We present a general equilibrium-mechanism design model with two-sided limited commitment that helps account 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 payout, as well as variation of firms’ investment and payout policies across both size and age.

JEL Code: E13, G30
Introduction

It is well known that investment and dividend policies vary substantially with firm size. Small firms tend to invest at a much higher rate and pay out less dividends (if any) compared with large firms. CEO compensation contracts are also tied to firm size. While the overall elasticity of managerial compensation with respect to firm size is significantly positive, it is far from uniform – it is larger in the left and right tails of size distribution and it is much smaller for medium-sized firms. In addition, both the distribution of firm size and the empirical distribution of CEO compensation are characterized by a power law. In this paper, we aim to understand the joint distribution of firm size and managerial compensation and the observed heterogeneity in firms’ investment and payout policies in a general equilibrium model with limited commitment. We take a mechanism design approach and explore the implications of the constrained efficiency on firms’ dynamics. We show that limited commitment on the shareholder and on the manager side can account for many salient features of the cross-sectional data.

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 (NPV) projects, and that managers cannot commit to wage contracts that result in life-time utility lower than their outside options.

We establish several important results. First, we show that two-sided limited commitment implies a positive relationship between CEO pay and firm size and a V-shaped size elasticity of CEO compensation. We show that under the optimal contract, CEO compensation remains constant as long as none of the commitment constraints binds, which is characteristic of medium-sized firms. A sequence of positive productivity shocks increases the value of the firm and the manager’s outside option and, once his constraint binds, forces shareholders to raise the level of compensation to retain the manager. A sequence of negative shocks lower the value of the firm and may put the firm at risk of bankruptcy. If the size of the firm continues to fall, shareholders cut the manager’s wage as they are not committed to bear additional losses. Hence, in our model CEO compensation is sensitive to firm size and more so for large and small firms, as in the data.

Second, our model endogenously generates a power law in firm size and CEO compensation. Given that the technology is constant return to scale, firm growth follows a process similar to the discrete time processes studied in Kesten (1973) and Saporta (2005) and results in a fat-tailed distribution of firm size. Because managers’ outside option rises with firm size, their compensation under the optimal contract has to increase accordingly.

1See also Gabaix (2009) and Benhabib, Bisin, and Zhu (2011).
Consequently, the power law in firm size translates into a power law in CEO compensation.\(^2\)

We further show that our model predicts an inverse relationship between investment rate and size, and a positive relationship between dividend payout and firm size. Limited commitment on the shareholder side creates strong incentives for small firms to invest. When a firm value is close to zero, the commitment constraint is likely to bind, which is welfare reducing. To avoid further downsizing and improve risk sharing, small firms choose to accelerate investment. By the same logic, small firms tend to have low or zero dividend yields as they spend most of their resources on funding investments. In contrast, large firms are likely to face a binding participation constraint on the manager side because managers’ outside options become more attractive as firms grow. To reduce the likelihood of a binding constraint, it is optimal for large firms to slow down their investment and growth. Hence, consistent with the data, small firms in the model pay out less, invest more, and grow faster compared with large firms.

We calibrate the model to match standard macroeconomic moments and volatility of output at the firm level and show that it can quantitatively account for the key characteristics of the joint distribution of firms’ size, investment, payout and CEO compensation policies. We also show that, despite its simplicity, our model has rich implications for investment and payout decisions conditional on both firm size and age.

We show that both types of limited commitment, on the manager side and on the shareholder side, are important for understanding empirical relationships among CEO compensation, firms’ investment and size. To highlight their importance, we first discuss the implications of 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 managerial compensation. Due to convex adjustment costs, all firms in this set-up have the same investment-to-capital ratio and identical expected growth rates. Gibrat (1931)’s law holds, and hence, this specification can account for the power law in firm size. Contrary to the data, though, it implies a zero correlation between CEO pay and size and rules out any dependence of firm growth on size. Introducing limited commitment on the manager side allows the model to generate a power law in CEO compensation. However, as in the frictionless case, the elasticity of CEO pay with respect to firm size is zero for small firms. In addition, quantitatively, firms’ investment and growth rates vary very little with size.

Our paper builds on the large literature on limited commitment and its implications

\(^2\)As we show, CEO compensation under the optimal contract is approximately a linear function of the running maximum of firm size. CEO pay features a fat-tailed distribution because the running maximum of the firm growth process obeys a power law similar to that in firm size.
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. Krueger and Uhlig (2006) consider a competitive risk sharing model with one-sided limited commitment. More recently, Lorenzoni and Walentin (2007) study the implications of limited commitment for 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 limited commitment on the manager side and study the link between the inequality of CEO compensation and productivity growth. Our model differs from the above literature in several respects. We work in a general-equilibrium setting and solve a mechanism-design problem with two-sided limited commitment whereas others typically focus on limited commitment on the agent side only. We use continuous-time methods to characterize the solution to the optimal contract and the cross-sectional distribution of firms as ordinary differential equations that allows for sharper analytical results and efficient numerical solutions. In addition, none of above mentioned papers attempts to explain power law in firm size and CEO compensation and their interaction.

A recent paper by Cooley, Marimon, and Quadrini (2012) considers a discrete-time model with two-sided limited commitment and study its implications on the size of the financial sector and compensation of financial executives. From the modeling perspective, the form of limited commitment on the shareholder side in our model differ from theirs. In their model, shareholders cannot commitment to any compensation plans that provides higher utility than managers’ outside option. As a result, in their model managers always receive their outside options. We consider the case where shareholders cannot commit to negative NPV projects, and the optimal contract allows for risk sharing. Also, unlike Cooley, Marimon, and Quadrini (2012), we focus on understanding firm dynamics and power law in firm size and CEO compensation.

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

---

4Our paper is also related to previous work that studies the implications of dynamic agency for firms’ investment and financing decisions, for example, Quadrini (2004), Clementi and Hopenhayn (2006), and DeMarzo and Fishman (2007).

4Many other papers also study firm decisions in models with limited commitment, for example Schmid (2008), Arellano, Bai, and Zhang (2012), and Li (2013).

5Lustig, Syverson, and Van Nieuwerburgh (2011) is an exception. Their model also implies a power law for the distribution of firm size.
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 taking size distribution as given.\footnote{For a survey on the literature of the economics of super stars, see Gabaix and Landier (2008).} Our model provides an alternative, mechanism-design based explanation of the level of CEO pay and its dependence on firm size. In our model, both the distribution of firm size and CEO compensation are endogenous outcomes of the optimal contract. An additional advantage of our dynamic model is that it can be used to study both the cross-sectional distribution and the life-cycle dynamics of firms’ investment, CEO compensation and dividend payout policies.

The continuous-time methodology of this paper builds on the fast growing literature of continuous time dynamic contracting and firm decisions, for example, Sannikov (2008), DeMarzo and Sannikov (2006), DeMarzo, Fishman, He, and Wang (2009), He (2009), Biais, Mariotti, and Villeneuve (2010). For an excellent survey of this literature, see Biais, Mariotti, Plantin, and Rochet (2004).

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, investment, dividend payout and CEO compensation policies. In Section II, we set up a general equilibrium model with limited commitment and show that the equilibrium allocation, if exist, is the solution to a mechanism design problem with two-sided limited commitment. We consider a frictionless Arrow-Debreu economy and discuss its implications in Section III. In Section IV, we augment our first best economy with limited commitment on the manager side and compare our setup to the Albuquerque and Hopenhayn (2004) model. We study our full model with two-sided limited commitment on both the shareholder and the manager side in Section V and evaluate the quantitative implications of the model for the cross-sectional variation of firms’ decision policies. Section VII concludes.

I Stylized Facts

We begin with a brief summary of the empirical distribution of firms’ size, their investment, payout, and CEO compensation policies. The theoretical framework we develop in subsequent sections is aimed at providing a coherent interpretation of the observed features of the cross-sectional data. The details of our empirical analysis and additional evidence are provided in Section VI.

1. The right tails of firm size and CEO compensation obey a power law.\footnote{The power law of the size distribution is well documented in the literature, for example, see Axtell (2001).} Firm size
is characterized by a power-law distribution with an exponent of about 1.1. The distribution of CEO compensation is approximated by a power law with a slope coefficient of around 1.7.

2. The elasticity of CEO pay with respect to firm size is close to $1/3$. The elasticity is V-shaped in size: it is larger in the tails and lower in the middle of firm-size distribution.

3. Small firms invest at a higher rate compared with large firms. The average investment-to-capital ratio of firms in the top size decile is about 10%. Small firms (those in the bottom decile of size distribution) have an average investment rate of around 17%.

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

We will keep this evidence as a reference when discussing the qualitative implications of the models that we lay out in following sections. In Section VI, we calibrate our benchmark model with two-sided limited commitment and formally evaluate its ability to account for these and other cross-sectional characteristics of the data.

II A General Equilibrium Model with Limited Commitment

In this section, we set up a general equilibrium model with heterogeneous firms and limited commitment.

A Preferences

Time is continuous and infinite. There are 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:

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

(1)

8See Roberts (1956), Baker, Jensen, and Murphy (1988), and Frydman and Saks (2010).
where $\beta > 0$ is the time discount rate, and $\gamma > 0$ is the relative risk aversion coefficient. $C_t$ 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.\footnote{Our model can be easily extended to incorporate the case where shareholders and managers have different time discount rate and/or different risk aversion. We do not entertain these extensions to maintain parsimony in our quantitative exercise.} \footnote{We refer to the shareholder as she and the manager as he in the rest of the paper.}

**B Production Technology**

Production in the economy takes place at a continuum of locations indexed by $j \in \mathcal{J}$, where $\mathcal{J}$ is the set of all possible locations. General output at location $j$ at time $t$, denoted by $y_{j,t}$, is produced using capital and labor through a Cobb-Douglas technology:

$$y_{j,t} = K_{j,t}^\alpha (z_t N_{j,t})^{1-\alpha},$$

where $K_{j,t}$ is the amount of capital and $N_{j,t}$ is the amount of labor at location $j$ and time $t$. $z_t$ is the labor-augmenting productivity. We set $z_t = z$ to be constant in our theoretical model for simplicity but allow for aggregate productivity growth in calibration.

The representative shareholder owns all capital in the economy, supplies one unit of labor inelastically per unit of time, but does not have access to production technology. Managers are the only type of agents that have access to production technology. For simplicity, we assume that managers do not supply any labor.

Labor market is competitive. We focus on the stationary equilibrium where market prices are time-invariant. Let $W$ denote the real wage and $\Pi (K)$ denote the operating profit function, that is,

$$\Pi (K) = \max_N \{ z K_t^\alpha N^{1-\alpha} - WN \}$$  \hspace{1cm} (2)

is the total revenue of a firm maximizing out labor input. Because the production technology is constant return to scale, and labor market is competitive, the operating profit function is linear: $\Pi (K) = AK$, where $A$ is the economy-wide (equilibrium) marginal product of capital. Because labor market is perfectly competitive, equation (2) implies that firms’ capital stock and total number of employees are proportional to each other in equilibrium. As a result, $K$ and $N$ are equivalent measures of firm size in our model.

The manager hired at location $j$ has access to a technology that accumulates capital according to the following law of motion:

$$dK_{j,t} = (I_{j,t} - \delta K_{j,t}) dt + K_{j,t} \sigma dB_{j,t},$$  \hspace{1cm} (3)
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. The term $I_{j,t}$ is investment made at location $j$ at time $t$. Investing $I$ at a location with total capital stock $K$ costs $h\left(\frac{I}{K}\right) K$ of general output, where

$$h(i) = 1 + \frac{1}{2}h_0i^2$$

is a strictly convex adjustment cost function.

### C Entry and Exit of Firms

A unit measure of managers arrives at the economy per unit of time. Upon arrival, a manager is endowed with an outside option that delivers life-time utility $\bar{U}$. In order to operate a production technology for the shareholder, the manager 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 realization of productivity shocks. 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 $U$. 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 and the initial promised utility to the manager of a new generation of firms to maximize profit. As we show below, the value function $V(K,U)$ is strictly decreasing in $U$. Therefore, the optimal choice of the initial promised utility to the manager is $\bar{U}$, and the optimal initial size of new firms, denoted by $\bar{K}$, is given by:

$$\bar{K} \in \arg\max_k \left\{ V(K,\bar{U}) - H(K) \right\}. \quad (4)$$

Managers are subject to random health shocks that follow a Poisson process with intensity $\kappa$. 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.

---

\footnote{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, Brownian motion $B_{j,t}$ can be interpreted as a combination of productivity shocks and capital depreciation shocks.}

\footnote{For simplicity, we do not explicitly specify the technology that delivers the reservation utility. The outside option is never taken under our assumptions.}
Profit Maximization under Limited Commitment

The economy we consider is one with overlapping generation of firms. At any point of time \( t \), a new generation of firms is created. We use \( C_{j,t+s}^t \), \( I_{j,t+s}^t \), and \( D_{j,t+s}^t \) to denote the managerial compensation, investment, and dividend payout policy, respectively, for the generation-\( t \) firms at location \( j \) with age \( s \). To save notation, we suppress the time index \( t \), and consider the decision problem of a typical firm \( j \) of age \( s \).

Let \( r \) denote the equilibrium interest rate. Taking \( r \) as given, firm \( j \) chooses a feasible contract \( \{C_{j,s}, I_{j,s}, D_{j,s}\}_{s=0}^{\infty} \) to maximize the present value of dividend payments:

\[
E_0 \left[ \int_0^\tau e^{-rs}D_{j,s}ds \right],
\]

where \( E_0 \) stands for expectation taken with respect to information available when the firm is at age 0, and \( \tau \) is the stopping time at which the manager of firm \( j \) is hit by a Poisson health shock.

