Firm Dynamics under Limited Commitment*

Hengjie Ai, Dana Kiku, and Rui Li †

August 1, 2016

We present a general equilibrium 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. Our model endogenously generates power laws in firm size and CEO compensation and explains the differences in their empirical distributions. We also show that the model is able to quantitatively account for the salient features of firms' growth dynamics, the observed negative relationship between firms' investment rate and size, and the positive relationship between firms' size and their dividend and CEO payout.

JEL Code: E13, G30

---

*This paper is previously circulated under the title, "A Mechanism Design Model of Firm Dynamics: The Case of Limited Commitment."

†Hengjie Ai (hengjie@umn.edu) is affiliated with the Carlson School of Management, University of Minnesota, Dana Kiku (dka@illinois.edu) is at the University of Illinois at Urbana-Champaign, and Rui Li (abclirui@gmail.com) is at the University of Massachusetts Boston. We would like to thank Daniel Andrei, Xavier Gabaix, Zhiguo He, Nengjiu Ju, Dmitry Livdan, Erzo Luttmer, Vincenzo Quadrini, Adriano Rampini, Vish Vishwanathan, and Noah Williams for their helpful comments on the paper. We also thank seminar participants at the American Finance Association 2013 Meeting, the China International Conference in Finance, the Kellogg Finance Conference, the Macro Finance Workshop, the Minnesota Macro-Asset Pricing Conference, the Cheung Kong Graduate School of Business, the University of Southern California, and the University of Wisconsin-Madison.

Introduction

The joint cross-sectional distribution of firm size, managerial compensation, investment and dividend policies is characterized by several stylized features. On average, small firms invest and grow at a higher rate and pay out less to their shareholders and managers compared with large firms. The right tail of the distribution of firm size follows closely a power law with the exponent of about one.[1] Dividend payout to shareholders and compensation to managers likewise exhibit power-law behavior: while dividends follow a power-law distribution that is similar to that of firm size, the power law in the compensation for chief executive officers (CEO) has a significantly thinner tail, with a slope of about two.

These features are hard to simultaneously explain in a frictionless neoclassical framework. As we show, under constant return-to-scale technologies, i.i.d. productivity growth and convex adjustment costs, all firms have identical investment and expected growth rates. As a result, Gibrat [28]'s law holds and the distribution of firm size follows a power law. However, contrary to the data, this framework rules out any dependence of investment and growth on firm size. Further, if shareholders are well-diversified and managers are risk averse, then the optimal compensation contract prescribes constant managerial pay, which is inconsistent with the large inequality in CEO compensation in the data. Although these data speak strongly against perfect risk sharing, the substantial difference in the power-law characteristics of CEO compensation and firm size suggests a yet significant degree of risk sharing between shareholders and managers. Motivated by this evidence, we present a model where the lack of commitment by shareholders and managers hinders perfect risk sharing and results in a sizable cross-sectional variation in investment and payout policies that is consistent with the data.

Our model is set in a general equilibrium framework where firms' investment, growth, dividend payout and managerial compensation decisions are endogenously determined as the result of the optimal contract subject to 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 show that limited commitment on both the shareholder and the manager sides can account for the key characteristics of the joint distribution of firm size and CEO compensation and the observed heterogeneity in investment and payout policies.

To understand the contracting implications of limited commitment, suppose that a firm's capital (both physical and intangible such as ideas, research, employee talent etc.) can be deployed efficiently only by its manager, who may also use it for his private benefit. The

[1] Specifically, the fraction of the U.S. firms with a size above K is proportional to $K^{-\xi}$, where ξ is approximately 1.1.

manger's outside option, therefore, increases with the size of the firm's capital stock. If the manager is offered a constant wage as a frictionless model would imply, it is ex post rational for him to walk away once his outside option exceeds the present value of what he expects to receive under the contract. Similarly, the shareholder may have an incentive to abandon the firm if the value of the firm becomes negative, i.e., if the present value of the firm's cash flow falls below the promised constant payment to the manager. Thus, if either the manager or the shareholder cannot fully commit, the ex-ante optimal constant compensation contract would not be ex-post incentive compatible. We show that the optimal contract under two-sided limited commitment prescribes a constant wage as long as none of the limited commitment constraints bind and a minimum modification in compensation necessary to satisfy these constraints whenever they bind.

We solve a general equilibrium model with two-sided limited commitment and establish several important results. First, we demonstrate that limited commitment on the manager side translates a power law in firm size into a power law in CEO compensation. In particular, we prove that managerial compensation follows a power law with an exponent that depends on the power law in firm size and the elasticity of managers' outside options to firm size.

Second, we show that limited commitment on the shareholder side results in an inverse relationship between firm investment and size and a positive relationship between dividend payout and size. When a firm's value declines, the shareholder commitment constraint is likely to bind. To avoid further losses and a potential default, small firms choose to defer dividends, cut down managerial pay and accelerate investment to grow out of the constraint. In contrast, large firms are likely to face a binding constraint on the manager side because managers' outside options become more attractive as firms grow. To reduce the likelihood of the managers' default, it is optimal for large firms to slow down their investment and growth and increase CEO compensation. Hence, consistent with the data, small firms in our model pay out less, invest more and grow faster compared with large firms.

We further show that the model calibrated to match a standard set of macroeconomic and aggregate moments can quantitatively account for the observed power-law behavior in firm size, dividends and executive compensation. In particular, it is able to replicate a wedge in the right-tail characteristics of the empirical distributions of firm size and CEO compensation. We also show that the two-sided limited commitment leads to a significant amount of heterogeneity in firms' investment and payout decisions that is quantitatively consistent with the sample variation in average investment and growth rates across size-sorted portfolios. In addition, we provide direct empirical evidence that corroborates the model-implied dynamics of CEO compensation and its response to fluctuations in firm size. Consistent with the model predictions, we show that in the data, small firms (especially those with weak performance) and large firms (especially those with superior performance) feature a significantly higher size elasticity of managerial compensation compared with the rest of the market.

The tradeoff between risk sharing and limited commitment in our model builds on the earlier work of Kehoe and Levine [36], Kocherlakota [39], and Alvarez and Jermann [4]. Our specification of managers' outside options is similar to the one in Kiyotaki and Moore [38], and Albuquerque and Hopenhayn [3]. Several more recent papers also study optimal contracting problems related to ours. Berk, Stanton, and Zechner [9] solve the optimal labor contract in a model with limited commitment and capital structure decisions. Eisfeldt and Papanikolaou [21, 22] emphasize that compensation of the key firm employees depends on their outside options. Rampini and Viswanathan [50, 51] study the implications of limited commitment for risk management and capital structure.[2]

Our paper is more closely related to two recent papers on limited commitment and firm dynamics. Cooley, Marimon, and Quadrini [15] consider a model with two-sided limited commitment where shareholders cannot commit to compensation plans that provide utility higher than managers' outside options. As a result, managers in their model always receive their outside options. In contrast, we assume that shareholders cannot commit to negative NPV projects. Therefore, the optimal contract in our model allows for risk sharing. Lustig, Syverson, and Van Nieuwerburgh [43] consider a model with limited commitment on the manager side and study the link between the inequality of CEO compensation and productivity growth. Their model also generates a power law in firm size, but unlike us, they do not characterize the power law in CEO compensation and its relationship with the power law in firm size. In addition, different from Lustig, Syverson, and Van Nieuwerburgh [43], in our model, investment decisions are endogenously determined by the optimal contract, which allows us to explore the implications of limited commitment for the cross-sectional distribution and life-cycle dynamics of firms' investment and growth.

This paper is related to the large literature on agency frictions and managerial compensation. Edmans and Gabaix [19] provide a comprehensive review of the earlier literature; more recent papers include Edmans, Gabaix, Sadzik, and Sannikov [20], Biais, Mariotti, and Villeneuve [11], and Bond and Axelson [13]. Our paper is also related to the literature on firm dynamics and power law in economics and finance (see Gabaix [26] and Luttmer [45] for a survey). The neoclassical model without frictions considered in our paper is essentially an interpretation of the model in Luttmer [44]. Terviö [57], and Gabaix

[2] A broader literature that focuses on the implications of dynamic agency problems for firms' investment and financing decisions includes Quadrini [49], Clementi and Hopenhayn [14], and DeMarzo and Fishman [16]. Limited commitment is also featured in Lorenzoni and Walentin [42], Schmid [56], Arellano, Bai, and Zhang [5], and Li [41].

and Landier [27] are assortative matching models that link CEO compensation to firm size taking size distribution as given. 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.

