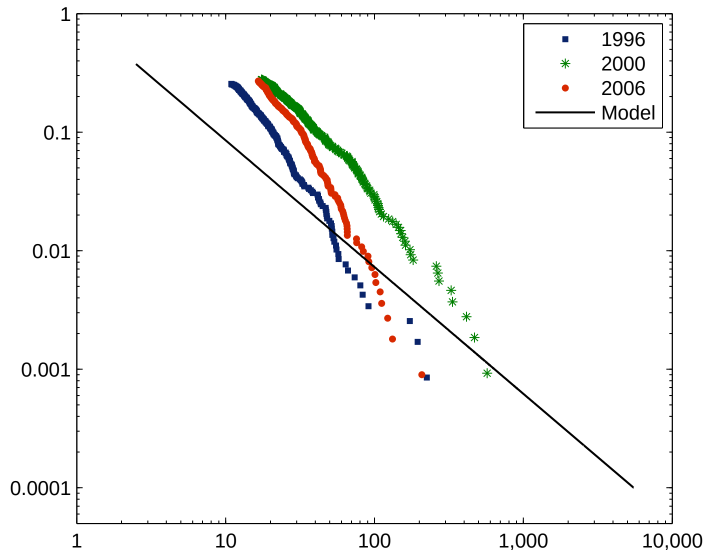
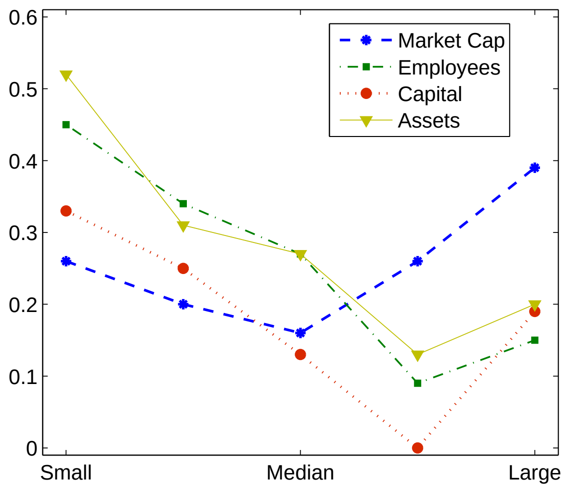
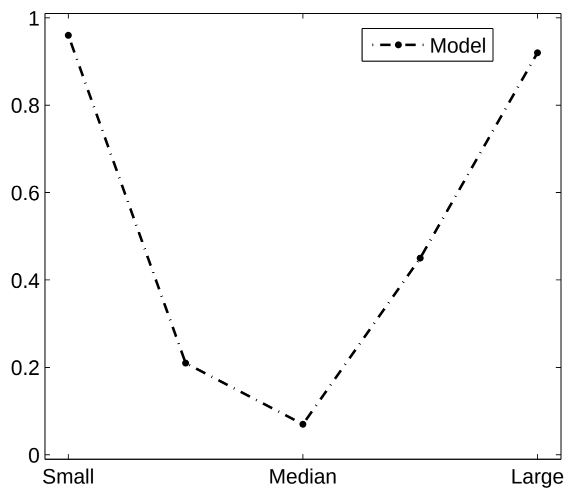


Figure 11. The Right Tail of the Distribution of CEO Compensation

Figure 11 plots the right tail of the distribution of CEO compensation in the data (for 1996, 2000 and 2006) and the slope implied by the model. CEO compensation is measured in million of dollars.