We restrict our attention to contracts that satisfy the resource constraint and make the manager at least as better off as his outside options. The resource constraint requires that managerial compensation, cost of investment (including adjustment cost), and dividend payment must not exceed the total operating profit:

\[
C_{j,s} + h \left( \frac{I_{j,s}}{K_{j,s}} \right) K_{j,s} + D_{j,s} \leq \Pi(K_{j,s}), \quad \text{for all } s \geq 0, \tag{6}
\]

where \( \Pi(K) \) is the operating profit function defined in equation (2). Note that we do not impose an exogenous nonnegativity constraint, \( D_s \geq 0 \). Rather, we let the optimal contract determine the amount of dividend payment. A negative dividend payment can be interpreted as equity issuance.

To incentivize managers to work, the compensation plan \( \{C_{j,s}\}_{s=0}^{\infty} \) must deliver a lifetime utility at least as high as their outside options, \( \bar{U} \):

\[
\left\{ E_0 \left[ \int_0^\tau e^{-\beta s} (\beta + \kappa) C_{j,s}^{1-\gamma} ds \right] \right\}^{\frac{1}{1-\gamma}} \geq \bar{U}, \tag{7}
\]

Here we normalize the utility of the manager so that it is measured in the same units as consumption.

The key agency friction in our model is limited commitment, which restricts the set of feasible allocations. Formally, a contract \( \{C_{j,s}, I_{j,s}, D_{j,s}\}_{s=0}^{\infty} \) is feasible if it satisfies equations (6) and (7) and the following two commitment constraints.
First, we assume that shareholders cannot commit to negative net present value projects. In this case, dividend policies that involve negative NPV of the firm at any point in time cannot be implemented. Limited commitment on the shareholder side restricts the set of feasible plans and requires the dividend policy \( \{D_{j,s}\}_{s=0}^{\infty} \) to satisfy:

\[
E_s \left[ \int_s^\tau e^{-r(t-s)} D_{j,t} \, dt \right] \geq 0, \quad \text{for all } s \geq 0.
\]  

(8)

That is, the value of the firm must be non-negative at all times in all states of the world. Intuitively, a binding limited commitment constraint on the shareholder side implies zero firm value and can be interpreted as bankruptcy (Ai and Li (2013)).

Second, following Kehoe and Levine (1993), Kiyotaki and Moore (1997), and Albuquerque and Hopenhayn (2004), we 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 outside option. Upon default, the manager can retain a fraction \( \theta \) of the capital stock and hire labor in 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 the capital stock he absconds with upon default, but cannot enter into any intertemporal risk sharing contract. As we show in the appendix, the utility of the manager upon default is a linear function of the capital stock of the firm: \( u_{\text{MIN}}(\theta) K_s \), where \( u_{\text{MIN}}(\theta) \) is a function of \( \theta \). Limited commitment on the manager side requires the compensation plan \( \{C_{j,s}\}_{s=0}^{\infty} \) to satisfy:

\[
\left\{ E_s \left[ \int_s^\tau e^{-\beta(t-s)} (\beta + \kappa) C_{j,t}^{1-\gamma} \, dt \right] \right\}^{1-\gamma} \geq u_{\text{MIN}}(\theta) K_s, \quad \text{for all } s \geq 0.
\]

(9)

Any compensation plan that violates condition (9) will result in manager’s default on the contract with positive probability. Default destroys risk sharing and implies a non-trivial welfare loss. As a result, efficiency requires that we focus attention on plans that satisfy condition (9).

To distinguish the two constraints, (8) and (9), in what follows we call the limited commitment constraint on shareholder side, equation (8), the bankruptcy constraint, and the limited commitment constraint on manager side, equation (9), the participation constraint.

### E General Equilibrium

A competitive equilibrium must specify the path of interest rates, \( \{r_t\}_{t \geq 0} \), and wages, \( \{W_t\}_{t \geq 0} \), consumption of the representative shareholder, \( \{C_t\}_{t \geq 0} \), consumption, investment, and dividend payout policies for all firms at all times, \( \left\{ \left( C_{j,t+s}, I_{j,t+s}, D_{j,t+s} \right)_{s=0}^{\infty} \right\}_{j \in J} \).
We focus our attention on stationary equilibria where the exit rate of firms equals the entry rate, and the cross-section distribution of firm characteristic is time-invariant. In this case, the interest rate and the wage rate are also time-invariant. Using the results in Ai and Li (2013), the solution to the optimal contracting problem can be constructed by policy functions of state variables \((K, U)\), where \(K\) is the total capital stock of the firm and \(U\) is the continuation utility promised to the manager. Because policies depend only on the state variables \((K, U)\), we can think of them as a summary of firms’ type. Using the language of Atkeson and Lucas (1992), the equilibrium allocation can be achieved by allocation rules:

\[
C(K,U), I(K,U), D(K,U), N(K,U), G(K,U),
\]

that are consistent with the policy functions of the dynamic contracting problem described in equations (5)-(9). Given the rules, allocations can be constructed by 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 rules using:

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

and

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

Equation (11) follows the formulation in Sannikov (2008), except that we use a non-expected utility representation of the preference so that utility is measured in consumption units (equation 7).

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

1. Taking interest rates as given, the allocation constructed from the allocation rules described above solves the firm’s optimal contracting problem that maximizes (5) subject to constraints (6)-(9).

2. The initial condition for news firms is \((\bar{K}, \bar{U})\), where \(\bar{K}\) solves the profit maximization problem in equation (4).

\[\text{13} \text{We prove the existence of such an equilibrium by construction.}\]
3. Taking real wages as given, the policy function \( N(K,U) \) solves the intra-temporal profit maximization problem in Equation (2) for all firms at all times.

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

5. Goods market clears:
\[
C + \int \left[ C(K,U) + h \left( \frac{I(K,U)}{K} \right) K \right] d\Phi(K,U) + H(\bar{K}) = \int K^\alpha (zN)^{1-\alpha} d\Phi(K,U). \tag{12}
\]

6. Labor market clears:
\[
\int N(K,U) d\Phi(K,U) = 1. \tag{13}
\]

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 (10) and (11).\(^{14}\)

\section*{F Normalized Continuation Utility}

The optimal contracting problem described in equations (5)-(9) is homogenous in the state variable \( K \). Let \( V(K,U) \) denote the value function of the optimal contracting problem, then homogeneity implies
\[
V(K,U) = v \left( \frac{U}{K} \right) K \tag{14}
\]
for some smooth function \( v \). Define \( u = \frac{U}{K} \) to be the normalized continuation utility. As shown in Ai and Li (2013), homogeneity of the optimal contracting problem implies that the policy functions satisfy
\[
C(K,U) = c(u) K; \quad I(K,U) = i(u) K \tag{15}
\]
for some continuous functions \( c(\cdot) \) and \( i(\cdot) \). The normalized value function, \( v(u) \), can be characterized by the solution to an ordinary differential equation with appropriate boundary conditions, which we describe in detail in Proposition 4 and 5. In addition, due to homogeneity of decision rules, the two dimensional measure, \( \Phi(K,U) \), in the market clearing conditions (12) and (13) can be replaced by a one-dimensional ”summary measure” as described in Ai (2012). The equilibrium is completely characterized by the two ordinary equations, (26) in Section IV and (48) in Appendix E.

\(^{14}\)Technically, \( \Phi(K,U) \) must satisfy a version of the Kolmogorov forward equation as we show in the appendix.
The Mechanism Design Interpretation of the Model

In this subsection, we provide a mechanism design interpretation of our model. We show that the competitive equilibrium described above is constrained efficient subject to limited commitment constraints. The rest of the paper can be read independently of this subsection.

Our result is similar to that in Atkeson and Lucas (1992). Note, however, Theorem 1 in Atkeson and Lucas (1992) does not apply directly here, because the limited commitment constraint on the shareholder side contains equilibrium prices. We show that constraint (8) can be stated as a rationality constraint on the part of shareholders in terms of allocations.

We first set up some notations. As in Atkeson and Lucas (1992), we index agents by their initial promised utility $U_0$ and the history of realizations of capital stock $\{K_t\}_{t=0}^\infty$, where $\{K_t\}_{t=0}^\infty$ is an element of $F[0, \infty)$, the space of continuous functions on $[0, \infty)$. To save notations, we use $\Psi$ to denote the set of all managers and $\psi \in \Psi$ to denote a generic element of $\Psi$. A subset $\tilde{\Psi} \subseteq \Psi$ is said to be $t$-measurable if it is of the form $\tilde{\Psi} = \{\psi = (U_0, \{K_s\}_{s=0}^\infty) : (U_0, \{K_s\}_{s=0}^t) \in \mathbb{B}\}$, where $\mathbb{B}$ is a Borel subset of $[0, \infty) \times F[0, t]$. Intuitively, $\tilde{\Psi}$ select agents from $\Psi$ based only on date-$t$ available information.

An allocation is denoted as $\{C_t, \tilde{K}_t\}_{t=0}^\infty$, $\{\{C_t(\psi) , I_t(\psi) , N_t(\psi)\}_{t=0}^\infty\}_{\psi \in \Psi}$, where $C_t$ is the consumption of the representative shareholder at time $t$, and $\tilde{K}_t$ is the initial size of firms built at time $t$. For $t \geq 0$, $C_t(\psi)$, $I_t(\psi)$, and $N_t(\psi)$ denote consumption, investment, and the amount of labor hired by manager $\psi$ at time $t$.

An allocation $\{C_t, \tilde{K}_t\}_{t=0}^\infty$, $\{\{C_t(\psi) , I_t(\psi) , N_t(\psi)\}_{t=0}^\infty\}_{\psi \in \Psi}$ is said to satisfy the limited commitment constraint for the shareholder if for any $t$,

$$E_t \left[ \int_0^\infty e^{-\beta s} \frac{1}{1-\gamma} C_{t+s}^{1-\gamma} ds \right] \geq E_t \left[ \int_0^\infty e^{-\beta s} \frac{1}{1-\gamma} \tilde{C}_{t+s}^{1-\gamma} ds \right]$$

for all continuation consumption profiles $\{\tilde{C}_{t+s}\}_{s=0}^\infty$ that satisfy

$$\tilde{C}_{t+s} + H(\tilde{K}_{t+s}) = \int_{\tilde{\Psi}} K_{t+s}^\alpha(\psi) \left(z\tilde{N}_{t+s}(\psi)\right)^{1-\alpha} - h \left( \frac{I_{t+s}(\psi)}{K_{t+s}(\psi)} \right) K_{t+s}(\psi) - C_{t+s}(\psi) d\psi,$$

for all $s$ and

$$\int_{\tilde{\Psi}} \tilde{N}_{t+s}(\psi) d\psi = 1,$$

for some $\tilde{\Psi} \subseteq \Psi$ that is $t$-measurable. The interpretation of equation (16) is that the shareholder can abandon any subset of managers at time $t$ based on available information and redeploy labor among all remaining firms.
The limited commitment constraint (9) is an individual rationality constraint that rules out the incentive for managers to default and take their outside options. The limited commitment constraint (17) is a coalition rationality constraint. It requires that the shareholders do not have jointly profitable deviations together with any subgroup of managers to default on the rest of the managers in the economy. We are now ready to define feasibility and constrained efficiency.

**Definition 1. (Constrained Efficiency)**

An allocation \( \{ C_t, K_t \}_{t=0}^{\infty}, \{ \{ C_t(\psi), I_t(\psi), N_t(\psi) \}_{s=0}^{\infty} \}_{\psi \in \Psi} \) satisfies the resource constraint if for all \( t \),

\[
C_t + H(K_t) = \int_{\Psi} \left[ K_t^\alpha(\psi) (zN_t(\psi))^{1-\alpha} - \left( \frac{I_t(\psi)}{K_t(\psi)} \right) K_t(\psi) - C_t(\psi) \right] d\psi, \tag{19}
\]

and

\[
\int_{\Psi} N_t(\psi) d\psi = 1. \tag{20}
\]

It is said to be feasible if it satisfies the resource constraints (19) and (20), the limited commitment constraint for managers in equation (9), and the limited commitment constraint for the shareholder in equation (16). It is said to be constrained efficient if there is no other feasible allocation that makes all shareholders and managers weakly better off and at least a subset of subgroup of them positive measure strictly better off.

**Proposition 1. (Mechanism Design Interpretation of the Model)**

Let \( \{ \bar{C}, \bar{K} \}, \{ \{ \bar{C}_t(\psi), \bar{I}_t(\psi), \bar{N}_t(\psi) \}_{t=0}^{\infty} \}_{\psi \in \Psi} \) be a stationary competitive equilibrium allocation that satisfies conditions 1-7 in subsection II.E. Then it is constrained efficient.

*Proof.* See Appendix A.

We analyze the optimal contract with limited commitment in the following sections. We start by presenting the implications and limitations of the first-best economy with full commitment. We then sequentially introduce limited commitment on the manager side and on the shareholder side and discuss how they improve upon the first-best model.

**III First Best**

Firms’ maximization problem in the first best case is as described in Section II.D except that constraints (8) and (9) do not apply. The profit maximization problem is separable and can
be solved in two steps. The first is to maximize the total value of the firm

$$E_0 \left[ \int_0^\tau e^{-rs} \left[ AK_s - h \left( \frac{I_s}{K_s} \right) K_s \right] ds \right]$$  \hspace{1cm} (21)$$

subject to the law of motion of capital in equation (3) by choosing the optimal investment policy. In the second step, we choose \( \{C_s\}_{s=0}^\infty \) to minimize the total cost of managerial compensation:

$$E_0 \left[ \int_0^\tau e^{-rs} C_s ds \right]$$  \hspace{1cm} (22)$$

subject to the promise keeping constraint in equation (7).

In the first best case, the firm value maximization problem (21) is standard as in Hayashi (1982). The solution to the cost minimization problem is straightforward: risk aversion of the manager and the fact that the principal and the agent have identical discount rates \( r = \beta \) imply a constant consumption of the manager: \( C_t = \bar{U} \). The solution to the firm’s problem is summarized in the following proposition.