The continuous-time methodology of our paper builds on the fast growing literature on continuous-time dynamic contracting, for example, Sannikov [55], DeMarzo and Sannikov [18], DeMarzo, Fishman, He, and Wang [17], He [35], Biais, Mariotti, and Villeneuve [11].[3] The optimal contracting design in our paper is related to the one in Ai and Li [2], and Bolton, Wang, and Yang [12].

The rest of the paper is organized as follows. In Section 1, we summarize the key stylized features of the joint empirical distribution of firm size, investment, dividend payout and CEO compensation policies. In Section 2, we set up a general equilibrium model with limited commitment. In Section 3, we consider a frictionless Arrow-Debreu economy and discuss its limitations. In Section 4, we introduce and discuss the implications of the manager-side and the shareholder-side limited commitment separately. Our full model with two-sided limited commitment is presented and analyzed in Section 5. We calibrate the model and evaluate its quantitative implications in Section 6. Concluding remarks are provided in Section 7.

1 Stylized Facts

We begin with a brief summary of the empirical distribution of firms' size, investment, dividend 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 data. The in-depth analysis of the cross-sectional characteristics and time-series dynamics are provided in Section 6.

1. Firm sizes and firm dividends are both characterized by a power-law distribution with an exponent of about 1.1.[4]
2. The right tail of CEO compensation is also well approximated by a power law with a larger slope coefficient of around 2.
3. Managerial compensation is positively related to firm size. Empirically, the average elasticity of CEO compensation to size is about one-third.[5]

[3] For an excellent survey of this literature see Biais, Mariotti, Plantin, and Rochet [10].
[4] A power law in the firm size distribution is well documented in the literature, see for example Axtell [7].
[5] See Roberts [52], Baker, Jensen, and Murphy [8], and Frydman and Saks [25].

4. Small firms invest at a higher rate and grow faster compared with large firms. The average investment-to-capital ratio of firms in the bottom size quintile is about 17%. Large firms (those in the top quintile of the size distribution) have an average investment rate of around 9%.

5. Small firms are less likely to make dividend payments to their shareholders compared with large firms. In the bottom size quintile, on average, only one out of ten firms pays dividends. The fraction of dividend-paying firms increases to more than 70% in the right tail of the 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 6, we calibrate our benchmark model with two-sided limited commitment and evaluate its ability to quantitatively account for these and other characteristics of the data.

2 A General Equilibrium Model with Limited Commitment

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

2.1 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 (CRRA) utility:

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

where $\beta > 0$ is the time discount rate, and $\gamma > 0$ is the relative risk aversion coefficient. $\mathbf{C}_t$ denotes consumption flow rate of the shareholder at time t.[6]

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

2.2 Production Technology

Production of market consumption goods takes place at a continuum of firms indexed by $j \in \mathcal{J}$, where $\mathcal{J}$ is the set of all firms. General output of firm j at time t, denoted by $y_{j,t}$, is produced using capital and labor though a Cobb-Douglas technology:

$$y_{j,t} = \mathbf{z} K_{j,t}^{\alpha} N_{j,t}^{1-\alpha}, \tag{2}$$

where $K_{j,t}$ is the amount of capital and $N_{j,t}$ is the amount of labor employed by firm j at time t, and α is the capital share. $\mathbf{z}$ is the total factor productivity, which is assumed to be constant over time.

The representative shareholder owns all capital in the economy and supplies one unit of general purpose labor inelastically per unit of time but does not have accesses to the production technology. Managers are the only type of agents that have access to the Cobb-Douglas technology, which can be used to produce market consumption goods with general purpose labor hired in the market (labor hereinafter) or to produce home consumption goods with managers' own home labor.

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

$$\Pi(K) = \max_{N} \left\{ \mathbf{z} K^{\alpha} N^{1-\alpha} - \mathbf{W} N \right\} \tag{3}$$

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: $\Pi(K) = \mathbf{A}K$, where $\mathbf{A}$ is the economy-wide (equilibrium) marginal product of capital. Equation (3) implies that in equilibrium, firms' capital stock and the total number of employees are proportional to each other and, therefore, are equivalent measures of firm size.

The manager hired by firm j has access to a technology that accumulates capital according to the following law of motion:

$$dK_{j,t} = \left(I_{j,t} - \delta K_{j,t}\right) dt + K_{j,t} \sigma dB_{j,t}, \tag{4}$$

where $\delta > 0$ is the instantaneous depreciation rate of capital. The standard Brownian motion, $B_{j,t}$, is i.i.d. across firms and represents productivity shocks to the capital accumulation technology. The term $I_{j,t}$ is investment made in firm j at time t. Investing I in a firm with total capital stock K costs $h\left(\frac{I}{K}\right) K$ of general output, where $h\left(\frac{I}{K}\right) = \frac{I}{K} + \frac{1}{2} h_0 \cdot \left(\frac{I}{K}\right)^2$ with $h_0 > 0$ being a standard quadratic adjustment cost function.

2.3 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 $\bar{U}$ and $\bar{n}$ units of home labor that can be used for home production only.[7] In order to operate a firm 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. 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 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 firm profit. As we demonstrate in the paper, 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 $\bar{K}$, satisfies:

$$\bar{K} \in \arg\max_{K} \left\{ V\left(K, \bar{U}\right) - H(K) \right\}. \tag{5}$$

Managers value consumption using the same CRRA preference and the same discount rate β as shareholders. Unlike shareholders who live forever, managers are subject to random health shocks that follow a Poisson process with intensity $\kappa > 0$. 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. The continuation utility of a manager who operates firm j at time t is given by:

$$U_{j,t} = \left\{ E_t \left[\int_0^{\tau_j} (\beta + \kappa) e^{-\beta s} C_{j,t+s}^{1-\gamma} ds \right] \right\}^{\frac{1}{1-\gamma}}, \tag{6}$$

where τ_j is the stopping time at which the manager is hit by a Poisson health shock. For notational convenience, we adopt a normalization of the CRRA preference so that utility is measured in the same units as consumption.

[7] For simplicity, we do not explicitly specify the technology that delivers the reservation utility and treat $\bar{U}$ as a parameter of the model. The outside option is never taken in the equilibrium.

2.4 Profit Maximization under Limited Commitment

The economy we consider is one with overlapping generations 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 a generation-t firms with age s indexed by $j \in \mathcal{J}$. To save notations, we suppress the time index t, and consider the decision problem of a typical firm j of age s.

We focus on stationary equilibria where the discount rate $\mathbf{r}$ is constant over time. Taking $\mathbf{r}$ 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, $E_0\left[\int_0^{\tau_j} e^{-\mathbf{r}s} D_{j,s} ds\right]$, where E_0 stands for expectation taken with respect to information available when the firm is at age 0.

The shareholders' budget constraint requires that managerial compensation, cost of investment (including the adjustment cost), and dividend payment do not exceed the total operating profit:

$$C_{j,s} + h\left(\frac{I_{j,s}}{K_{j,s}}\right) K_{j,s} + D_{j,s} \leq \Pi\left(K_{j,s}\right), \text{ for all } s \geq 0. \tag{7}$$

Note that we do not impose an exogenous nonnegativity constraint, $D_{j,s} \geq 0$. Rather, we let the optimal contract determine the amount of dividend payment. A negative dividend payment can be interpreted as equity issuance. We focus on contracts that satisfy shareholders' budget constraint and managers' participation constraint that requires that the manager's life-time utility is at least as high as his outside option, that is:

$$U_{j,0} \geq \bar{U}. \tag{8}$$

So far we have described a frictionless economy with Arrow-Debreu markets. Implicit in the assumption of the Arrow-Debreu markets, is an infinitely severe punishment for any violation of pre-specified contracts. In other words, both shareholders' and managers' outside options are $-\infty$ at all t for any $t > 0$. As a result, all agents can fully commit once a contract is signed at date zero. The participation constraint in equation (8) pins down a unique allocation on the Pareto frontier.