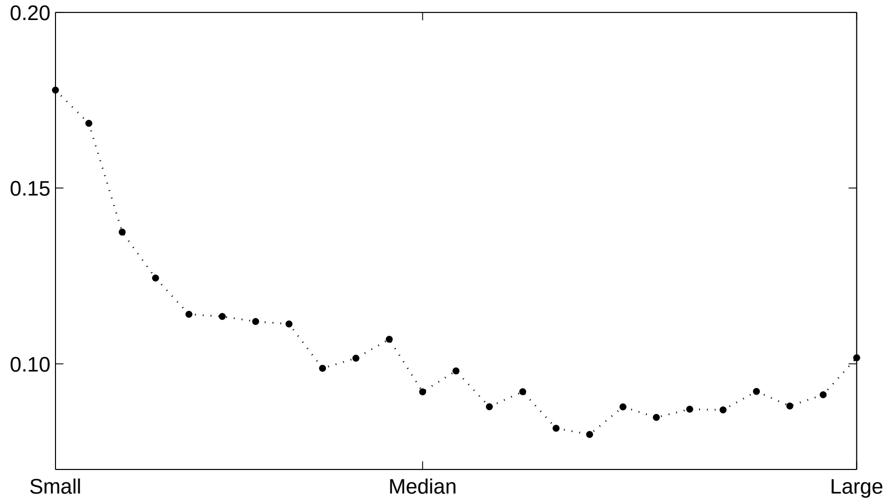
(a) Data



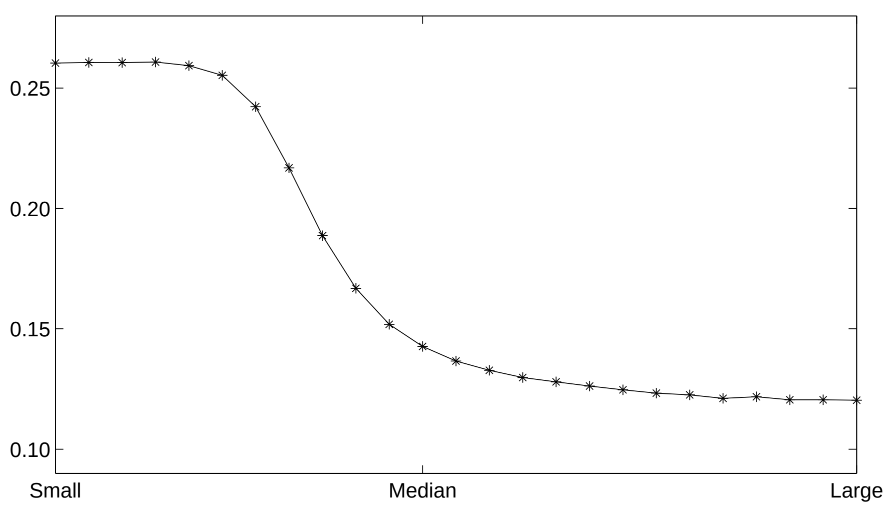
(b) Model

Figure 12. Elasticity of CEO Compensation w.r.to Size conditional on Size

Figure 12 plots the estimates of elasticity of CEO compensation with respect to firm size in the data (Panel (a)) and in the model (Panel (b)). We consider several measures of firm size in the data: market capitalization (" Market Cap"), the number of employees ("Employees"), total capital stock ("Capital"), and book value of assets ("Assets").



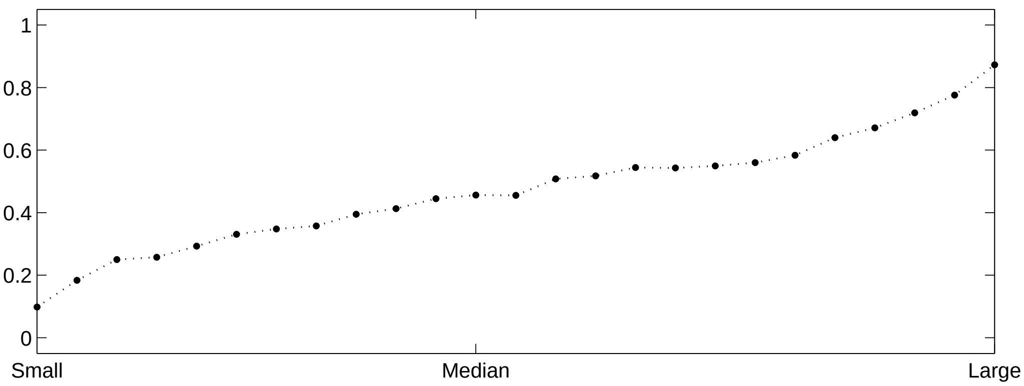
(a) Data



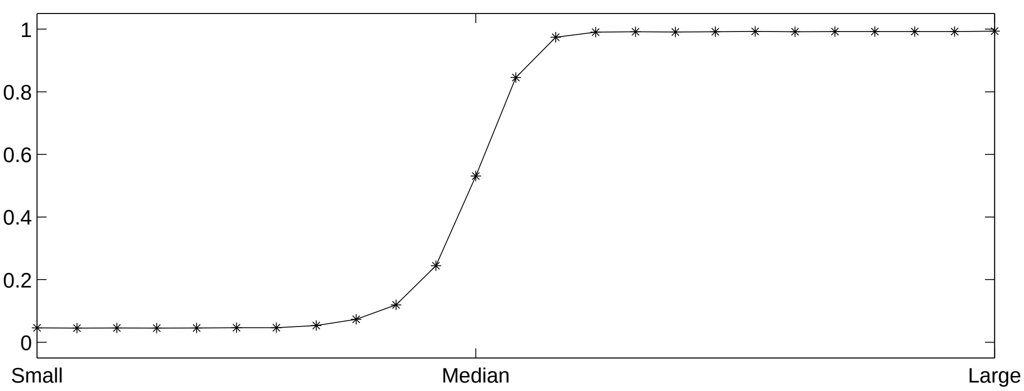
(b) Model

Figure 13. Investment-to-Capital Ratio across Size Percentiles

Figure 13 plots the average ratio of investment to capital across 25 size-sorted portfolios. In the data, size is measured by the number of firm employees.



(a) Data



(b) Model

Figure 14. Fraction of Dividend(Interest)-Paying Firms across Size Percentiles

Figure 14 plots the fraction of firms that make dividend and/or interest payments across 25 size-sorted portfolios. In the data, size is measured by the number of firm employees.