**Proposition 2. (The First-Best Case)**

The value function of the firm is given by:

$$V(K, U) = \bar{v}K - \frac{1}{r + \kappa} U,$$  \hspace{1cm} (23)$$

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

$$\hat{i} = \arg \max_i \frac{A - h(i)}{\hat{r} - \hat{i}} = \hat{r} - \sqrt{\hat{r}^2 - \frac{2}{h_0} (A - \hat{r})} \in (0, \hat{r}),$$  \hspace{1cm} (24)$$

where \( \hat{r} = \kappa + r + \delta \).

**Proof.** See Appendix B.

The term \( \bar{v}K = h'(i) K \) in equation (23) is the firm value in the neoclassical model with capital adjustment cost, and \( \frac{1}{r + \kappa} U \) is the present value of manager’s compensation. Perfect risk sharing implies a constant managerial compensation: \( C_t = \bar{U} \) for all \( t \);\(^{15}\) therefore, the present value of the cost of managerial compensation is simply given by Gordon (1959)’s formula.

Equation (24) implies that the investment-to-capital ratio is constant across all firms. This results in a Gibrat’s law in firm growth: growth rates are i.i.d. and do not depend

\(^{15}\)Recall that we normalize the utility function of the manager so that life-time utility is measured in consumption units.
on size. As a result, the model features a power law of the size distribution as in Luttmer (2007), which is summarized in the following proposition.

**Proposition 3. Power Law of Firm Size**

Given firms’ initial size, $\bar{K}$, and their optimal investment policy, $\hat{i}$, the total measure of firms in the stationary equilibrium is $\frac{1}{\kappa}$ and the total amount of capital stock is

$$K = \frac{\bar{K}}{\kappa + \delta - \hat{i}}. \quad (25)$$

Furthermore, the density of firm size distribution is given by:

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

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

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

In particular, the right tail of firm size obeys power law with exponent $\alpha_2$.

**Proof.** See Appendix C.

The first best model has several predictions. As shown in Proposition 3, this specification generates a power law in firm size. In addition, it implies an inverse relationship between the propensity of dividend payout and firm size. Both of these predictions are qualitatively consistent with the empirical evidence presented in Section I. Note that firms’ cash flow, $A K_t$, and their investment, $h(i) K_t$, are proportional to size. Because managerial pay is constant in time series and in the cross section, total dividend payout, $D_t = (A - h(i)) K_t - \bar{U}$, is high for large firms and low or even negative for small firms.\(^{16}\)

Other implications of the first best model, however, are grossly inconsistent with the data. First, it predicts a flat investment-size relationship. Second, the distribution of CEO compensation is degenerate and hence the model generates a zero elasticity of managerial compensation with respect to size and cannot account for the observed fat tail in CEO pay.

\(^{16}\)These implications continue to hold in models with limited commitment as we discuss below.
IV Limited Commitment on the Manager Side

In this section we allow for limited commitment on the manager side. Firms’ profit maximization problem is described by equations (5)-(7) and the limited commitment constraint, equation (9). The solution to the optimal contracting problem is characterized by the following proposition.

Proposition 4. Limited Commitment on the Manager Side

1. The normalized value function satisfies the following ODE on $[u_{MIN}, \infty)$

$$0 = \max_{c,i,g} \left\{ \left[ z - c - h(i) + v(u) \left[ t - r - \delta \right] \right] + uv' (u) \left[ t - (\frac{r}{1-\gamma}) \left( 1 - \left( \frac{r}{z} \right)^{\frac{1}{1-\gamma}} \right) - (i - \delta) + \frac{\gamma g^2 \sigma^2}{2} \right] + \frac{1}{2} u^2 v'' (u) (g - 1) \sigma^2 \right\}$$

with boundary conditions $\lim_{u \to u_{MIN}} (u) = -\infty$ and $\lim_{u \to \infty} \left[ v(u) - \left( \bar{v} - \frac{1}{r+\kappa} u \right) \right] = 0$.

2. Under the optimal contract,

$$C_t = \max \left\{ c(u_{MIN}) K_t^*, C_0 \right\},$$

where $c(\cdot)$ is the normalized compensation policy as defined in (15).

Proof. See Appendix D.

The optimal contract under limited commitment on the manager side alters the implications of the first best model along several dimensions. First, the continuation utility of managers stays above their outside options at all times. This implies that the normalized utility, $u_t$, is higher than $u_{MIN}$ for all $t$. We illustrate this graphically in Figure 1. Note that the normalized value function for the one-sided limited commitment case stays below that for the first best case for all $u$ and converges to the latter as $u \to \infty$.\(^{17}\)

Second, the compensation contract is downward rigid, as in Harris and Holmstrom (1982). Compensation has to increase to match the manager’s outside option whenever the participation constraint binds, and must remain constant otherwise due to risk sharing. Thus, managerial compensation is a linear function of the running maximum of firm size.\(^{18}\)

---

\(^{17}\)There is a substantial distance between the two value functions in the figure, indicating that convergence is slow.

\(^{18}\)See also Lustig, Syverson, and Van Nieuwerburgh (2011), Grochulski and Zhang (2011), and Miao and Zhang (2013).
As we show below, this feature of the model allows the power law in firm size to translate into a power law in CEO compensation.

While the optimal contracting problem here is similar to that in Albuquerque and Hopenhayn (2004), there are two important differences between our models. First, the Albuquerque and Hopenhayn (2004) model features decreasing return to scale technologies and stationary productivity shocks. As a result, firms eventually reach their optimal size, where the limited commitment constraint is not binding. Therefore, neither the distribution of firm size nor that of the managerial compensation in their model have fat right tails. In our model, firms grow without a bound due to the constant return to scale technology. In addition, because managers’ outside options are increasing in firm size, the limited commitment constraint binds even in the long run. These features of the model produce a power law in both firm size and CEO compensation. The following example considers a special case of our model, where we can explicitly derive the power law coefficient of firm size and CEO compensation.

Example 1. (Power Law in CEO Compensation)

Assume that the adjustment cost function \( h(i) \) takes the following form:

\[
    h(i) = \begin{cases} 
        i & \text{if } 0 \leq i \leq \hat{i} \\
        \infty & \text{if } i > \hat{i}
    \end{cases}
\]

Assume also that the parameters of the model satisfy

\[
    \frac{A - \hat{i}}{r + \delta + \kappa - \hat{i}} \geq 1 + \frac{1}{r + \kappa} \frac{\gamma}{a - 1} \left( \frac{a + \gamma - 1}{a} \right)^{\frac{\gamma}{a}} u_{MIN},
\]

where

\[
    a = \sqrt{\left( \frac{\hat{i} - \delta}{\sigma^2} - \frac{1}{2} \right)^2 + \frac{2 (\kappa + r)}{\sigma^2}} - \left( \frac{\hat{i} - \delta}{\sigma^2} - \frac{1}{2} \right).
\]

Then the distribution of firm size is the same as in Proposition 3. In addition, there exists \( \bar{x} \) such that the total measure of firms with CEO pay higher than \( x \) is proportional to \( x^{\alpha_2} \) for all \( x > \bar{x} \). That is, the right tail of CEO pay has the same power law coefficient as firm size.\(^{19}\)

With more general adjustment cost functions, it is not possible to obtain an explicit expression for the power law coefficient of CEO compensation. However, our simulation evidence suggests that power law in CEO compensation is a very robust feature of the model with limited commitment on the manager side.

\(^{19}\)A formal proof of the above result is available from the authors upon request.
The second key difference between our model and Albuquerque and Hopenhayn (2004) is in terms of the cross-sectional pattern of the size elasticity of CEO compensation. In Albuquerque and Hopenhayn (2004), managers are risk-neutral. As a result, they receive no payment as long as the limited commitment constraint binds, and payment policy is indeterminant once the constraint no longer binds. In contrast, in our model, managerial pay is typically smooth due to risk aversion. It stays constant when the limited commitment constraint does not bind, and increases whenever the constraint binds to match managers’ outside options. Consequently, the elasticity of CEO pay with respect to firm size is high in the right tail of size distribution, where managers’ limited commitment constraint is more likely to bind. As we show in the next section, after incorporating limited commitment on the shareholder side, our model generates a V-shaped elasticity of CEO pay with respect to firm size, as in the data.

V Two-Sided Limited Commitment

In this section, we present our full model with limited commitment on both the shareholder and the manager side.\(^{20}\) In this set-up, shareholders cannot commit to negative NPV projects and this requires that firm value stays positive at all times in all states of the world under the optimal contract.

While limited commitment on the manager side affects mainly the optimal contract for large firms where managers’ outside options are more attractive, limited commitment on the shareholder side primarily impacts the dynamics of small firms that are close to bankruptcy. Risk sharing requires that payment to the manager must stay constant whenever the constraints are not binding. Limited commitment on the shareholder side implies that managerial compensation must be cut whenever the firm value hits zero. Because a binding limited commitment constraint is inefficient, small firms invest at a higher rate when firm value is close to zero. From the shareholder’s perspective, there is little down side to investment due to limited liability, and increasing investment is optimal when firm size is small.

The properties of the optimal compensation contract are summarized in the following proposition.

**Proposition 5. (Two-Sided Limited Commitment)**

1. There exists a \(u_{\text{MAX}} > 0\) such that under the optimal contract, \(u_{\text{MIN}} \leq u_t \leq u_{\text{MAX}}\) for

\(^{20}\)The implications of the model with one-sided limited commitment on the part of the shareholder are discussed in the appendix.
all $t$ and the normalized value function, $v(u)$ satisfies the HJB equation in the interior of $[u_{MIN}, u_{MAX}]$. In addition, the limited commitment constraint on the manager or the shareholder side binds if and only if $u_t = u_{MIN}$ or $u_t = u_{MAX}$, respectively.

2. Under the optimal contract, $u_t$ is decreasing in productivity shocks.

3. The optimal compensation-to-capital ratio, $c(u_t) = \frac{C_t}{K_t}$ takes the following form:

$$\log c(u_t) = \log C_0 - \log K_t + l^{+}_t - l^{-}_t,$$

where $\{l^{+}_t, l^{-}_t\}_{t=0}^{\infty}$ are the minimum increasing processes such that $c(u_{MIN}) \leq c(u_t) \leq c(u_{MAX})$ for all $t$.

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

Proof. See Appendix D.

Parts 1 and 2 of the proposition characterize the dynamics of the normalized continuation utility. In the two sided limited commitment case, $u_t$ must stay in the interval $[u_{MIN}, u_{MAX}]$. The participation constraint in equation (9) requires $U_t \geq u_{MIN} K_t$. In normalized utility terms, it is equivalent to $u_t \geq u_{MIN}$. In addition, limited commitment on the shareholder side implies $u_t \leq u_{MAX}$. While $u_{MIN}$ is determined by the exogenous outside option of the manager, $u_{MAX}$ is endogenously determined by the boundary condition $v(u_{MAX}) = 0$.

The normalized value function for the two-sided limited commitment case is plotted in Figure 1 as the solid line. Under two-sided limited commitment, the normalized utility is bounded away from zero by $u_{MIN}$, where the limited commitment constraint on the manager side binds, and bounded from above by $u_{MAX}$, where firm value hits zero. Clearly, introducing the additional contracting friction reduces the efficiency of risk sharing and lowers the value function. As a result, the value function for the two-sided limited commitment case lies under the value functions for the first best case and for the manager-side limited commitment case.

We illustrate the dynamics of the state variable, $u_t$, under the optimal contract in Figure 3. The top panel of Figure 3 is the expected change, or the drift coefficient of $\log u$ as a function of $u$. The bottom panel is volatility, or the diffusion coefficient of $\log u$. 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 $u_{MIN}$.

\[ For illustrative purposes, in Figures 1 and 2, we assume that the marginal product of capital is the same across all economies. Hence, the comparison between the first best case and cases with limited commitment is a partial equilibrium one. In general equilibrium, fixing preference and technology parameters of the model and adding limited commitment will result in an endogenous change in the steady-state level of capital and, therefore, a different marginal product of capital. \]
and stays positive in the region close to $u_{MIN}$, indicating a tendency for $u_t$ to return to it steady state mean, $\bar{u}$, when $u_t$ is small. Similarly, the drift of $u$ achieves its minimum at $u_{MAX}$ and remains negative in regions close to $u_{MAX}$. This implies that $u_t$ tends to revert back to $\bar{u}$ when it is small. The pattern of the drift of log $u$ 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.

The diffusion coefficient of log $u$ on the Brownian motion $dB_t$ is negative but above $-1$. Note that $u_t = \frac{U_t}{K_t}$. The Brownian productivity shocks $dB_t$ affect $K_t$ directly and impact $U_t$ indirectly through the optimal contract. Perfect risk sharing (first best) implies a diffusion coefficient of $-1$: because continuation utility remains constant as $K_t$ moves with productivity shocks. In contrast, no risk sharing at all corresponds to a diffusion coefficient of 0 because continuation utility moves one for one with productivity shocks. A diffusion coefficient between $-1$ and 0 is the consequence of imperfect risk sharing: a positive productivity shock raises $K_t$ and $U_t$ simultaneously, but continuation utility $U_t$ is less sensitive to shocks. Note that the diffusion of $u_t$ tends to be zero on the boundaries $u_{MIN}$ and $u_{MAX}$, indicating that risk sharing is poor when the limited commitment constraints are binding. In this case, normalized utility returns to the interior with probability one because it becomes driven solely by the drift term (top panel).

Part 3 of Proposition 4 implies that the optimal compensation contract inherits some properties from the one sided limited commitment case. In particular, managerial compensation must stay constant whenever none of the commitment constraints binds. It increases by a minimum amount to keep the manager from defaulting when the manager’s participation constraint binds, and falls by a minimum necessary amount to prevent firm value from being negative when the bankruptcy constraint binds. Formally, the logarithm of compensation-to-capital ratio, $\log c(u_t)$ can be obtained from $\log C_0 - \log K_t$ by imposing a two-side regulator, $\{l^+_t, l^-_t\}_{t=0}^\infty$.