As we show in Section 3 below, the ex ante efficient allocation in the frictionless economy may not be ex post incentive compatible if shareholder and managers have non-trivial outside options and cannot fully commit. Formally, we say that a contract $\{C_{j,s}, I_{j,s}, D_{j,s}\}_{s=0}^{\infty}$ is feasible if it satisfies shareholders' budget constraint in equation (7), managers' participation constraint in equation (8) and the following two limited commitment constraints.

First, shareholders cannot commit to negative net present value (NPV) projects. In this case, dividend policies that involve negative NPV of the firm at any point in time will result in shareholder default. Limited commitment on the shareholder side therefore restricts the set of feasible plans and requires that

$$V(K_{j,s}, U_{j,s}) \equiv E_s \left[\int_s^{\tau_j} e^{-\mathbf{r}(t-s)} D_{j,t} dt \right] \geq 0, \text{ for all } s \geq 0. \tag{9}$$

That is, the value of the firm must be non-negative at all times in all states of the world.

Second, following Kehoe and Levine [36], Kiyotaki and Moore [38], and Albuquerque and Hopenhayn [3], we assume that managers have an option to default and cannot commit to compensation contracts that yield life-time utility lower than that provided by their outside option. We further assume that the outside option is a function of capital under management and denote it by $U_{OUT}(K_{j,s})$. Limited commitment on the manager side requires the compensation plan to satisfy:

$$U_{j,s} \geq U_{OUT}(K_{j,s}), \text{ for all } s \geq 0. \tag{10}$$

Upon default, the manager can retain all of the firm's capital stock.[8] However, he is forever excluded from the credit market. That is, after default, he can only consume the output produced from the capital stock he absconds with but cannot enter into any intertemporal risk sharing contract.[9] We consider two specifications that differ in terms of managers' labor market participation upon default.

In the first specification that we refer to as Specification 1, after default, managers are allowed to hire general labor in the market to use in the production process. Here, for simplicity, we also assume that $\bar{n} = 0$ so that home production technology is never used. As we show in the appendix, the utility of the manager upon default under Specification 1 is a linear function of the firm's capital: $U_{OUT}(K) = \varpi_1 K$, where ϖ_1 is a constant given in equation (B.2).

In the second specification (Specification 2), upon default, managers are not only excluded from the credit market but also from the labor market. In this case, after default, a manager can only use capital and home labor to produce home consumption goods. As we show in the appendix, the utility of the manager upon default is $U_{OUT}(K) = \varpi_2 K^{\alpha}$, where ϖ_2 is a

[8] In general, we can allow managers to retain only a fraction of the firm's capital as in Kiyotaki and Moore [38], and Rampini and Viswanathan [50]. However, because managers are risk averse in our setup, exclusion from intertemporal risk sharing arrangement represents a non-trivial punishment to managers. For parsimony, following Kehoe and Levine [36], we assume that managers can abscond with all of their capital under management.

[9] Note that default never occurs in equilibrium because it destroys risk sharing and is inefficient.

constant given in equation (B.3).

To save notation, we will write

$$U_{OUT}(K) = \varpi K^{\nu}, \tag{11}$$

with the understanding that $\varpi = \varpi_1$ and $\nu = 1$ in Specification 1, and $\varpi = \varpi_2$, $\nu = \alpha$ in Specification 2. Also, in what follows we refer to constraint in equation (9) as the shareholder-side limited-commitment constraint and condition in equation (10) as the manager-side limited-commitment constraint.

Note that K in our model should be interpreted as a broad notion of capital, which include not only physical capital, but more importantly, human capital and other forms of intangible capital, which are likely to be more important determinants of managers' outside options. Although we do not distinguish different types of capital here, we present an extension of our model in the technical appendix, where production requires both physical and human capital, but managers' outside options depend only on human capital. We show that this extension is observational equivalent to our main model.

2.5 General Equilibrium

A competitive equilibrium must specify the path of interest rates, $\{\mathbf{r}_t\}_{t\geq 0}$, wages, $\{\mathbf{W}_t\}_{t\geq 0}$, consumption of the representative shareholder, $\{\mathbf{C}_t\}_{t\geq 0}$, managerial compensation, investment, and dividend payout policies for all firms at all times,

$$\left\{\left[\left(C^t_{j,t+s}, I^t_{j,t+s}, D^t_{j,t+s}\right)_{s=0}^{\infty}\right]_{t=0}^{\infty}\right\}_{j\in\mathcal{J}}.$$

We focus our attention on stationary equilibria where the exit rate of firms equals the entry rate, and the cross-sectional distribution of firm characteristic is time-invariant. In such equilibria, the interest rate and the wage rate are also time-invariant. The optimal contracting problem can be solved recursively by using (K, U) as state variables, where K is the total capital stock of the firm and U is the continuation utility promised to the manager defined in equation (6). The advantage of the continuous-time and Brownian motion setup is that the dynamics of continuation utility are completely determined by specifying a consumption flow, $C(K, U)$ and a local sensitivity of the continuation utility with respect to Brownian shocks, $G(K,U)$. In particular,

$$dU = \left[-\frac{\beta+\kappa}{1-\gamma} \left(C^{1-\gamma} U^{\gamma} - U \right) + \frac{1}{2} \frac{G(K,U)^2}{U} \right] dt + G(K,U)\, dB.^{10} \tag{12}$$

Because policies depend only on the state variables (K,U), we can think of them as a summary of the firm's type. Following the convention of Atkeson and Lucas [6], hereinafter, we suppress j and index firms by their type (K,U) and construct the equilibrium allocation by using the 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 the last section. Specifically, for firms of each type, we first specify the flow rate of managerial compensation, investment, dividend payout and the amount of labor hired (at the current instant) using the allocation rules, $\{C(K,U),\ I(K,U),\ D(K,U),\ N(K,U)\}$. Next, we construct the law of motion of the state variables from the allocation rules using the law of motion of capital,

$$dK = K \left[\left(\frac{I(K,U)}{K} - \delta \right) dt + \sigma dB \right], \tag{13}$$

and the law of motion of continuation utility in equation (12).

Formally, an equilibrium consists of an interest rate, $\mathbf{r}$, a real wage, $\mathbf{W}$, allocation rules, $\{C(K,U),\ I(K,U),\ D(K,U),\ N(K,U),\ G(K,U)\}$, consumption of the representative shareholder, $\mathbf{C}$, and a cross-sectional 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 firm value subject to the limited-commitment constraints on the shareholder and the manager sides.

2. The initial condition for new firms is $\left(\bar{K}, \bar{U}\right)$, where $\bar{K}$ maximizes the profit of the creation of new firms, as in equation (5).

3. Taking real wages as given, the policy function $N(K,U)$ solves the intra-temporal profit maximization problem in equation (3) for all firms at all times.

[10] This formulation is similiar to the representation in Sannikov [55], except that we use a monotonic transformation so that utility is measured in consumption units. We provide the details of the derivation in Appendix C.

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

5. Goods market clears:

$$\mathbf{C}+\int\left[C\left(K,U\right)+h\left(\frac{I\left(K,U\right)}{K}\right)K\right]d\Phi\left(K,U\right)+H\left(\bar{K}\right)=\mathbf{z}\int K^{\alpha}N\left(K,U\right)^{1-\alpha}d\Phi\left(K,U\right). \tag{14}$$

6. Labor market clears:

$$\int N\left(K,U\right)d\Phi\left(K,U\right)=1. \tag{15}$$

7. The cross-sectional distribution of types, $\Phi\left(K,U\right)$, is consistent with the law of motion of $\left(K,U\right)$ implied by the allocation rules, as in equations (12) and (13).[11]

Note that under Specification 1, due to the linearity of managers' outside options, firms' value and policy functions are homogenous of degree one in K. It is, therefore, convenient to work with normalized functions. Let $u=\frac{U}{K}$ and $v\left(u,K\right)=\frac{V(K,uK)}{K}$ denote the normalized utility and value functions, respectively. Because in Specification 1, $v\left(u,K\right)$ does not depend on K, with a slight abuse of notations, we will denote it as $v\left(u\right)$. Similarly, let $c\left(\cdot\right)$ and $i\left(\cdot\right)$ represent the normalized compensation and investment policies; that is,

$$C\left(K,U\right)=c\left(u\right)K;\qquad I\left(K,U\right)=i\left(u\right)K. \tag{16}$$

The normalized value function can be characterized by the solution to an ordinary differential equation (ODE) with appropriate boundary conditions. In addition, due to the homogeneity of decision rules, the two dimensional measure $\Phi\left(K,U\right)$ in the market clearing conditions (14) and (15) can be replaced by a one-dimensional "summary measure" as described in Appendix D.