In Figure 4, we illustrate the sample path of a firm starting from the interior of $[u_{MIN}, u_{MAX}]$. The top panel in Figure 4 is the trajectory of the log size of the firm, $\log K_t$, and the second panel is the path of the normalized utility, $u_t$. 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, $\log C_t$. 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 lower 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 $u_{MAX}$.

\[^{22}\text{log } K_t \text{ is a Brownian motion with a drift, therefore its sample path has an unbounded variation. To illustrate the basic properties of the optimal contract we plot smooth sample paths.}\]
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). During this period, \( u_t \) stays at \( u_{MAX} \), where the bankruptcy constraint (8) binds, as shown in the second panel of Figure 4. The firm value remains at zero and does not cross over to the negative region due to reductions in managerial compensation, which keeps decreasing until the firm starts to experience positive productivity shocks. From time \( t = 2 \) to \( t = 3 \), the firm receives 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 \( C_0 \). 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 keeps decreasing, until the firm starts to receive positive productivity shocks for the next time.

In Figure 5, we plot a sample path of a firm with \( u_0 \) close to the left boundary, \( u_{MIN} \). The top panel is the realization of the log size of the firm, \( \log K_t \). The second panel is the path of the normalized utility, \( u_t \). 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, \( \log C_t \). 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 increase 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 and the size-normalized firm value, \( v(u_t) \), increase. At time 0.5, the normalized continuation utility 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 normalized 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 5). 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 of the firm, \( K_t \), reaches its previous running maximum at \( t = 5 \), at which time, the participation constraint starts to bind again, and manager’s compensation increases as a result (bottom panel).

By part 4 of the proposition, the optimal investment policy, \( i(u) \) increases with normalized continuation utility. This is shown in Figure 2 as the solid line. Note that investment rate in
both the manager-side limited commitment and the two sided limited commitment cases is an increasing function of \( u \). However, in the one-sided limited commitment case, investment rate stays below the first level, because higher investment rate increases firm size and therefore the outside option of the manager. In addition, as \( u \to \infty \), \( i(u) \to i \), which is the optimal investment level in the frictionless economy. In the two-sided limited commitment case, \( u \) is bounded by the bankruptcy point, \( u_{\text{MAX}} \). Moreover, as \( u \) increases, investment rate increases above the first best level, \( i \). It is optimal to invest at a level higher than \( i \) because a binding bankruptcy constraint is associated with inefficient risk sharing and therefore welfare reducing. Because normalized continuation utility is negatively correlated with firm size, in the model with two sided limited commitment, small firms over-invest and large firms under-invest relative to the first-best specification.

In the next section, we calibrate our model and show that two-sided limited commitment accounts for a wide range of stylized features of firms’ investment, CEO compensation, and dividend payout, including those we listed in Section I.

VI Quantitative Results

A Calibration and Simple Statistics

We calibrate the model to match key macroeconomic statistics of the post-war U.S. data. The parameter values and the targeted moments are listed in Tables 1 and 2. In calibration, we allow for aggregate economic growth by setting \( z_t = z_0 e^{\mu t} \).\(^{23}\) We choose \( \mu = 2\% \) to match the average annual growth rate in the U.S. economy. We set risk aversion at 2 and assume a zero rate of time discount. Together with the assumed growth rate our preference configuration implies a 4% rate of return, which is approximately an average of equity and bond returns in the data.\(^{24}\)

On the production side, following the macro literature, we set capital share, \( \alpha \), at 0.33. We choose capital adjustment cost parameter \( \phi = 5 \) to match the average market-to-book ratio (Tobin’s Q) in the data of about 1.7. We calibrate the firm death rate, \( \kappa \), at 5% to be consistent with the average exit rate. We choose capital depreciation \( \delta = 7\% \) that, together with the death rate, implies a 12% effective depreciation rate of capital. We set the volatility parameter \( \sigma = 35\% \) to match the average volatility of firms’ sale growth rates in our data set. To ensure balanced growth, the cost of building new firms is assumed to be homogenous.

\(^{23}\)The solution of the model with a deterministic aggregate growth is a straightforward extension of the stationary model discussed in earlier sections. Details of the extended model are available upon request.

\(^{24}\)Our calibration requires a discount rate of zero to avoid the well-known risk-free rate puzzle (Weil (1989)). Kocherlakota (1996) discusses the difficulty of matching asset return moments with standard expected utility.
in aggregate capital. In particular, the cost function is specified as:

\[ H(K_t, K_t) = \frac{\psi_0}{1 + \psi_1} \left( \frac{K_t}{K_t} \right)^{1+\psi_1} K_t. \]  

(28)

We choose the parameters \(\psi_0\) and \(\psi_1\) so that the initial size of the generation-0 firms is normalized to one and the profit of setting up a new firm is zero.

We calibrate the equilibrium marginal product of capital, \(A\), to match the average investment-to-output ratio of 20% without explicitly specifying the initial level of aggregate productivity, \(z_0\). Note that along the balanced growth path, aggregate investment must be related to aggregate capital stock by \(I_t = (\mu + \kappa + \delta)K_t\). The output-to-investment ratio is then given by:

\[ \frac{Y_t}{I_t} = \frac{z_t K_t^\alpha}{(\mu + \kappa + \delta) K_t} = \frac{1}{\mu + \kappa + \delta} z_t K_t^{\alpha-1}. \]

(29)

Therefore, the equilibrium marginal product of capital can be calibrated directly to match the aggregate investment-to-output ratio:

\[ A = \alpha z_t K_t^{\alpha-1} = \alpha (\mu + \kappa + \delta) \frac{Y_t}{I_t}. \]

(30)

The magnitude of agency frictions in our model is determined by two parameters, managers’ reservation utility, and the fraction of firms’ asset that managers can abscond with upon default, \(\theta\). To guarantee balance growth, we assume that the reservation utility of managers born at time \(t\) is \(e^{\mu t} \bar{U}\). We calibrate the parameter \(\bar{U} = 0.1107\) to roughly match the average CEO pay-to-output ratio in our sample. We set \(\theta = 1\) to follow Kehoe and Levine (1993)’s specification of the default technology. That is, upon default, managers can take off with 100% of firm’s capital stock but are permanently excluded from the financial market.

After solving the model numerically, we simulate it out for 300 years, and discard the first 100 years of data. The model is continuous time, and all quantities are aggregated to the annual level. Our simulated sample consists of two million firms, which allows us to obtain tight estimates of various quantities of interest. Below, we present the cross-sectional implications of our simulated model and discuss its ability to account for the key characteristics of the empirical distributions of firm size, CEO compensation, investment and dividend policies. The description and details of the empirical data are provided in the appendix.
B Power Law in Size and CEO Compensation

It has been shown in the literature that firm size follows a power law distribution (for example, Axtell (2001), Gabaix (2009), and Luttmer (2007)). We confirm this evidence and show that the empirical distribution of CEO compensation is also fat-tailed. We describe the model-implied distribution of size and managerial compensation and discuss how well our model can account for the right tail characteristics of the cross-sectional data.

Following Luttmer (2007) and Gabaix (2009), we use the following parametrization of power law. A distribution of random variable $X$ obeys power law if its density is governed by

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

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

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

The parameter $\zeta$ is called the power-law exponent.

Table 3 shows year-by-year estimates of the power-law parameter for firm size and CEO compensation. We measure size by the number of employees and estimate the slope of the right tail using data on the top 300 firms in a given year.\(^{25}\) Both firm size and CEO compensation conform to a power-law distribution with an average exponent of 1.1 and 1.7, respectively. Our estimates for size distribution are consistent with Luttmer (2007), who reports a power-law exponent of 1.07 using the total number of employees from the U.S. Census data. Gabaix and Landier (2008) measure firm size by market value and find comparable estimates of around one.

Consistent with the data, our calibrated model implies a power-law distribution of firm size with an exponent of about 1.09. Figure 6 provides a visual comparison of the tail behavior of the empirical and the model-implied distributions. 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). Under power law, the log-log plot is a straight line with a slope equal to the negative exponent. As the figure shows, the model-implied slope matches quite well the slope observed in the data.

The right tail of CEO compensation in the model is distributed in virtually the same way as the that of firm size, with a power-law parameter of 1.09. Intuitively, in the model, CEO compensation in large firms is roughly a linear function of the running maximum of firm

\(^{25}\)It is common in the literature to fix the mass of the right tail to a pre-specified number of firms. We follow this tradition in Table 3. In the appendix, we present estimates of both parameters of power-law distribution (the exponent and the right-tail cutoff) and formally test the power-law null.
Because growth dynamics generate power law in the right tail of size distribution, the running maximum of the same process obeys a similar power law. That is, the optimal contract under limited commitment makes the power law in firm size translate into a similar power law in CEO compensation. Thus, our model provides an alternative explanation of power law in CEO compensation to the assortative matching model of Gabaix and Landier (2008).

Note that because the elasticity of managers’ outside option with respect to firm size is assumed to be one, our model implies a fatter right tail of managerial compensation relative to the data. We illustrate the fit of the model in Figure 7 by presenting the complementary cumulative distribution function of CEO pay, side by side in the data and in the model. Both distributional plots are approximately linear on the log-log scale, but the model-implied slope is less steep compared with the data. Generalizing the outside option will potentially allow the model to better fit the right tail of the empirical distribution of CEO compensation.

C Cross-Sectional Characteristics of CEO Compensation

The cross-sectional distribution of CEO compensation is characterized by several stylized features. Generally, managerial compensation tends to increase with firm size. However, size elasticity of CEO pay is not uniform and varies considerably with size.

As Table 4 shows, the average elasticity of CEO pay with respect to firm size in the data is about 0.32 and is largely similar across alternative measures of size. We find that the dependance of managerial pay on size is mostly driven by small and large firms, and is considerably weaker across medium-sized firms. The top panel of Figure 8 illustrates the variation in size elasticity of CEO compensation across five size-sorted portfolios. We consider four measures of size: market capitalization (stars), the number of employees (squares), capital (circles) and book value of assets (triangles). For all size measures, the elasticity is V-shaped — CEO compensation is more sensitive to size for firms in the left and in the right tails of size distribution.

The V-shape in size elasticity turns out to be a signature characteristic of our two-sided limited commitment model. In the model, CEO compensation varies with firm size either due to a binding bankruptcy constraint or a binding participation constraint. Because small firms tend to be close to the bankruptcy point, and large firms tend to face a binding participation constraint, CEO compensation is more sensitive to size for very small and very large firms. The bottom panel of Figure 8 shows the cross-sectional variation in size elasticity implied by

---

26 In fact, as we show in Appendix E, in a model with limited commitment on manager side only, CEO compensation is exactly a linear function of the running maximum of firm size.

27 Others report similar estimates of around one-third (see Gabaix (2009) for a survey of the literature).
our model. Similar to the data, the model-implied elasticity is close to zero for medium-sized firms, and is quantitatively large for firms in the bottom and top size quintiles. Quantitatively, size elasticity of CEO compensation is equal to 0.99, 0.02 and 0.43 for firms in the first, third and fifth size sorted portfolios, respectively. On average, the elasticity of CEO compensation to firm size predicted by the model is about 0.24, which is roughly similar to the empirical estimate.

We also find that controlling for firm size, the level of CEO compensation varies significantly with firm age. Table 5 reports the average level of CEO compensation for size and age sorted portfolios. We construct portfolios by sorting firms into three size groups, and dividing each size bin into three age-sorted portfolios. We find that in the data (shown in the top panel), the level of managerial compensation declines with firm age for small and medium-sized portfolios, but increases with age for large firms. On average, across small and medium-sized firms, CEOs in old firms earn about 40% less compared with managers in young firms. In contrast, managerial compensation in old big firms is about 30% larger than that in young big firms.

As shown in the bottom panel of Table 5, our model implies similar cross-sectional trends. To understand these patterns, it is important to note that CEO compensation in the model is history dependent and that firms, on average, grow. Consider first two small firms of different generations: a young and an old. To end up small, the old firm must have experienced a lot of negative productivity shocks, sufficient to offset its deterministic growth. Thus, the old firm is likely to have spent much more time in the bankruptcy region and is likely to have cut CEO compensation more frequently than the young one. As a result, across small firms, the model-implied CEO compensation in old firms is on average 45% lower than that in young firms. The situation is reversed for the large size cohort. While small old firms in the model represent losers, large old firms represent winners – these are firms that have experienced long-run growth and have come to face a binding participation constraint more often than their young counterparts. Consequently, consistent with the data, among large firms, the model-implied CEO compensation is increasing in age.

Table 6 shows the cross-sectional variation of CEO compensation as a percentage of firm size in the data (Panel A) and in the model (Panel B). Empirically, controlling for age, CEO pay-to-capital ratio declines with firm size. For example, among young firms, the ratio falls from 12.3% for small firms to 1.4% for large firms; among old firms, it declines from 4.2% to about 0.8% as size increases. Similarly, controlling for size, managerial compensation as a fraction of firm capital tends to decrease with age. The model features a similar cross-sectional pattern.

In the model, the negative relationship between CEO pay-to-capital ratio and firm size is
the implication of (imperfect) risk sharing under the optimal contract. Risk sharing reduces sensitivity of CEO compensation to productivity shocks. Therefore, as firms become large, CEO compensation as a fraction of their size declines. Also, as discussed above, age in the model is slightly negatively correlated with size for small firms and is positively correlated with size for large firms. As a result, in the cross-section of small firms, the pattern in CEO pay-to-capital ratio across age mimics the pattern in the level of CEO pay; for large firms, the cross-sectional pattern in CEO compensation-to-capital ratio is driven mostly by size. In all, similar to the data, the model-implied ratio of CEO compensation to capital is declining with age for all size cohorts.

D Cross-Sectional Distribution of Investment

The inverse relationship between investment (and more generally growth rate) and size is one of the stylized features of the cross-sectional data – small firms, on average, invest at a much higher rate compared with large firms. Panel A of Figure 9 shows the empirical distribution of investment-to-capital ratio across 20 size-sorted portfolios. The average investment rate of the smallest percentile in the data is about 18%, while firms in the top portfolio have an average investment-to-capital ratio of only 10%. This evidence is consistent with Evans (1987) and Hall (1987) who document a negative relationship between firm growth and firm size. In our model, realized growth rates depend on both investment rate and productivity shocks but firms’ expected growth is a function of investment rate only. Hence, differences in investment rates across firms translate directly into differences in average growth rates. We focus on investment rates because their measurement is less prone to selection biases.28

The distribution of investment rates implied by our model is plotted in the bottom panel of Figure 9. Consistent with the data, investment rate in the model monotonically declines from about 25% for small firms to almost 13% for firms in the very right tail of size distribution. As discussed above, limited commitment on the shareholder side makes it optimal for small firms to invest at a high rate to avoid sliding into bankruptcy. Thus, the rate of investment in the model is inversely related to firm size.

Table 7 characterizes the empirical distribution of investment rates across both size and age. For all age cohorts, investment rate in the data declines with size (Panel A). Similarly, controlling for size, average investment-to-capital ratio decreases with age. Quantitatively, per each unit of capital, small young firms invest twice more relative to large old firms. This evidence is consistent with findings in Davis, Haltiwanger, and Schuh (1996) and Cooley

28Observed 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 relative to large firms.
and Quadrini (2001), who show that, controlling for the other characteristic, firm growth is negatively correlated with size and age.

The quantitative implications of our model for the cross-sectional distribution of investment rates across size and age, presented in Panel B of Table 7, are largely consistent with the data. Controlling for age, the model-implied investment rate is a decreasing function of firm size. Heterogeneity in firms’ investment policy in the same age group reveals the key 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 productivity shocks. Controlling for age, large firms are those that received a good deal of positive productivity shocks; small firms, in contrast, are firms that were unlucky. To prevent bankruptcy and rebuild their capital, small firms choose much more aggressive investment policies compared with large firms. As a result, investment rate in the model decreases with size for all age cohorts and matches the pattern documented in the data.

Further, consistent with the data, the model-implied investment rate declines with age, especially for small firms. While all small firms invest at a high rate, younger ones have particularly strong incentives for investment and rapid growth because the managers’ equity stake in these firms is the highest. As firms grow older, the managers’ equity share tends to decline and so does their investment rate. This trend is also present among medium-sized and large firms, although it becomes less pronounced as the sensitivity of investment policy to the state variable (normalized utility) decreases outside the bankruptcy region.

E Cross-Sectional Characteristics of Payout Policy

Empirically, small firms are less likely to pay out dividends compared with large firms. Panel A of Figure 10 illustrates this evidence by plotting the fraction of dividend- and/or interest-paying firms across 20 size sorted portfolios. As the figure shows, the vast majority of small firms makes neither type of payments. In fact, only 10% of firms in the bottom size percentile pay dividends or make any interest payments. In contrast, almost 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 almost 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 pattern observed in the data. Small firms in our model invest at a high rate and tend not to pay out dividends, whereas large firms invest less and pay out more frequently (effectively, most of the time).
The model is also able to account for a strong positive correlation between dividend yields and firms’ size. On average, dividend yields of small, median and large firms in the data are about 0.8%, 1.6% and 2.2% respectively. The model-implied dividend yields are also monotonically increasing in size. In fact, the net payout of small firms in the model is virtually zero as they use all available funds to accelerate investment and growth. Further, in the data, the positive relationship between dividend yields and size persists after controlling for firms’ age, and conditional on size, older firms pay out more than younger ones. The model is fully consistent with the observed heterogeneity in dividend yields across double-sorted portfolios – in the model, controlling for the other characteristic, dividend yields increase with size and age.

**VII 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.
Appendix

A The Mechanism Design Interpretation of the Model

Proof of Proposition 1 Our goal is to show that competitive equilibrium allocations subject to the limited commitment constraints, (8) and (9), as defined in Section II.E are constrained efficient. The difficulty is that the limited commitment constraint on shareholder side (9) contains equilibrium prices and is not equivalent to (16) in general. To show that the equilibrium allocation defined in Section II.E is constrained efficient, we proceed in two steps. We first define a notion of competitive equilibrium with limited commitment where all constraints are stated in terms of allocations. We show that competitive equilibrium with limited commitment is constrained efficient. In the second step, we show that the equilibrium allocation defined in Section II.E satisfies our requirement for competitive equilibrium with limited commitment.

We first set up some notations. For all \( \psi \in \Psi \), we use \( t(\psi) \) to denote the stopping time at which manager \( \psi \) arrives at the economy, and use \( \tau(\psi) \) to denote the stopping time at which manager \( \psi \) exits the economy. The definition of competitive equilibrium with limited commitment is stated as follows.

Definition 2. (Competitive Equilibrium with Limited Commitment)

An allocation \( \{ \tilde{C}_t, \tilde{K}_t \}_{t=0}^{\infty}; \left\{ \left\{ \tilde{C}_t(\psi), \tilde{I}_t(\psi), \tilde{N}_t(\psi) \right\}_{t=0}^{\infty} \right\}_{\psi \in \Psi} \) is said to be a competitive equilibrium with limited commitment under the supporting price system \( \{ r_t, W_t \}_{t=0}^{\infty} \) and managers utility frontier, \( \{ \tilde{U}(\psi) \}_{\psi \in \Psi} \) if:

1. Firms of all generations maximize profit. This include:
   i) For all initial generation firms, the allocation \( \{ \tilde{C}_t(\psi), \tilde{I}_t(\psi), \tilde{N}_t(\psi) \}_{t=0}^{\infty} \) maximizes shareholder’s profit, (5), subject to: 1) the initial condition of \( K_0(\psi) \); 2) the participation constraint (7); 3) The limited commitment constraint (9) and (16).
   ii) For all firms constructed at time \( t \), the allocation \( \tilde{K}_t; \left\{ \tilde{C}_{t+s}(\psi), \tilde{I}_{t+s}(\psi), \tilde{N}_{t+s}(\psi) \right\}_{s=0}^{\infty} \) maximizes shareholders’ profit,

\[
E_t \left[ \int_0^{\infty} e^{-rs} \left[ K_{t+s}^{\alpha}(\psi) (z N_{t+s}(\psi))^{1-\alpha} - W_{t+s} N_{t+s}(\psi) - h \left( \frac{I_{t+s}(\psi)}{I_{t+s}(\psi)} \right) - C_{t+s}(\psi) \right] ds \right] K_t - H(K_t),
\]

subject to 1) the participation constraint (7); 2) The limited commitment constraint (9) and (16).

2. The consumption process \( \{ \tilde{C}_t \}_{t=0}^{\infty} \) maximizes the representative shareholder’s utility subject to the budget constraint:

\[
\int_0^{\infty} e^{-r_t} C_t dt = PV + \int_0^{\infty} e^{-r_t} W_t dt,
\]

\[
\int_0^{\infty} e^{-r_t} C_t dt = PV + \int_0^{\infty} e^{-r_t} W_t dt,
\]
where we denote the present value of the profit from firm $\psi$ to be $PV(\psi)$, and $PV$ the total profit from all firms:

$$PV = \int_{\psi} e^{-r_t(t)} PV(\psi) \, d\psi.$$ 

3. The resource constraints, (19) and (20) are satisfied.

In what follows, we show that competitive equilibrium allocations with limited commitment as defined above are constrained efficient. This is formally stated in Lemma 1 below.

**Lemma 1.** Let $\{\tilde{C}_t, \tilde{K}_t\}_{t=0}^{\infty}$; $\{\tilde{C}_t(\psi), \tilde{I}_t(\psi), \tilde{N}_t(\psi)\}_{t=0}^{\infty} \not\in \psi \in \Psi$ be a competitive equilibrium allocation with limited commitment (with constraints (9) and (16)), then it is constrained efficient.

**Proof.** Let $\{C_t, K_t\}_{t=0}^{\infty}$; $\{C_t(\psi), \tilde{I}_t(\psi), \tilde{N}_t(\psi)\}_{t=0}^{\infty} \not\in \psi \in \Psi$ be a competitive equilibrium allocation with limited commitment (subject to constraints (9) and (16)). Let $r_t, W_t_{t=0}^{\infty}$ be the supporting price system. We assume that the proposed allocation is Pareto dominated by an alternative feasible allocation $\{C_t^*, K_t^*\}_{t=0}^{\infty}$; $\{C_t^*(\psi), I_t^*(\psi), N_t^*(\psi)\}_{t=0}^{\infty} \not\in \psi \in \Psi$ and prove by contradiction.

Because all preferences are strictly monotone, we can without loss generality assume that all managers are weakly better off and the representative shareholder is strictly better off, that is,

$$E\left[\int_0^{\infty} e^{-\beta s} \frac{1}{1-\gamma} C_{s+1-\gamma} ds\right] > E\left[\int_0^{\infty} e^{-\beta s} \frac{1}{1-\gamma} \tilde{C}_{s+1-\gamma} ds\right].$$

This implies that the "*" allocation must violate the budget constraint for the shareholder, that is,

$$\int_0^{\infty} e^{-r_t} C_t^* \, dt > \widetilde{PV} + \int_0^{\infty} e^{-r_t} W_t \, dt,$$

where $\widetilde{PV}$ denotes the total profit from all firms under the "^*" allocation. Let $PV^*$ denote the total profit of all firms under the "*" allocation, then profit maximization implies $\widetilde{PV} \geq PV^*$. Therefore, inequality (31) implies

$$\int_0^{\infty} e^{-r_t} C_t^* \, dt > PV^* + \int_0^{\infty} e^{-r_t} W_t \, dt.$$ 

Using the definition of $PV^*$, the above implies

$$\int_0^{\infty} e^{-r_t} C_t^* \, dt > \int_\Psi E[\int_0^{\infty} e^{-r_t} \left[K_t^* (\psi) (z N_t (\psi))^{1-\alpha} - W_t N_t (\psi) - h \left( \frac{I_t (\psi)}{K_t (\psi)} \right) K_t (\psi) - C_t (\psi) \right] dt \, d\psi$$

$$- \int_0^{\infty} e^{-r_t} H(\tilde{K}_t) \, dt + \int_0^{\infty} e^{-r_t} W_t \, dt.$$ 

Using the resource constraint (20), we have:

$$\int_0^{\infty} e^{-r_t} \left[C_t^* + H(\tilde{K}_t)\right] \, dt > \int_\Psi E[\int_0^{\infty} e^{-r_t} \left[K_t^* (\psi) (z N_t (\psi))^{1-\alpha} - h \left( \frac{I_t (\psi)}{K_t (\psi)} \right) K_t (\psi) - C_t (\psi) \right] dt \, d\psi,$$

which contradicts the resource constraint (19) for the "*" allocation. \qed
We are now ready to prove Proposition 1. As in the main text of the paper, we focus on stationary equilibria, where all prices and aggregate quantities are constant. We denote such an equilibrium as \( \{ \hat{C}, \hat{K} \} ; \left\{ \{ C_t (\psi) , I_t (\psi) , N_t (\psi) \} \right\}_{t=0}^{\infty} \) \( \psi \in \Psi \). Note that in a stationary equilibrium we must have \( r = \beta \). Finally, we use \( \hat{W} \) to denote the equilibrium wage.

**Proof.** Let \( \hat{W} \) be the equilibrium wage. In light of Lemma 1, we only need to show that \( \{ \hat{C}, \hat{K} \} \), \( \left\{ \{ \hat{C}_t (\psi) , \hat{I}_t (\psi) , \hat{N}_t (\psi) \} \right\}_{t=0}^{\infty} \) \( \psi \in \Psi \) constitutes a competitive equilibrium with limited commitment with constraints (9) and (16). It is enough to show: i) The proposed allocation, \( \left\{ \{ \hat{C}_t (\psi) , \hat{I}_t (\psi) , \hat{N}_t (\psi) \} \right\}_{t=0}^{\infty} \) \( \psi \in \Psi \) satisfies constraint (16); ii) Given \( C_t = \hat{C} \) for all \( t \), any allocation, \( \left\{ \{ C_t (\psi) , I_t (\psi) , N_t (\psi) \} \right\}_{t=0}^{\infty} \) \( \psi \in \Psi \) that satisfies constraints (7), (9) and (16) yields a lower profit than the proposed allocation, \( \left\{ \{ \hat{C}_t (\psi) , \hat{I}_t (\psi) , \hat{N}_t (\psi) \} \right\}_{t=0}^{\infty} \) \( \psi \in \Psi \). To this end, it is enough to show that \( \left\{ \{ C_t (\psi) , I_t (\psi) , N_t (\psi) \} \right\}_{t=0}^{\infty} \) \( \psi \in \Psi \) also satisfies (8).