Under Specification 2, the homogeneity property no longer holds because managers' outside options are no longer linear in firm size. However, we will continue using the state pair $\left(u,K\right)$ to facilitate comparison across the two specifications.

We conclude this section by making two observations that are useful for understanding the general equilibrium. First, in a stationary equilibrium, consumption of the representative shareholder is constant over time because she is well diversified and is not exposed to idiosyncratic shocks. As a result, the shareholder's intertemporal maximization problem implies that the risk-free interest rate must equal the shareholder's time discount rate: $\mathbf{r}=\beta$.

[11] Technically, $\Phi\left(K,U\right)$ must satisfy a version of the Kolmogorov forward equation as we show in the appendix.

Second, health shocks make managers effectively less patient than shareholders because they evaluate utility using an effective discount rate of $\beta + \kappa$ (see equation (6)). However, because the cash flows completely evaporate after managers are hit by health shocks, shareholders value firm's cash flow with the same discount rate of $\mathbf{r} + \kappa$, i.e., $V(K_t, U_t) = E_t\left[\int_0^\infty e^{-(\mathbf{r}+\kappa)s} D_{t+s} ds\right]$. From the optimal contract design perspective, our model should therefore be interpreted as one where the principal and the agent have the same discount rate because $\beta = \mathbf{r}$.[12]

3 The First-Best Case

In the first-best case, shareholders maximize the present value of firm's cash flow subject to the manager's participation constraint in equation (8). Using the budget constraint in equation (7), the present value of cash flow can be written as:

$$E_0\left[\int_0^\infty e^{-(\mathbf{r}+\kappa)t}\left(\mathbf{A}K_t - h\left(\frac{I_t}{K_t}\right)K_t\right)dt\right] - E_0\left[\int_0^\infty e^{-(\mathbf{r}+\kappa)t} C_t dt\right]. \tag{17}$$

Note that the participation constraint affects only the choice of managerial compensation in the second term. As a result, the profit maximization problem is separable and can be solved in two steps. The first step is to maximize the total value of the firm in the first term of equation (17) by choosing the optimal investment policy. The second step is to select the optimal managerial compensation to minimize the cost subject to the manager's participation constraint.[13]

The firm value maximization problem in the first step is standard as in Hayashi [34]. The solution to the cost minimization problem is also straightforward: risk aversion of the manager and the fact that the principal and the agent have identical discount rates imply a constant consumption of the manager: $C_t = \bar{U}$ for all t.[14] We make the following assumptions to guarantee that firm value is finite and the maximization problem is well defined.

Assumption 1. *The parameter values of the model satisfy:*

$$\mathbf{A} > \mathbf{r} + \delta + \kappa > \frac{-1 + \sqrt{1 + 2h_0\mathbf{A}}}{h_0} \tag{18}$$

Then we summarize the solution to the firm's problem in the following proposition.

[12]Our model can be easily extended to allow for different discount rates.

[13]This separation is no longer possible in the case with agency frictions because the limited-commitment constraints impose joint restrictions on the sequence of C_t and K_t.

[14]Because the risk-averse shareholders hold the claims of all firms in the economy, they are effectively risk-neutral with respect to idiosyncratic shocks, which can be completely diversified.

Proposition 1. *The first-best Case*

Under the Assumption 1, firm value is finite and is given by:

$$V(K, U) = \bar{v}K - \frac{1}{\mathbf{r} + \kappa}U, \tag{19}$$

where $\bar{v} = h'(\hat{\imath})$ *and* $\hat{\imath} \in (0, \hat{r})$ *is the optimal investment-to-capital ratio given by:*

$$\hat{\imath} = \arg\max_{i<\hat{r}} \frac{\mathbf{A} - h(i)}{\hat{r} - i} = \hat{r} - \sqrt{\hat{r}^2 - \frac{2}{h_0}(\mathbf{A} - \hat{r})}, \tag{20}$$

where $\hat{r} \equiv \mathbf{r} + \kappa + \delta$.

Proof. See Appendix A. □

The first term $\bar{v}K = h'(\hat{\imath})K$ in equation (19) is the firm value in the neoclassical model with capital adjustment costs. The second term is the present value of the cost of managerial compensation. In the absence of aggregate uncertainty, perfect risk sharing implies a constant managerial compensation, the resent value of which is simply given by the Gordon [29]'s formula: $\frac{1}{\mathbf{r}+\kappa}U$.[15]

To close the model, we note that the rate at which new capital is created in this economy equals the initial size of firms, $\bar{K}$, because the total amount of entrant firms is normalized to measure one per unit of time. The fact that existing firms grow at rate $\hat{\imath}$ and that the aggregate depreciation rate (physical depreciation plus involuntary exit) is $\kappa + \delta$ imply a steady-state capital stock of $\mathbf{K} = \frac{\bar{K}}{\kappa+\delta-\hat{\imath}}$. In addition, because total labor endowment is one, the equilibrium marginal product of capital is

$$\mathbf{A} = \alpha\mathbf{z}\left(\frac{1}{\mathbf{K}}\right)^{1-\alpha} = \alpha\mathbf{z}\left(\frac{\kappa + \delta - \hat{\imath}}{\bar{K}}\right)^{1-\alpha}. \tag{21}$$

Also, the first-order condition for the creation of new firms implies that

$$H'\left(\bar{K}\right) = \bar{v} = 1 + h_0\hat{\imath}. \tag{22}$$

The three equations (20), (21), and (22) jointly determine three equilibrium quantities: $\mathbf{A}$, $\hat{\imath}$, and $\bar{K}$. Given them, we can construct the equilibrium wage as the marginal product of labor, $\mathbf{W} = (1 - \alpha)\mathbf{z}\mathbf{K}^{\alpha}$, and the aggregate consumption of the representative shareholder is the total output less the cost of investment.[16]

[15] Recall that we normalize the utility function of the manager so that life-time utility is measured in consumption units.

[16] We need to make assumptions on the primitive parameters of the model so that the assumption in

Equation (20) in Proposition 1 shows that the investment-to-capital ratio in this economy is constant across firms. As a result, Gibrat's law holds, growth rates are i.i.d. across firms and the distribution of firm size follows a power law as in Luttmer [44], which is summarized in the following proposition.

Proposition 2. *Power Law of Firm Size*

Given firms' initial size, $\bar{K}$, and their optimal investment policy, $\hat{\imath}$, the density of the firm size distribution is given by:

$$\phi(K) = \begin{cases} \frac{1}{\sqrt{\left(\hat{\imath}-\delta-\frac{1}{2}\sigma^2\right)^2+2\kappa\sigma^2}} \bar{K}^{-\xi} K^{\xi-1} & K \geq \bar{K} \\ \frac{1}{\sqrt{\left(\hat{\imath}-\delta-\frac{1}{2}\sigma^2\right)^2+2\kappa\sigma^2}} \bar{K}^{-\eta} K^{\eta-1} & K < \bar{K}, \end{cases} \tag{23}$$

where $\eta > \xi$ are the two roots of the quadratic equation $\kappa + \left(\hat{\imath} - \delta - \frac{1}{2}\sigma^2\right) x - \frac{1}{2}\sigma^2 x^2 = 0$. In particular, the right tail of firm size obeys a power law with exponent ξ.

Proof. See Appendix A.2. □

To summarize, the first-best model generates a power law in firm size, which is consistent with the right-tail behavior of the empirical distribution. However, it fails to account for other important features on the data. First, it rules out any cross-sectional variation in investment rates and, hence, fails to explain a robustly negative relationship between firm size and investment. Similarly, it cannot account for the observed cross-sectional differences in growth rates. Second, the distribution of CEO compensation in the first-best case is degenerate. Hence, in contrast to the data, it implies a zero elasticity of managerial compensation with respect to firm size and obviously cannot account for the observed fat tail in CEO pay.

4 One-Sided Limited Commitment

In this section, we discuss the implications of manager-side and shareholder-side limited commitment separately. We first consider the manager-side limited commitment and provide an analytical proof of the power law in managerial compensation. Then, we briefly summarize the implications of the shareholder-side limited commitment to prepare for the discussion of the full model with two-sided limited commitment.