We first prove i). Let \( \{ \hat{C}_{t+s} \}_{s=0}^{\infty} \) be any continuation consumption plan that satisfies (17) and (18). By the concavity of the utility function of the representative shareholder, we have

\[
E_t \left[ \int_0^\infty e^{-\beta s} \frac{1}{1-\gamma} \hat{C}^{1-\gamma} - 1 \frac{1}{1-\gamma} \hat{C}^{1-\gamma} ds \right] \geq E_t \left[ \int_0^\infty e^{-\beta s} C^{-\gamma} (\hat{C} - \hat{C}_{t+s}) ds \right].
\]  

(32)

Because \( \{ \hat{C}_{t+s} \}_{s=0}^{\infty} \) satisfies the resource constraint (19), and \( \{ \hat{C}_{t+s} \}_{s=0}^{\infty} \) satisfies (17), we have

\[
\hat{C}_{t+s} - \hat{C}_{t+s} = \int_\Psi \left[ \hat{K}_{t+s}^\alpha (\psi) \left( z \hat{N}_{t+s} (\psi) \right)^{1-\alpha} - h \left( \frac{\hat{I}_{t+s} (\psi)}{\hat{K}_{t+s} (\psi)} \right) \hat{K}_{t+s} (\psi) - \hat{C}_{t+s} (\psi) \right] d\psi
\]

\[
- \int_\Psi \left[ \hat{K}_{t+s}^\alpha (\psi) \left( z \hat{N}_{t+s} (\psi) \right)^{1-\alpha} - h \left( \frac{\hat{I}_{t+s} (\psi)}{\hat{K}_{t+s} (\psi)} \right) \hat{K}_{t+s} (\psi) - \hat{C}_{t+s} (\psi) \right] d\psi.
\]

Because labor is allocated optimally under the \( \ast \ast \) allocation,

\[
\int_\Psi \hat{K}_{t+s}^\alpha (\psi) \left( z \hat{N}_{t+s} (\psi) \right)^{1-\alpha} d\psi = z^{1-\alpha} \left( \int_\Psi \hat{K}_{t+s} (\psi) d\psi \right)^\alpha,
\]

by lemma 2 below. Also,

\[
\int_\Psi \hat{K}_{t+s}^\alpha (\psi) \left( z \hat{N}_{t+s} (\psi) \right)^{1-\alpha} d\psi \leq z^{1-\alpha} \left( \int_\Psi \hat{K}_{t+s} (\psi) d\psi \right)^\alpha.
\]

32
The above results imply
\[
\int_{\Psi} \tilde{K}_{t+s}^{\alpha} (\psi) \left( z \tilde{N}_{t+s} (\psi) \right)^{1-\alpha} \, d\psi - \int_{\Psi} \tilde{K}_{t+s}^{\alpha} (\psi) \left( z \tilde{N}_{t+s} (\psi) \right)^{1-\alpha} \, d\psi \\
\geq z^{1-\alpha} \left( \int_{\Psi} \tilde{K}_{t+s} (\psi) \, d\psi \right)^{\alpha} - z^{1-\alpha} \left( \int_{\Psi} \tilde{K}_{t+s} (\psi) \, d\psi \right)^{\alpha} \\
\geq z^{1-\alpha} \alpha \left( \int_{\Psi} \tilde{K}_{t+s} (\psi) \, d\psi \right)^{\alpha-1} \left[ \int_{\Psi} \tilde{K}_{t+s} (\psi) \, d\psi - \int_{\Psi} \tilde{K}_{t+s} (\psi) \, d\psi \right] \\
= A \int_{\Psi \setminus \tilde{\Psi}} \tilde{K}_{t+s} (\psi) \, d\psi,
\]
where the second inequality is due to concavity of the production function. Therefore,
\[
\tilde{C}_{t+s} - \tilde{C}_{t+s} \geq \int_{\Psi \setminus \tilde{\Psi}} \left[ A \tilde{K}_{t+s} (\psi) - h \left( \frac{\tilde{I}_{t+s} (\psi)}{\tilde{K}_{t+s} (\psi)} \right) \tilde{K}_{t+s} (\psi) \right] \, d\psi \geq 0,
\]
(33)

Because the proposed allocation satisfies the limited commitment constraint (8), that is,
\[
E_t \left[ \int_0^\infty e^{-rs} \left[ A \tilde{K}_{t+s} (\psi) - h \left( \frac{\tilde{I}_{t+s} (\psi)}{\tilde{K}_{t+s} (\psi)} \right) \tilde{K}_{t+s} (\psi) - \tilde{C}_{t+s} (\psi) \right] \, ds \right] \geq 0 \text{ for all } \psi.
\]

To summarize, (32) and (33) together imply
\[
E_t \left[ \int_0^\infty e^{-\beta s} \frac{1}{1-\gamma} \tilde{C}_{t+s}^{1-\gamma} - \frac{1}{1-\gamma} \tilde{C}_{t+s}^{1-\gamma} \, ds \right] \\
\geq \int_{\Psi \setminus \tilde{\Psi}} C^{-\gamma} E_t \left[ \int_0^\infty e^{-rs} \left( A \tilde{K}_{t+s} (\psi) - h \left( \frac{\tilde{I}_{t+s} (\psi)}{\tilde{K}_{t+s} (\psi)} \right) \tilde{K}_{t+s} (\psi) - \tilde{C}_{t+s} (\psi) \right) \, ds \right] \, d\psi \\
\geq 0,
\]
as needed.

To prove ii), given \( C_t = \tilde{C} \) for all \( t \). Let \( \left\{ \tilde{C}_{t+\psi} (\psi), \tilde{I}_{t+\psi} (\psi), \tilde{N}_{t+\psi} (\psi) \right\}_{s=0}^{\infty} \}_{\psi \in \Psi} \) be any allocation that satisfies (16). We need to show that \( \left\{ \tilde{C}_{t+\psi} (\psi), \tilde{I}_{t+\psi} (\psi), \tilde{N}_{t+\psi} (\psi) \right\}_{s=0}^{\infty} \}_{\psi \in \Psi} \) also satisfies (8), that is, for almost every \( \psi \),
\[
E_t \left[ \int_0^\infty e^{-rs} \left( A \tilde{K}_{t+s} (\psi) - h \left( \frac{\tilde{I}_{t+s} (\psi)}{\tilde{K}_{t+s} (\psi)} \right) \tilde{K}_{t+s} (\psi) - \tilde{C}_{t+s} (\psi) \right) \, ds \right] \geq 0. \quad (34)
\]
We assume that (34) is not true and prove by contradiction.

Note that the condition expectation in (34) depends only on state variables \( K, U \). If (34) is not true, then there exists \( \left( \hat{K}, \hat{U} \right) \) and \( \delta, \varepsilon > 0 \) such that
\[
E \left[ \int_0^\infty e^{-rs} \left( A \tilde{K}_{t+s} (\psi) - h \left( \frac{\tilde{I}_{t+s} (\psi)}{\tilde{K}_{t+s} (\psi)} \right) \tilde{K}_{t+s} (\psi) - \tilde{C}_{t+s} (\psi) \right) \, ds \bigg| K, U \right] < -\delta.
\]
for all \((K, U) \in B_\varepsilon(K, \hat{U})\), where \(B_\varepsilon(K, \hat{U})\) is an open ball centered at \((K, \hat{U})\) with measure \(\varepsilon\).

To derive a contradiction, we construct an alternative consumption plan, \(\{\tilde{C}_{t+s}\}_{s=0}^\infty\) that satisfies conditions (17), (18) and yields a strictly higher utility for the shareholder. Define \(\tilde{C}_s = \tilde{C}_{t+s}\) for all \(s \leq t\), and

\[
\tilde{C}_{t+s} = \max \left\{ \int_{\Phi\setminus B_\varepsilon(\psi)} \tilde{K}_{t+s}^\alpha(\psi) \left( z \tilde{N}_{t+s}(\psi) \right)^{1-\alpha} - h \left( \frac{\tilde{t}_{t+s}(\psi)}{\tilde{K}_{t+s}(\psi)} \right) \tilde{K}_{t+s}(\psi) - \tilde{C}_{t+s}(\psi) \right\} d\psi - H(\tilde{K}_{t+s})
\]

subject to : \(\int_{\Phi\setminus B_\varepsilon(\psi)} \tilde{N}_{t+s}(\psi) d\psi = 1\),

Note that \(\tilde{C}_t \to \tilde{C}\) for all \(t\) as \(\varepsilon \to 0\). Therefore, for \(\varepsilon\) close to zero,

\[
E \left[ \int_0^\infty e^{-\beta s} \left( \frac{1}{1-\gamma} \tilde{C}_{1-\gamma} - \frac{1}{1-\gamma} \tilde{C}_{t+s}^{1-\gamma} \right) ds \right] \mid K, U
\]

(35)

\[
E \left[ \int_0^\infty e^{-\beta s} \left\{ \tilde{C}^{-\gamma} (\tilde{C} - \tilde{C}_{t+s}) + o(\tilde{C} - \tilde{C}_{t+s}) \right\} ds \right] \mid K, U
\]

(36)

By the definition of \(\{\tilde{C}_t\}_{t=0}^\infty\),

\[
\tilde{C} - \tilde{C}_{t+s} = \int_{\Phi} \left[ \tilde{K}_{t+s}^\alpha(\psi) \left( z \tilde{N}_{t+s}(\psi) \right)^{1-\alpha} - h \left( \frac{\tilde{t}_{t+s}(\psi)}{\tilde{K}_{t+s}(\psi)} \right) \tilde{K}_{t+s}(\psi) - \tilde{C}_{t+s}(\psi) \right] d\psi
\]

\[
- \int_{\Phi\setminus B_\varepsilon(\psi)} \left[ \tilde{K}_{t+s}^\alpha(\psi) \left( z \tilde{N}_{t+s}(\psi) \right)^{1-\alpha} - h \left( \frac{\tilde{t}_{t+s}(\psi)}{\tilde{K}_{t+s}(\psi)} \right) \tilde{K}_{t+s}(\psi) - \tilde{C}_{t+s}(\psi) \right] d\psi
\]

Using lemma 2,

\[
\int_{\Phi} \tilde{K}_{t+s}^\alpha(\psi) \left( z \tilde{N}_{t+s}(\psi) \right)^{1-\alpha} d\psi - \int_{\Phi\setminus B_\varepsilon(\psi)} \tilde{K}_{t+s}^\alpha(\psi) \left( z \tilde{N}_{t+s}(\psi) \right)^{1-\alpha} d\psi
\]

\[
= z^{1-\alpha} \left( \int_{\Phi} \tilde{K}_{t+s}(\psi) d\psi \right)^\alpha - z^{1-\alpha} \left( \int_{\Phi\setminus B_\varepsilon(\psi)} \tilde{K}_{t+s}(\psi) d\psi \right)^\alpha
\]

\[
= z^{1-\alpha} \left( \int_{\Phi} \tilde{K}_{t+s}(\psi) d\psi \right)^{\alpha-1} \left( \int_{B_\varepsilon(\psi)} \tilde{K}_{t+s}(\psi) d\psi + o \left( \int_{B_\varepsilon(\psi)} \tilde{K}_{t+s}(\psi) d\psi \right) \right)
\]

\[
= \int_{B_\varepsilon(\psi)} A \tilde{K}_{t+s}(\psi) d\psi + o \left( \int_{B_\varepsilon(\psi)} \tilde{K}_{t+s}(\psi) d\psi \right)
\]

where \(A\) is the marginal product of capital. Therefore,

\[
\tilde{C} - \tilde{C}_{t+s}
\]

\[
= \int_{B_\varepsilon(\psi)} A \tilde{K}_{t+s}(\psi) - h \left( \frac{\tilde{t}_{t+s}(\psi)}{\tilde{K}_{t+s}(\psi)} \right) \tilde{K}_{t+s}(\psi) - \tilde{C}_{t+s}(\psi) \right] d\psi + o \left( \int_{B_\varepsilon(\psi)} \tilde{K}_{t+s}(\psi) d\psi \right)
\]

\[
< -\delta \varepsilon + o(\varepsilon)
\]

(37)
Combining (36) and (37), we have, for \( \varepsilon \) small enough,
\[
E \left[ \int_0^\infty e^{-\beta s} \left( \frac{1}{1-\gamma} \hat{C}_1 - \frac{1}{1-\gamma} \hat{C}_{i+s} \right) ds \right] K, U
\]
\[
< E \left[ \int_0^\infty e^{-\beta s} \left\{ \hat{C}_{-\gamma} (-\delta \varepsilon) + o(\varepsilon) \right\} ds \right]
\]
\[
< 0.
\]
This contradicts the assumption that \( \left\{ \left\{ \hat{C}_{t(\psi)+s} (\psi), \tilde{I}_{t(\psi)+s} (\psi), \tilde{N}_{t(\psi)+s} (\psi) \right\}_{s=0}^\infty \right\}_{\psi \in \Psi} \) also satisfies the limited commitment constraint, (16), as needed.

**Lemma 2.** Given \( \{ \hat{K}(\psi) \}_{\psi \in \Psi} \), let \( \Psi \) be any measurable subset of \( \Psi \), then
\[
\max_{\psi} \int_{\Psi} \hat{K}(\psi) \left( z \tilde{N}(\psi) \right)^{1-\alpha} d\psi \bigg|_{\text{subject to } \int_{\Psi} \tilde{N}(\psi) d\psi = 1} = z^{1-\alpha} \left( \int_{\Psi} \hat{K}(\psi) d\psi \right)^\alpha
\]

**Proof.** Using first order condition for the maximization problem. \( \square \)

### B Optimal Solution for the First Best Case

We make the following assumption which guarantees the first best investment to be strictly positive and the firm value to be finite.

**Assumption 1**
\[
A > \beta + \delta + \kappa \tag{38}
\]
and
\[
\frac{-1 + \sqrt{1 + 2h_0A}}{h_0} < \hat{r} < r + \delta + \kappa. \tag{39}
\]

**Proof of Proposition 2:** Homogeneity implies that the value function of the optimization problem (21) is of the form \( \bar{v}K \), where \( \bar{v} \) must solve the following HJB equation.
\[
(r + \kappa) \bar{v} = \max_i \left\{ A - h(i) + \left( i - \hat{r} \right) \bar{v} \right\}. \tag{40}
\]

The first order condition of the right hand side implies
\[
h'(i) = \bar{v} \text{ or } \bar{v} = 1 + h_0i.
\]

By plugging the above equation into (40), we have that \( i \) satisfies
\[
\frac{1}{2}h_0i^2 - h_0\hat{r}i + (A - \hat{r}) = 0.
\]

The relevant solution is
\[
i = \hat{r} - \sqrt{\hat{r}^2 - \frac{2}{h_0} (A - \hat{r})}.
\]

35
Notice under Assumption 1, \( \hat{i} \) defined above satisfies

\[
\hat{i} = \arg \max_{i \in [0, i_A]} \frac{A - h(i)}{\hat{r} - i},
\]

as needed.

C Power Law of Firm Size in the First Best Case

We first state a lemma that characterizes the stationary distribution of a Brownian motion with drift. A unit measure of particles enters the real line at \( x_0 \) at each point in time. They evaporate at a Poisson rate \( \kappa \) per unit of time. Conditioning on survival, each particle follows a Brownian motion with drift after entrance:

\[
dx_t = \mu dt + \sigma dB_t.
\]

Denote the density of the stationary distribution of the particles as \( m(x|x_0) \), we have:

**Lemma 3.** Let \( \alpha_1 > 0 \) and \( \alpha_2 < 0 \) denote the two roots of the quadratic equation, \( \kappa + \mu \alpha - \frac{1}{2} \sigma^2 \alpha^2 = 0 \).

1. The stationary distribution is given by:

\[
m(x|x_0) = \begin{cases} 
\frac{1}{\sqrt{\mu^2 + 2\kappa \sigma^2}} e^{\alpha_2 (x-x_0)} & x \geq x_0 \\
\frac{1}{\sqrt{\mu^2 + 2\kappa \sigma^2}} e^{\alpha_1 (x-x_0)} & x < x_0
\end{cases}
\]

2. Assume \( \kappa > \mu + \frac{1}{2} \sigma^2 \), then \( \alpha_2 < -1 \), and

\[
\int_{-\infty}^{\infty} e^y m(x|x_0) dy = \frac{e^{x_0}}{\kappa - (\mu + \frac{1}{2} \sigma^2)} < \infty
\]

3. The density of \( X = e^x \), denoted \( M(X|x_0) \) is given by:

\[
M(X|x_0) = \begin{cases} 
\frac{1}{\sqrt{\mu^2 + 2\kappa \sigma^2}} e^{-\alpha_2 x_0} X^{\alpha_2 - 1} & X \geq e^{x_0} \\
\frac{1}{\sqrt{\mu^2 + 2\kappa \sigma^2}} e^{-\alpha_1 x_0} X^{\alpha_1 - 1} & X < e^{x_0}
\end{cases}
\]

In particular, the right tail of \( X \) obeys power law with slope \( \alpha_2 \).

**Proof.** See Luttmer (2007).

**Proof of proposition 3:** By proposition 2, \( \log K \) is a Brownian motion:

\[
d\log K = \left( \hat{i} - \delta - \frac{1}{2} \sigma^2 \right) dt + \sigma dB_t,
\]

with initial condition \( \log \bar{K} \). Proposition 2 can be proved by applying Lemma 3 directly.
D Optimal Contraction with Limited Commitment

We first derive an expression for the manager’s outside option, \( u_{MIN}(\theta) \). We need the following assumption to guarantee that the utility level of managers is finite upon default.

**Assumption 2** If \( \gamma < 1 \), then

\[
\frac{1}{2} \sigma^2 < \frac{1}{\gamma} \left[ \delta + \frac{1}{1 - \gamma} (\theta \mathbf{A} - (\beta + \kappa)) \right],
\]

(alternatively, if \( \gamma > 1 \), then

\[
\frac{1}{2} \sigma^2 < \frac{1}{\gamma - 1} (\beta + \kappa) - \delta.
\]

Given a level of capital stock upon the manager’s default, \( K_0 \), the manager maximizes his utility by making the following optimal investment and consumption decision.

\[
\max_{\{c_t, i_t\}_{t=0}^{\infty}} \left\{ E \left[ \int_0^\tau (\beta + \kappa) e^{-\beta t} C_t^{1-\gamma} dt \right] \right\}^{1/1-\gamma}
\]

such that

\[
C_t + h(I_t/K_t) K_t = \theta \mathbf{A} K_t
\]

\[
dK_t = K_t [(i_t - \delta) dt + \sigma dB_t].
\]

We define \( c_t = \frac{C_t}{K_t} \) and \( i_t = \frac{I_t}{K_t} \). Using the expected utility representation of the same preference, for a given \( \theta \), let

\[
W(K_0; \theta) = \max_{\{c_t, i_t\}_{t=0}^{\infty}} \left\{ E \left[ \int_0^\tau (\beta + \kappa) e^{-\beta t} \frac{1}{1-\gamma} (c_t K_t)^{1-\gamma} dt \right] \right\}.
\]

Then the function \( W(\cdot; \theta) \) satisfies the following HJB differential equation.

\[
0 = \max_i (\beta + \kappa) \left[ \frac{(\theta \mathbf{A} - h(i)) K^{1-\gamma}}{1 - \gamma} - W(K; \theta) \right] + K(i - \delta) W_t(K; \theta) + \frac{1}{2} K^2 \sigma^2 W_{KK}(K; \theta). \tag{43}
\]

Homogeneity of the optimization problem implies that \( W(K; \theta) = \omega(\theta) K^{1-\gamma} \) for some function \( \omega(\cdot) \), and the optimal investment to capital ratio and the consumption to capital ratio are constant across time which we denote as \( i(\theta) \) and \( c(\theta) \). Then (43) implies

\[
0 = \max_i (\beta + \kappa) \left[ \frac{(\theta \mathbf{A} - h(i)) K^{1-\gamma}}{1 - \gamma} \right] + \left( (1 - \gamma)i - \hat{\beta} \right) \omega(\theta)
\]

(44)

with \( \hat{\beta} = \beta + \kappa + \delta (1 - \gamma) + \frac{1}{2} \gamma (1 - \gamma) \sigma^2 \). The first order condition of the maximization problem on
the right hand side implies

\[ (eta + \kappa) \frac{(\theta A - h(i))^{1-\gamma}}{(1-\gamma)\omega(\theta)} = \frac{\theta A - h(i(\theta))}{h'(i(\theta))}. \]

By plugging the above equality into (44), we have that \( i(\theta) \) satisfies \( f(i(\theta)) = 0 \), where

\[ f(x) \equiv \theta A - h(x) + h'(x) \left[ (1-\gamma)x - \beta \right]. \]  \hspace{1cm} (45)

The optimal investment policy, \( i(\theta) \) is given by the following lemma

**Lemma 4.** Under Assumption 1 and 2, \( i(\theta) \) is the unique solution to \( f(x) = 0 \) over \([0,i_A]\), where \( i_A = \frac{-1+\sqrt{1+2h_0A}}{h_0} \).

**Proof.** Since \( h'(i) > 0 \) for \( i > 0 \) and \( \theta A - h(i(\theta)) \in [0,A-h(i(\theta))] \), so \( i(\theta) \in [0,i_A] \). We prove that equation \( f(x) = 0 \) has a unique solution over \([0,i_A]\). Notice that

\[ f'(x) = -\left\{ \gamma(1+h_0x) + h_0x\beta \right\}. \]

First we consider the case with \( \gamma < 1 \). It is straight forward to show that \( f(x) < 0 \) over \((0,i_A)\) because (39) implies \( i_A < \beta + \delta + \kappa \) which in turn implies \( \beta > 0 \). In addition, (41) implies that

\[ f(0) = \theta A - \beta > 0 \]

and (38)

\[ f(i_A) = -h'(i_A) \left[ \beta - (1-\gamma)(i_A - \delta) + \frac{1}{2}\gamma(1-\gamma)\sigma^2 \right] < 0. \]

Therefore, continuity of \( f(\cdot) \) implies the desired result. If \( \gamma > 1 \), (42) implies \( \beta > 0 \) and then \( f'(i) < 0 \) over \((0,i_A)\). (42) also implies \( f(0) > 0 \) and \( f(i_A) > 0 \). Consequently, continuity of \( f(\cdot) \) implies the desired result in this case. \( \square \)

Using the above lemma, we can derive the optimal consumption and investment policies under autarky:

\[ i(\theta) = \frac{\frac{2}{\gamma} \left( \theta A - \beta \right)}{\left( 1 - \frac{h_0}{\gamma} \beta \right) + \sqrt{\left( 1 - \frac{h_0}{\gamma} \beta \right)^2 - 4h_0 \left( \frac{1}{2\gamma} - 1 \right) \left( \frac{1}{\gamma} \left( \theta A - \beta \right) \right)}} \]

and

\[ c(\theta) = \theta A - h(i(\theta)). \]

Therefore \( \omega(\theta) = [\theta A - h(\theta)]^{-\gamma} (1 + h_0i(\theta)) \) and \( u_{MIN}(\theta) = \omega(\theta)^{\frac{1}{1-\gamma}} \).

The proof for the HJB equation and the characterization of optimal compensation and investment policies under limited commitment can be adapted from Lemma 2, Lemma 3, and Proposition 3 in Ai and Li (2013).
E  Aggregation and the Summary Measure $m$

In this section, we describe a procedure to solve the cross-sectional distribution of firm characteristics. We refer the reader to Ai (2012) for the technical details and proofs. Because firm types can be summarized by $(K, u)$, the cross-section distribution of firm characteristic can be equivalently characterized by a distribution on the $(K, u)$ space. We denote this measure by $\tilde{\Phi}(K, u)$. Using $\tilde{\Phi}(K, u)$, the resource constraints (12) and (13) can be written as:

$$C + \int [c(u) K + h(i(u)) K] \, d\Phi(K, u) + H(\bar{K}) = \int K^\alpha (zn(u) K)^{1-\alpha} \tilde{\Phi}(K, u),$$  \hspace{1cm} (46)

and

$$\int n(u) K d\Phi(K, u) = 1.$$ \hspace{1cm} (47)

We define the summary measure $m$ on the space of normalized promised utility space, $(0, 1)$. Let $m(\cdot)$ be the density such that for any Borel set $B \subseteq (0, 1)$,

$$\int_B m(u) \, du = \int I_{\{u \in U, \, K \in (0, \infty)\}} d\tilde{\Phi}(K, u).$$

Using the summary measure, $m(\cdot)$, the resource constraints (46) and (47) can be written as:

$$C + \int [c(u) + h(i(u))] m(u) \, du + H(\bar{K}) = \int (zn(u))^{1-\alpha} m(u) \, du,$$

and

$$\int n(u) m(u) \, du = 1,$$

respectively. Given Proposition 5, the equilibrium can be completely characterized by one-dimensional policy functions, $(c(u), i(u), n(u))$ and the one dimension measure, $m(u)$.

In addition, as shown in Ai (2012), $m(u)$ must satisfy the following ODE:

$$[\kappa + \delta - i(u)] m(u) = \frac{1}{2} \sigma^2 \frac{d^2}{d \log u^2} \left\{ m(u) [g(u) - 1]^2 \right\} + \frac{d}{d \log u} \left\{ b(u) m(u) \right\},$$  \hspace{1cm} (48)

where

$$b(u) = \frac{\beta + \kappa}{1 - \gamma} \left[ 1 - \left( \frac{c(u)}{u} \right)^{1-\gamma} \right] - [i(u) - \delta] + \frac{1}{2} \sigma^2 \left[ g^2(u) - (g(u) - 1)^2 \right].$$

F  Limited Commitment on Shareholder Side Only

In this section, we discuss the model with limited commitment on the shareholder side only. Firms’ profit maximization problem is described by equations (5)-(7) and the limited commitment constraint, equation (8). The key properties of the policy functions are summarized in Proposition 4 below. The proof of the proposition can be adapted from Ai and Li (2013).
Proposition 6. One-Sided Limited Commitment

1. There exists a $u_{\text{MAX}} > 0$ such that under the optimal contract, $0 < u_t \leq u_{\text{MAX}}$ for all $t$, and $v(u) \geq 0$ for all $u \in (0, u_{\text{MAX}}]$. In addition, the limited commitment in equation constraint (8) binds if and only if $u_t = u_{\text{MAX}}$.

2. The optimal compensation-to-capital ratio, $c(u_t) = \frac{C_0}{K_t}$, takes the following form:

$$\log c(u_t) = \log C_0 - \log K_t - l^-_t,$$

where $\{l^-_t\}_{t=0}^\infty$ is the minimum increasing process such that $c(u_t) \leq c(u_{\text{MAX}})$ for all $t$.

3. The optimal investment rate, $i(u)$, is a strictly increasing function of $u$, and $\lim_{u \to 0} i(u) = i$, where $i$ is the optimal investment level in the friction-less case.

The optimal contract under limited commitment on the shareholder side alters the implications of the first best model along several dimensions. First, firm value remains non-negative at all times. In the first best case, firms’ value function is given by $V(K, U) = \bar{v}_K + \frac{1}{1+r} U$. Note that $U_t = U$ for all $t$ is determined by the initial condition. Therefore, a sequence of negative productivity shocks lowers the size of the firm, $K_t$ and may eventually result in negative firm value. This is the implication of perfect risk sharing. When shareholders cannot commit to negative NPV projects, perfect risk sharing is no longer feasible, and firm value must stay non-negative at all times under the optimal contract. The above proposition implies that there exists a $u_{\text{MAX}}$, such that the normalized promised utility $u_t$ stays in the interval $(0, u_{\text{MAX}}]$ at all times, and $v(u) \geq 0 \forall u \in (0, u_{\text{MAX}}]$.

Second, under the optimal contract, managerial compensation is upward rigid: it decreases with negative productivity shocks as the limited commitment constraint binds but never increases. As a result, the elasticity of CEO compensation with respect to firm size is positive but there will be no power law in CEO pay.

Intuitively, part 2 of the above proposition implies that the optimal contract provides a constant compensation to the manager whenever firm value is strictly positive and involves a minimum necessary reduction in managerial compensation to keep the firm value non-negative whenever the shareholder’s commitment constraint binds. Optimal risk sharing implies that compensation has to stay constant whenever the commitment constraint does not bind, and limited commitment requires a reduction in managerial compensation whenever firm value hits zero. Because CEO pay is upward rigid, $C_0 = C(\bar{K}, \bar{U})$ is the maximum level of CEO compensation in the economy, and managers of all firms that have not experienced a binding bankruptcy constraint are paid $C_0$. Hence, there is no power law in CEO pay.