Proposition 1 is satisfied and $V\left(\bar{K}, \bar{U}\right) \geq H\left(\bar{K}\right)$ to ensure a positive entry in equilibrium. Although these assumptions are not here listed to conserve space, we make sure that they are satisfied in our calibration.

4.1 Limited Commitment on the Manager Side

In the case of the lack of commitment on the manager side, shareholders maximize firm value subject to managers' time-zero participation constraint (equation (8)), and the requirement that managers' continuation utility must be higher than their outside options at all times (equation (10)). Our main objective in this section is to show that limited commitment on the manager side generates a power law in managerial compensation and to relate the power law in CEO pay to the power law in firm size.

To facilitate closed-form solutions, we assume a special form of the adjustment cost function:

$$h(i) = \begin{cases} i & \text{if } 0 \le i \le \hat{\imath} \\ \infty & \text{if } i > \hat{\imath} \end{cases},$$

where $\hat{\imath} > \delta$. That is, we assume that the marginal cost of investment is one if $\frac{I}{K} \le \hat{\imath}$, and is infinite if $\frac{I}{K} > \hat{\imath}$. Under the above assumption, firms' optimal investment policy takes a very simple form: either invest at the maximum rate $\hat{\imath}$ or do not invest at all. We also make the following assumption.

Assumption 2.

$$\mathbf{A} > \mathbf{r} + \kappa + \delta > \hat{\imath}, \tag{24}$$

and

$$\frac{\mathbf{A} - \hat{\imath}}{\mathbf{r} + \kappa + \delta - \hat{\imath}} - \frac{\nu\gamma}{(\mathbf{r} + \kappa)(\zeta_1 - 1)} \zeta_1^{\frac{\nu\gamma - 1}{\nu(\gamma - 1)}} (\zeta_1 - (1 - \gamma))^{\frac{-\nu\gamma}{\nu(\gamma - 1)}} \varpi^{\frac{1}{\nu}} \ge 1, \tag{25}$$

where

$$\zeta_1 = \sqrt{\left(\frac{\hat{\imath} - \delta}{\sigma^2} - \frac{1}{2}\right)^2 + \frac{2(\kappa + \mathbf{r})}{\sigma^2}} - \left(\frac{\hat{\imath} - \delta}{\sigma^2} - \frac{1}{2}\right).$$

Inequality (24) imposes a lower bound on the marginal product of capital, $\mathbf{A}$. Inequality (25) is a restriction on the magnitude of managers' outside options, ϖ, which is a measure of the severity of agency frictions. Together, they guarantee that the technology is productive enough and that the agency frictions are not too large so that it is always optimal for firms to invest at the maximum rate of $\hat{\imath}$. The above assumptions allow us to simplify firms' investment decisions and focus on the implications of the model for power laws in firm size and CEO compensation. We summarize our main results in the following proposition.

Proposition 3. *Power Law in CEO Compensation*

1. *Under Assumptions 2, CEO compensation under the optimal contract is given by:*

$$C_t = \max\left\{\hat{c} \max_{0 \le s \le t} K_s^{\nu},\ C_0\right\}, \tag{26}$$

where the constant $\hat{c}$ is defined in equation (B.4) in the appendix. The optimal investment-to-capital is constant: $I_t = \hat{\imath} K_t$ for all t.

2. *The right tail of CEO compensation obeys a power law with a slope coefficient of $\frac{\xi}{\nu}$ with ξ being defined in Proposition 2.*

Proof. See Appendix B. □

In the model with limited commitment on the manager side, the compensation contract is downward rigid, as in Harris and Holmstrom [32]. Compensation has to increase to match the manager's outside option whenever the limited commitment constraint binds. Otherwise, due to risk sharing, it must remain constant. Because the manager's outside option is an increasing function of firm size, the above dynamics imply that managerial compensation must be an increasing function of the running maximum of firm size.[17]

Due to the special form of the adjustment cost function, firm investment rate is constant and Gibrat's law holds. As a result, Proposition 2 applies and firm size follows a power law with slope ξ. It is straightforward to show that if the distribution of K follows a power law with slope coefficient ξ, the distribution of K^ν obeys a power law with slope coefficient $\frac{\xi}{\nu}$. By part 1 of Proposition 3, managerial compensation is a linear function of the running maximum of K_t^ν. Intuitively, the running maximum of a power law process obeys a power law with the same slope coefficient. Therefore, managerial compensation in our model follows a power law with slope $\frac{\xi}{\nu}$. Proposition 3 thus links the power law in CEO pay to the power law in firm size and the elasticity of CEOs' outside options with respect to firm size. In our calibration exercise, we show that this relationship generalizes to the case with smooth adjustment costs, where Gibrat's law does not hold.

Under Specification 1, because managers are allowed to participate in the labor market after default, the constant return to scale of the production function implies that the profit function is linear in size. Therefore, the manager's outside option is also linear in K. In this case, Proposition 3 implies firm size and CEO compensation have the same power law, ξ. Under Specification 2, after default, the manager can only use his home labor to produce home consumption goods but is not allowed to participate in the labor market. In this case, total output after default is $\mathbf{z}K^\alpha \bar{n}^{1-\alpha}$ and the manager's outside option is proportional to K^α. Because managers are excluded from the labor market after default, capital is less scalable and the distribution of CEO pay has a thinner tail, with a slope of $\frac{\xi}{\alpha}$.[18] Hence,

[17] See also Lustig, Syverson, and Van Nieuwerburgh [43], Grochulski and Zhang [30], and Miao and Zhang [47].

[18] Recall that under Specification 2, $\nu = \alpha$ (see equation (11)).

qualitatively, Specification 2 is able to account for the observed differences in the right tails of the distributions of managerial compensation and firm size.

It is straightforward to show that dividend payout must follow a power law with the same slope as firm size, ξ. Assuming firm size is large enough so that the limited commitment constraint for managers bound at least once in the past, then $C_t = \hat{c} \max_{0 \le s \le t} K_s^{\nu}$ and $D_t = \mathbf{A}K_t - I_t - C_t = \mathbf{A}K_t - \hat{\imath}K_t - \hat{c} \max_{0 \le s \le t} K_s^{\nu}$. Because $\nu \le 1$, it follows that

$$(\mathbf{A} - \hat{\imath})\, K_t - \hat{c} \max_{0 \le s \le t} K_s \le D_t \le (\mathbf{A} - \hat{\imath})\, K_t.$$

Since both sides of this inequality follow a power law with slope ξ, dividends must obey the same power law.

Finally, we note two main differences between our framework with one-sided commitment and Albuquerque and Hopenhayn [3]. First, the Albuquerque and Hopenhayn [3] model features decreasing return to scale technologies and stationary productivity. As a result, firms eventually reach their optimal size, where the limited-commitment constraint does not bind. Therefore, neither the distribution of firm size nor that of the managerial compensation in their model have fat right tails. Second, managers are risk-neutral in Albuquerque and Hopenhayn [3]. Therefore, they receive no payment as long as the limited commitment constraint binds, and payment policy is undetermined once the firm grows out of the constraint.

4.2 Limited Commitment on the Shareholder Side

In the model with limited commitment on the shareholder side, shareholders maximize firm value subject to managers' time-zero participation constraint (equation (8)) and the requirement that firm value must stay non-negative at all times (equation 9). Because the objective function and the constraints are linear in size, the value and policy functions are homogenous in K and it is convenient to work with the normalized value and policy functions defined in equation (16). As we show in Appendix E, the normalized value function $v(u)$ is strictly decreasing in the normalized continuation utility of the manger. Intuitively, as we use more of the firm's cash flows to support a higher continuation utility of the manager, shareholder value declines. If we define u_{MAX} to be the highest normalized utility that can be supported by the optimal contract without violating shareholder's limited-commitment constraint, that is,

$$u_{MAX} = \sup \{u : v(u) \ge 0\}, \tag{27}$$

then under the optimal contract the shareholder constraint, $E_t\left[\int_0^\infty e^{-(\mathbf{r}+\kappa)t+s} D_{t+s} ds\right] \geq 0$, is equivalent to $u_t \leq u_{MAX}$. We describe two properties of the normalized value function and the normalized investment policy but relegate the details of the optimal contract to Appendix E.