Third, due to limited commitment on the shareholder side, firms’ investment rate is decreasing in size, which is qualitatively consistent with the data. As stated in Proposition 6, investment is increasing in $u$. Intuitively, as capital stock gets depleted and the normalized utility goes up, the firm gets closer to the bankruptcy point $u_{\text{MAX}}$, which involves imperfect risk sharing and is welfare
reducing. To avoid bankruptcy, the firm must increase investment to re-build its capital stock and move away from the distress region. Note that due to risk sharing, promised utility is less sensitive to productivity shocks than firm size. Therefore, under the optimal contract, firm size is negatively correlated with normalized utility. Hence, under limited commitment on the shareholder side, small firms invest more and feature higher growth rates compared with large firms.

G Data Description

The cross-sectional data that we use consist of US non-financial firms and come from the Center for Research in Securities Prices (CRSP). We measure executive compensation by the total compensation figure from ExecuComp database, which comprises of salary, bonuses, the value of restricted stock granted, the Black-Scholes-based value of options granted and long-term incentive payouts. Our benchmark measure of size is based on the number of firm employees. In addition, for each firm we collect market capitalization, the book value of firms’ assets, the gross value of property, plant and equipment to measure capital, capital expenditure to measure investment, and common dividends and interests on short- and long-term debt to measure total payout. We measure firm age by the number of years since the firm’s founding date. Our evidence remain virtually intact if we use the date of incorporation to define firm age. Nominal quantities are converted to real using the consumer price index compiled by the Bureau of Labor Statistics. The data are sampled on the annual frequency and cover the period from 1992 till 2011.

H Power-Law Estimates

As stated in the paper, the probability distribution function of a continuous power-law random variable $x$ is given by:

$$f(x) = k \zeta x^{-(1+\zeta)},$$  \hspace{1cm} (49)

where $k = x_{\text{min}}^\zeta$, $x_{\text{min}}$ is the lower bound of the power-law behavior, and $\zeta$ is the power-law exponent. It is common in empirical work to treat $x_{\text{min}}$ as if it were known (typically by choosing a point beyond which the empirical distribution appears approximately linear on a log-log plot) and estimate the scaling parameter $\zeta$ by maximum likelihood. However, unless the right-tail cutoff is chosen at or close to the true value, the estimates of the exponent may be significantly biased. To address this issue, we estimate both parameters by minimizing the Kolmogorov-Smirnov (KS) distance. In particular, for each potential lower bound $\tilde{x}$, we estimate the power-law exponent using the data above $\tilde{x}$ as:

$$\hat{\zeta} = N \left[ \sum_{i=1}^{N} \log \frac{x_i}{\tilde{x}} \right]^{-1}, \quad x_i \geq \tilde{x}, \quad i = 1, \ldots, N.$$  \hspace{1cm} (50)
Our estimates of $x_{\text{min}}$ and $\zeta$ is the pair that yields the power-law distribution that provides the best fit to the observed data according to the KS criteria, i.e.,

$$\{\hat{x}_{\text{min}}, \hat{\zeta}\} = \min_{\tilde{x}, \tilde{\zeta}} \left\{ \text{KS-distance} \right\} \equiv \min_{\tilde{x}, \tilde{\zeta}} \left\{ \max_{x \geq \tilde{x}} |F(x; \tilde{x}, \tilde{\zeta}) - \hat{F}(x)| \right\}, \quad (51)$$

where $F(x; \tilde{x}, \tilde{\zeta})$ is the candidate power-law cumulative distribution function and $\hat{F}(x)$ is the empirical distribution. Year-by-year estimates for size and CEO compensation are presented in Table A.1. To facilitate the comparison with Table 3, the estimates of the right-tail cutoff are quoted in terms of the number of firms in the right tail. The table also reports p-values of the Kolmogorov-Smirnov goodness-of-fit test constructed via bootstrap.

As the table shows, there is quite a lot of variation in power-law estimates across sample years. For the number of employees, the estimate of the exponent varies between 1 and 1.9, and averages about 1.26. The average of the scaling parameter of CEO compensation is about 2.1. While the mass of the right tail also varies considerably across years, interestingly, on average, the number of firms in the right tail of the size distribution is 272, which is quite close to the 300 cutoff that we use in Table 3. For CEO compensation, the right tail amounts to just above 200 firms. Overall, the goodness-of-fit test does not reject the power-law null – the p-value is above the conventional five-percent level in 65% and 75% of sample years for size and CEO compensation, respectively.
References


Table 1
Model Parameters

<table>
<thead>
<tr>
<th>Description</th>
<th>Notation</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Risk aversion</td>
<td>$\gamma$</td>
<td>2</td>
</tr>
<tr>
<td>Discount rate</td>
<td>$\beta$</td>
<td>0</td>
</tr>
<tr>
<td>Capital share</td>
<td>$\alpha$</td>
<td>0.33</td>
</tr>
<tr>
<td>Marginal product of capital</td>
<td>$A$</td>
<td>0.224</td>
</tr>
<tr>
<td>Adjustment cost</td>
<td>$\phi$</td>
<td>5</td>
</tr>
<tr>
<td>Productivity growth rate</td>
<td>$\mu$</td>
<td>2%</td>
</tr>
<tr>
<td>Volatility</td>
<td>$\sigma$</td>
<td>35%</td>
</tr>
<tr>
<td>Firm death rate</td>
<td>$\kappa$</td>
<td>5%</td>
</tr>
<tr>
<td>Capital depreciation</td>
<td>$\delta$</td>
<td>7%</td>
</tr>
</tbody>
</table>

Table 1 presents calibration of the key model parameters.
Table 2
Targeted Moments

<table>
<thead>
<tr>
<th>Moment (mean)</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aggregate Growth</td>
<td>2%</td>
</tr>
<tr>
<td>Rate of Return</td>
<td>4%</td>
</tr>
<tr>
<td>Tobin’s Q</td>
<td>1.7</td>
</tr>
<tr>
<td>Total Capital Depreciation</td>
<td>12%</td>
</tr>
<tr>
<td>Volatility of Sales Growth</td>
<td>35%</td>
</tr>
<tr>
<td>Aggregate Investment-to-Output Ratio</td>
<td>20%</td>
</tr>
<tr>
<td>Aggregate CEO Pay-to-Output Ratio</td>
<td>1.4%</td>
</tr>
</tbody>
</table>

Table 2 shows the model-implied moments that are targeted in calibration.
Table 3

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

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

Table 3 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 4
Elasticity of CEO Compensation with respect to Firm Size

<table>
<thead>
<tr>
<th>Size Measures</th>
<th>Employees</th>
<th>Market Cap</th>
<th>Capital</th>
<th>Assets</th>
</tr>
</thead>
<tbody>
<tr>
<td>Elasticity</td>
<td>0.31</td>
<td>0.33</td>
<td>0.28</td>
<td>0.37</td>
</tr>
<tr>
<td>SE</td>
<td>0.02</td>
<td>0.02</td>
<td>0.02</td>
<td>0.02</td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.20</td>
<td>0.27</td>
<td>0.20</td>
<td>0.23</td>
</tr>
</tbody>
</table>

Table 4 reports estimates of the 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”), book value of capital (“Capital”), and book value of assets (“Assets”).
Table 5 presents the distribution of CEO compensation across firms of different size and age. Panel A reports average CEO compensation in the data, Panel B presents model-implied statistics. Size in the data is measured by the number of firm employees.

<table>
<thead>
<tr>
<th>Panel A: Data</th>
<th>Age</th>
<th>Young</th>
<th>Median</th>
<th>Old</th>
</tr>
</thead>
<tbody>
<tr>
<td>Size</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Small</td>
<td>10.2</td>
<td>8.1</td>
<td>6.4</td>
<td></td>
</tr>
<tr>
<td>Median</td>
<td>13.6</td>
<td>12.4</td>
<td>7.9</td>
<td></td>
</tr>
<tr>
<td>Large</td>
<td>18.8</td>
<td>29.4</td>
<td>24.0</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Panel B: Model</th>
<th>Age</th>
<th>Young</th>
<th>Median</th>
<th>Old</th>
</tr>
</thead>
<tbody>
<tr>
<td>Size</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Small</td>
<td>4.8</td>
<td>3.4</td>
<td>2.7</td>
<td></td>
</tr>
<tr>
<td>Median</td>
<td>5.7</td>
<td>5.0</td>
<td>4.4</td>
<td></td>
</tr>
<tr>
<td>Large</td>
<td>5.8</td>
<td>7.0</td>
<td>28.6</td>
<td></td>
</tr>
</tbody>
</table>
Table 6
Distribution of CEO Pay-to-Capital across Size and Age

Panel A: Data

<table>
<thead>
<tr>
<th>Size</th>
<th>Age</th>
<th>Young</th>
<th>Median</th>
<th>Old</th>
</tr>
</thead>
<tbody>
<tr>
<td>Small</td>
<td></td>
<td>0.123</td>
<td>0.075</td>
<td>0.042</td>
</tr>
<tr>
<td>Median</td>
<td></td>
<td>0.022</td>
<td>0.023</td>
<td>0.016</td>
</tr>
<tr>
<td>Large</td>
<td></td>
<td>0.014</td>
<td>0.011</td>
<td>0.008</td>
</tr>
</tbody>
</table>

Panel B: Model

<table>
<thead>
<tr>
<th>Size</th>
<th>Age</th>
<th>Young</th>
<th>Median</th>
<th>Old</th>
</tr>
</thead>
<tbody>
<tr>
<td>Small</td>
<td></td>
<td>0.096</td>
<td>0.080</td>
<td>0.070</td>
</tr>
<tr>
<td>Median</td>
<td></td>
<td>0.046</td>
<td>0.034</td>
<td>0.029</td>
</tr>
<tr>
<td>Large</td>
<td></td>
<td>0.013</td>
<td>0.007</td>
<td>0.005</td>
</tr>
</tbody>
</table>

Table 6 presents the cross-sectional variation of the ratio of CEO compensation to capital. Panel A reports data moments, Panel B presents model-implied statistics. Size in the data is measured by the number of firm employees.
Table 7
Distribution of I/K across Size and Age

Panel A: Data

<table>
<thead>
<tr>
<th>Size</th>
<th>Young</th>
<th>Median</th>
<th>Old</th>
</tr>
</thead>
<tbody>
<tr>
<td>Small</td>
<td>0.24</td>
<td>0.19</td>
<td>0.14</td>
</tr>
<tr>
<td>Median</td>
<td>0.17</td>
<td>0.16</td>
<td>0.10</td>
</tr>
<tr>
<td>Large</td>
<td>0.16</td>
<td>0.14</td>
<td>0.11</td>
</tr>
</tbody>
</table>

Panel B: Model

<table>
<thead>
<tr>
<th>Size</th>
<th>Young</th>
<th>Median</th>
<th>Old</th>
</tr>
</thead>
<tbody>
<tr>
<td>Small</td>
<td>0.24</td>
<td>0.20</td>
<td>0.19</td>
</tr>
<tr>
<td>Median</td>
<td>0.16</td>
<td>0.15</td>
<td>0.15</td>
</tr>
<tr>
<td>Large</td>
<td>0.14</td>
<td>0.14</td>
<td>0.13</td>
</tr>
</tbody>
</table>

Table 7 presents the cross-sectional distribution of investment-to-capital ratio. 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.
Table A.1 presents the estimates of the exponent and the right-tail mass of power-law distribution for firm size and CEO compensation, and p-values of the Kolmogorov-Smirnov goodness-of-fit test. The right tail is quoted in terms of the number of firms in the right tail. Size in the data is measured by the number of firm employees.
Figure 1 plots the normalized value functions for the first best case (dash-dotted line), for the case of limited commitment on the manager side (dashed line), and for the two-sided limited commitment case (solid line). \( \bar{v} \) is the marginal/average Q in the Hayashi (1982) model. The domain of the normalized value function for the first best case is \([0, \infty)\) and \(u^\ast\) is the point where \(v(u)\) crosses zero. Under limited commitment on the manager side, the domain of the normalized value function is \([u_{MIN}, \infty)\) and the normalized continuation utility stays above \(u_{MIN}\). Under two-sided limited commitment, the normalized continuation utility stays between \(u_{MIN}\) and \(u_{MAX}\), where \(v(u_{MAX}) = 0\), and returns to the interior with probability one at both boundaries.

**Figure 1. Normalized Value Function**
Figure 2. Investment Policy

Figure 2 plots the optimal investment rate \( \left( \frac{I}{K} \right) \) as a function of the state variable \( u \) for the first best case (dash-dotted line), for the case with limited commitment on the manager side (dashed line), and for the two-sided limited commitment case (solid line).
Figure 3. Dynamics of Normalized Continuation Utility: Two-Sided Limited Commitment

Figure 3 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 4. Sample Path of CEO Compensation: The Bankruptcy Constraint

Figure 4 plots 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, $u_{MAX}$. 
Figure 5. Sample Path of CEO Compensation: The Participation Constraint

Figure 5 plots 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, $u_{MIN}$. 
Figure 6. The Right Tail of Size Distribution

Figure 6 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 7. The Right Tail of the Distribution of CEO Compensation

Figure 7 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.
Figure 8. Elasticity of CEO Compensation w.r.t. Size conditional on Size

Figure 8 plots the estimates of the 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”).
Figure 9. Investment-to-Capital Ratio across Size Percentiles

Figure 9 plots the average ratio of investment to capital across 20 size-sorted portfolios. In the data, size is measured by the number of firm employees.
Figure 10. Fraction of Dividend(Interest)-Paying Firms across Size Percentiles

Figure 10 plots the fraction of firms that make dividend and/or interest payments across 20 size-sorted portfolios. In the data, size is measured by the number of firm employees.