1. The normalized value function, $v(u)$, is strictly decreasing and concave on $(0, u_{MAX}]$ with $v(u_{MAX}) = 0$ and $\lim_{u \to 0} v(u) = \bar{v}$, where $\bar{v}$ is the firm value in the first-best case defined in Proposition 1.

2. The investment-to-capital ratio, $i(u)$, is strictly increasing in u and $\lim_{u \to 0} i(u) = \hat{\imath}$, where $\hat{\imath}$ is the first-best investment level defined in equation (20).

To understand the intuition of the above observations, suppose that a firm starts at $\bar{U}$ and experiences a sequence of positive productivity shocks. As the firm size increases, the normalized utility declines. In the limit, as $u_t = \frac{U_t}{K_t} \to 0$, the probability of hitting a binding constraint vanishes; therefore, both the investment-to-output ratio and firm value converge to their first-best levels. In contrast, a sequence of negative productivity shocks reduces K_t pushing the normalized utility towards u_{MAX}, where the limited-commitment constraint on the shareholder side binds. To avoid hitting the constraint and improve risk-sharing, firms optimally increase their investment as u approaches u_{MAX}. Hence, under limited commitment on the shareholder side, small firms invest more and growth faster compared with large firms.

To summarize, the limited commitment on the manager side generates a power law in managerial compensation and the limited commitment on the shareholder side generates an inverse relationship between investment rate and firm size. In the next section, we incorporate both frictions into our model and analyze the optimal contract and its implications in details.

5 Two-Sided Limited Commitment

In this section we present our full model, where managers cannot commit to compensation plans that deliver lower continuation utility than their outside options and shareholders cannot commit to negative NPV projects.

5.1 Optimal Contract with Two-sided Limited Commitment

The optimal contract maximizes firm value subject to managers' time-0 participation constraint (equation (8)), the shareholder-side limited-commitment constraint (equation (9)), and the manager-side limited-commitment constraint (equation (10)). Note that the limited-commitment constraint on the manager side requires $U_t \geq \varpi K^{\nu}$, or $\frac{U_t}{K_t} \geq \varpi K_t^{\nu-1}$ for all t. This motivates the following definition of the lower boundary of the normalized continuation utility: $u_{MIN}(K) = \varpi K^{\nu-1}$. The manager-side limited commitment constraint can therefore be written as $u_t \geq u_{MIN}(K_t)$. Note that in Specification 1, $u_{MIN}(K) = \varpi_1$ and does not depend on K. Similarly, generalizing equation (27), we can also define an upper bound for the normalized continuation utility (that is, the highest normalized utility of the manager that can be supported by the optimal contract):

$$u_{MAX}(K) = \sup\{u : V(uK, K) \geq 0\}.$$

The properties of the optimal contract are summarized in the following proposition.[19]

Proposition 4. *Two-Sided Limited Commitment*

1. *Under the optimal contract, $u_{MIN}(K_t) \leq u_t \leq u_{MAX}(K_t)$ for all t. In the domain $\left\{(K,U) : u_{MIN}(K) \leq \frac{U}{K} \leq u_{MAX}(K)\right\}$, the value function satisfies an HJB equation, which is listed in Appendix C together with the associated boundary conditions. Given the value function, the optimal CEO compensation and optimal investment can be constructed from the first-order conditions in equation (C.5).* [20]

2. *Suppose none of the limited-commitment constraint binds between time t_1 and t_2, then $C(K_t, U_t) = C(K_{t_1}, U_{t_1})$ for all $t \in [t_1, t_2]$.*

3. *In Specification 1, the value function and policy functions are homogenous of degree one in K. The boundaries $u_{MAX}(K) = \varpi_1$ and $u_{MIN}(K) = u_{MIN}$ do not depend on K. The normalized compensation policy satisfies*

$$\ln c(u_t) = \ln C_0 - \ln K_t + l_t^+ - l_t^-, \tag{28}$$

where $\left\{l_t^+, l_t^-\right\}_{t=0}^{\infty}$ are the minimum increasing processes such that $c(u_{MIN}) \leq c(u_t) \leq c(\varpi_1)$ for all t.[21] *The optimal investment, $i(u)$, is a strictly increasing function of u.*

[19] A recent paper by Bolton, Wang, and Yang [12] shows how similar contacts can be implemented by corporate liquidity and risk management policies.

[20] Using the recursive formulation for the optimal contracting problem, it is straightforward to show that the optimal contracts are renegotiation- proof in all versions of our model. For example, in the case with two-sided limited limited commitment, the optimal contract is renegotiation-proof because the value function $V(K_t, U_t)$ together with the associated continuation policy functions solves the optimal contracting problem subject to the limited-commitment constraints (8) and (10) with initial conditions (K_t, U_t).

[21] Formally, l^+ and l^- are the unique pair of continuous and nondecreasing processes that satisfy the following conditions: i) $l_0^+ = l_0^+ = 0$, ii) $\ln c(u_{MIN}) \leq \ln C_0 - \ln K_t + l_t^+ - l_t^- \leq \ln c(\varpi_1)$ for all t, iii) l^+ increases only when $\ln C_0 - \ln K_t + l_t^+ - l_t^- = \ln c(u_{MIN})$ and l^+ increases only when $\ln C_0 - \ln K_t + l_t^+ - l_t^- = \ln c(\varpi_1)$. For alternative equivalent constructions of two-sided regulators, see Harrison [33].

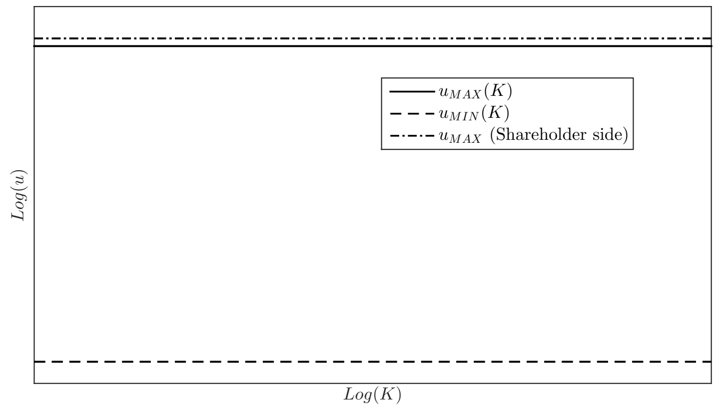
Proof. See Appendix C. □

5.2 Domain of the State Variables

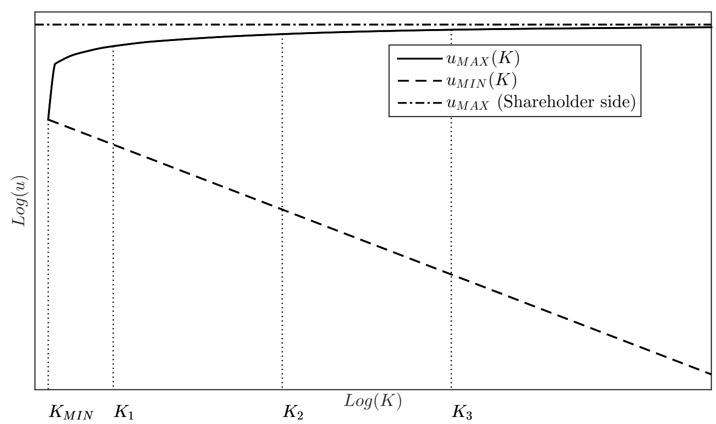
In Figure 1, we plot the boundaries of the normalized utility, $u_{MIN}(K)$ (dashed line) and $u_{MAX}(K)$ (solid line), for Specification 1 in the top panel and Specification 2 in the bottom panel. For convenience, both normalized utility and firm size are plotted in log units. Under Specification 1, because the value function and policy functions are homogenous of degree one in K, the boundary for the normalized utility do not depend on K, as shown in Panel (a) of Figure 1. Intuitively, $u_{MAX}(K)$ is the highest level of the normalized utility that can be supported by the optimal contract. Higher levels of u result in negative firm value. The lower boundary, $u_{MIN}(K)$, is the lowest level of the normalized utility that is required to retain the manager and to prevent him from taking the outside options. Note that under Specification 1, it is never optimal for the shareholder to default because in that case capital cannot be productively deployed. Also, in equilibrium, the manager never takes the outside option because it destroys risk-sharing. Therefore, under the optimal contract, the normalized continuation utility always stays within the two boundaries.

The boundaries $u_{MIN}(K)$ and $u_{MAX}(K)$ for Specification 2 are plotted in Panel (b) of Figure 1. Recall, that in this specification, the minimum promised utility required to match managers' outside option is $\varpi_2 K^\alpha$. Therefore, $\log u_{MAX}(K) = \log\left(\frac{\varpi_2 K^\alpha}{K}\right) = \log \varpi_2 - (1-\alpha)\log K$, represented by the dashed line, has a negative slope of $1-\alpha$. As firm size gets smaller, managers' outside option shrinks at a lower rate than firm's cash flow. Eventually, as $K \searrow K_{MIN}$, the cash flow is not enough to support managerial pay and it is optimal to shut down the firm and let the manager take his outside option. As a result, Specification 2 allows for endogenous firm exit. Intuitively, as K gets smaller, the home production technology dominates the market technology and it is optimal to dissolve the manager-firm match. Hence, K_{MIN} is the smallest firm size in this economy. At K_{MIN}, the net cash flow of the firm is actually negative and the firm value is zero (note that the firm value is still non-negative because is incorporates the option value of future growth). A further decline in capital lowers the firm value and results in the optimal shut-down of the firm. Therefore, Specification 2 implies that small firms exit markets more frequently relative to large firms.

As K increases, the market production technology becomes more efficient than the home technology and $u_{MAX}(K)$ increases. In the limit, as $K \to \infty$, $u_{MAX}(K)$ converges to the dash-dotted line, which is the upper bound on the normalized utility for the case with the shareholder-side limited commitment only. Intuitively, as firm size increases, the manager's outside option becomes negligible compared with the firm's cash flow, and limited



(a) Specification 1



(b) Specification 2

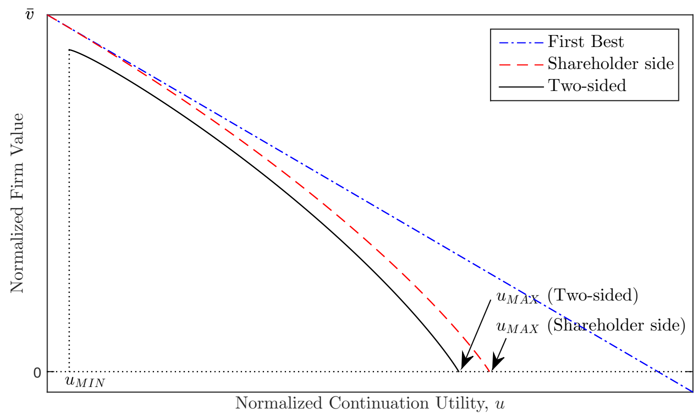
Figure 1: Domain of the State Variables in the $\ln K$–$\ln u$ Space

Figure 1 plots the domain of the state variables in the $\ln K$—$\ln u$ space for Specification 1 (top panel) and Specification 2 (bottom panel). The solid line represents the upper bound of normalized utility under the optimal contract, $u_{MAX}(K)$. It is the highest level of normalized utility of the manager that can be supported by the optimal contract. Higher levels of u results in negative firm value and shareholder default. The dashed line is the lower bound of normalized utility under the optimal contract, $u_{MIN}(K)$. It is the lowest level of normalized utility that keeps the managers from taking their outside options. The dash-dotted line is the upper bound on normalized utility in a model with one-sided limited commitment only. Under Specification 1, $u_{MAX}(K)$ and $u_{MIN}(K)$ do not depend on K due to the homogeneity of the problem. In Specification 2, the manager side limited commitment constraint is less likely to bind as firm size increases, and as a result, $u_{MAX}(K)$ asymptotes to that for the model with shareholder side limited commitment only as $K \to \infty$.

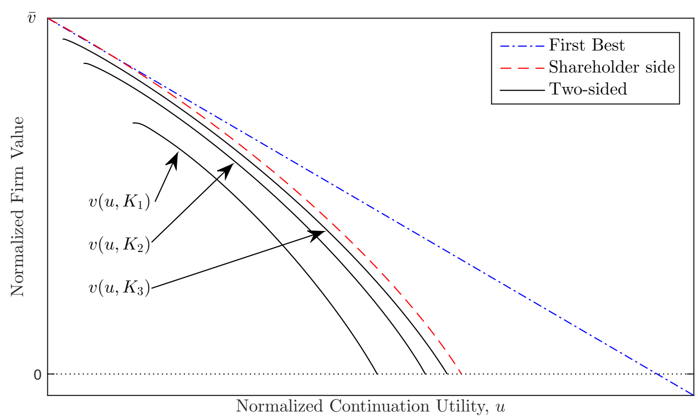
commitment on the manager side almost never binds.

5.3 Value Functions

The normalized value function, $v(u, K)$, for the two-sided limited commitment model is plotted in Figure 2 and is represented by a solid line. For the purpose of comparison, in the



(a) Specification 1



(b) Specification 2

Figure 2: Normalized Value Functions

Figure 2 plots the normalized value function for Specification 1 (top panel) and Specification 2 (bottom panel). The dash-dotted linear line is the value function for the first-best case and the dashed line represents the normalized value function for the case with one-sided limited commitment on shareholder side only. The solid lines are the normalized value function for our model with two-sided limited commitment. In the top panel, and normalized value function does not depend on capital stock due to homogeneity. In the bottom panel, we plot the normalized value function for three different levels of capital stock, $K_1 < K_2 < K_3$. As the level of K increases, the normalized value function approaches that of the model with shareholder side limited commitment constraint only.

same figure, we also plot the value function for the first-best case (dash-dotted line) and that for the case with the shareholder-side limited commitment only (dashed line).[22]

The top panel corresponds to Specification 1, under which $v(u)$ does not depend on K and is defined over $u \in [u_{MIN}, u_{MAX}]$. In the first-best case, the normalized value function is linear with slope $-\frac{1}{\mathbf{r}+\kappa}$ on its domain $(0, \infty)$, as shown in equation (19). In the case of limited commitment on the shareholder side only, u is bounded from above by the limited commitment constraint on shareholders. As $K \to \infty$, because optimal risk sharing implies that changes in U are slower than changes in K, $u = \frac{U}{K} \to 0$ in both the first-best case and the shareholder-side limited commitment case. As u approaches 0, the probability of a binding limited-commitment constraint vanishes and the normalized value function converges to that in the first-best case, $\bar{v} - \frac{1}{\mathbf{r}+\kappa}u$. As the figure illustrates, more contracting frictions reduce the efficiency of risk sharing, lower the value of the firm and shrink the region of feasible continuation utilities, u.

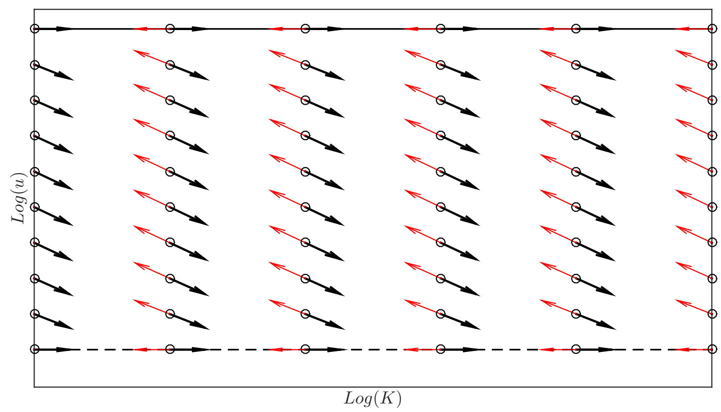
In Panel (b) of Figure 2, we plot the normalized value function, $v(u, K)$, for Specification 2 for for three levels of firm size: $K_1 < K_2 < K_3$. Note that as K increases, the domain of the normalized utility widens and the level of firm value rises. This is because as firm size increases, the market technology, which is linear in K, becomes more efficient compared to concave home production technology. As a result, the limited-commitment constraint on the manager side is less likely to bind, and risk sharing becomes more efficient. In the limit, as $K \to \infty$, the value function converges to that for the case with the shareholder-side limited commitment only.

5.4 Dynamics of State Variables

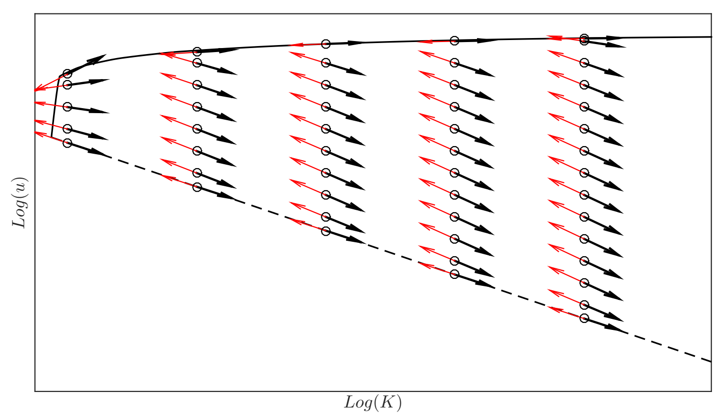
In Figure 3, we plot the diffusion of the state variables in the $(\ln K, \ln u)$ space for both model specifications. The thick arrows indicate the direction of the state variables upon a positive realization of the Brownian motion shock and the thin arrows represent movements of the state variables upon a negative realization of dB_t. The arrows always point toward the southeast or the northwest direction in the interior of the domain, indicating that u and K are negatively correlated. This is the implication of the optimal risk sharing: an increase in K is associated with a decrease in $u = \frac{U}{K}$ because continuation utility (U) is less sensitive to shocks than firm size (K). As a result, in the stationary distribution, small firms tend to be in the northwest region of the domain, and large firms tend to be in the southeast region of the domain. Also, because $(\ln K, \ln u)$ must stay in its domain with probability one (unless

[22] 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.

at K_{MIN} where firms exit the economy) for the limited commitment constraint to hold, the direction of the diffusion must be tangent to $u_{MIN}(K)$ and $u_{MAX}(K)$ on the boundaries.



(a) Specification 1

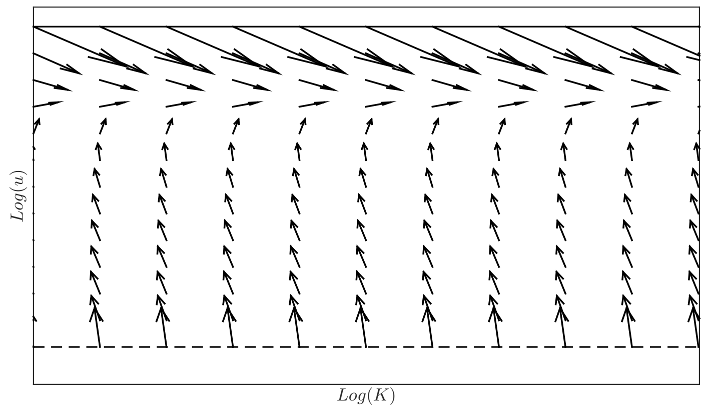


(b) Specification 2

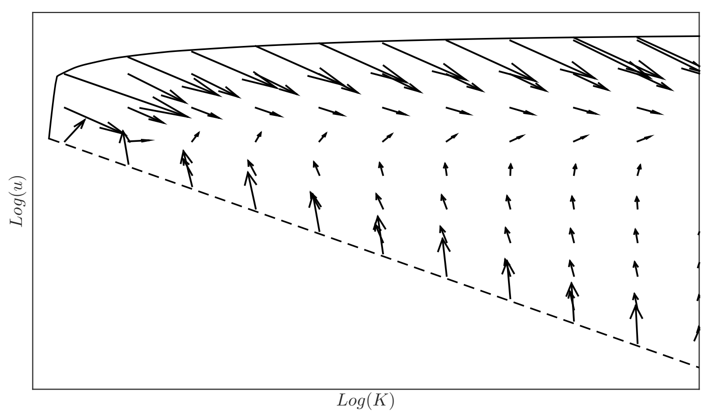
Figure 3: Movement of State Variable upon Unexpected Brownian Motion Shocks

Figure 3 plots the direction of movement of state variables for Specification 1 (top panel) and that for Specification 2 (bottom panel). The movement of the state variables upon a positive Brownian motion shock is illustrated by thick arrows and that upon a negative shock is depicted by thin arrows. On the boundaries, the arrows must be tangent to the boundary so that the limited commitment constraint is never violated. In the interior, the arrows generally point to southeast upon a positive shock and northwest upon a negative shock due to optimal risk-sharing.

Figure 4 plots the expected change (or the drift) of the state variables in the $(\ln K, \ln u)$ space. Note that at the two boundaries, the limited-commitment constraints bind and prevent perfect risk sharing. As a result, the arrows point away from these boundaries because efficiency requires that firms stay away from the binding constraints as much as possible. In addition, arrows point strongly to the right for large values of u. Because the length of the arrows is proportional to their magnitude, this pattern indicates rapid growth in size for firms close to $u_{MAX}(K)$. By comparison, the magnitude in growth rates is much smaller for firms close the manager-side limited commitment constraint, $u_{MIN}(K)$. As we show below, these patterns are due to dynamically optimal risk-sharing between shareholders and managers and are consistent with the inverse relationship between size and growth rates in the data.



(a) Specification 1



(b) Specification 2

Figure 4: The Expected Movement (Drift) of State Variables

Figure 4 illustrates the direction of expected movement (drift) of the state variables for Specification 1 (top panel) and that for Specification 2 (bottom panel). The state variables have a strong tendency to move away from the boundaries because the latter represent binding limited commitment constraints and limit risk sharing. Optimal risk sharing implies that firms close to the shareholder side limited commitment constraint ($u_{MAX}(K)$) invest more and growth faster in the $\ln K$ direction to allow for better diversification for managers. For the same reason, firms close to the manager side limited commitment constraint reduce investment to limit firm size and therefore managers' outside options.

5.5 Compensation Policy

Proposition 4 summarizes two important properties of the compensation policy. First, part 2 of the proposition implies that compensation must stay constant whenever the limited-commitment constraints do not bind. The intuition for this result is best illustrated using an expected utility representation of the manager's continuation utility. If we define $\hat{U} = \frac{1}{1-\gamma}U^{1-\gamma}$, it follows from equation (6) that $\hat{U}$ is additively separable with respect to consumption across time and states. It is not hard to prove that $\frac{\partial V}{\partial \hat{U}} = \frac{1}{C^{-\gamma}}$ (see equation (C.3) in Appendix C). Intuitively, this condition says that the marginal cost of utility provision for the shareholder must equal the inverse of the marginal utility of the manager: providing one addition unit of utility to the manager requires $1/C^{-\gamma}$ unit of consumption goods.[23] Note that optimality requires the marginal cost of utility provision to be equalized across time and future states if incentive compatibility constraint does not require otherwise. In our model, this means that in the interior of the domain, $\frac{\partial V}{\partial \hat{U}}$ must remain constant. Hence, C must also stay constant.[24]

Second, whenever the limited-commitment constraints bind, the optimal contract implements a minimum modification to managerial pay to keep the constraints from being violated. This result is due to the requirement of the optimal risk sharing and is best formalized by the regulated Brownian characterization of compensation policy in Specification 1 of our model. Using the definition of normalized compensation, $\frac{C_t}{K_t} = c(u_t)$, equation (28) in part 3 of Proposition 4 implies

$$\ln C_t = \log C_0 + l_t^+ - l_t^- . \tag{29}$$

Intuitively, l_t^+ is the minimum raise in managerial compensation needed to keep managers from taking their outside options, and l_t^- represents the minimum reduction in managerial pay to keep firm value from being negative. As regulators of Brown motions, l_t^+ and l_t^- increase only at discrete time points and never decrease. Therefore, managerial compensation in our model stays constant most of the time, and moves only occasionally to keep the limited commitment constraints from being violated.

In Figure 5, we illustrate the sample path of a firm starting from the interior of $[u_{MIN}, u_{MAX}]$. The top panel is the trajectory of the log size of the firm, and the second panel is the path of the normalized utility. The third panel is the corresponding realizations of the

[23] This is commonly known as the "inverse Euler equation" in discrete time dynamic contracting problems (see, for example, Rogerson [53]).

[24] Note that the above agrement is not true if we replace $\hat{U}$ with U. That is $\frac{\partial V}{\partial U}$ does not need to remain constant across states and over time because unlike $\hat{U}$, the aggregation of U is not additively separable with respect to time and states.

value of the firm, and the bottom panel shows the log managerial compensation.[25] At time

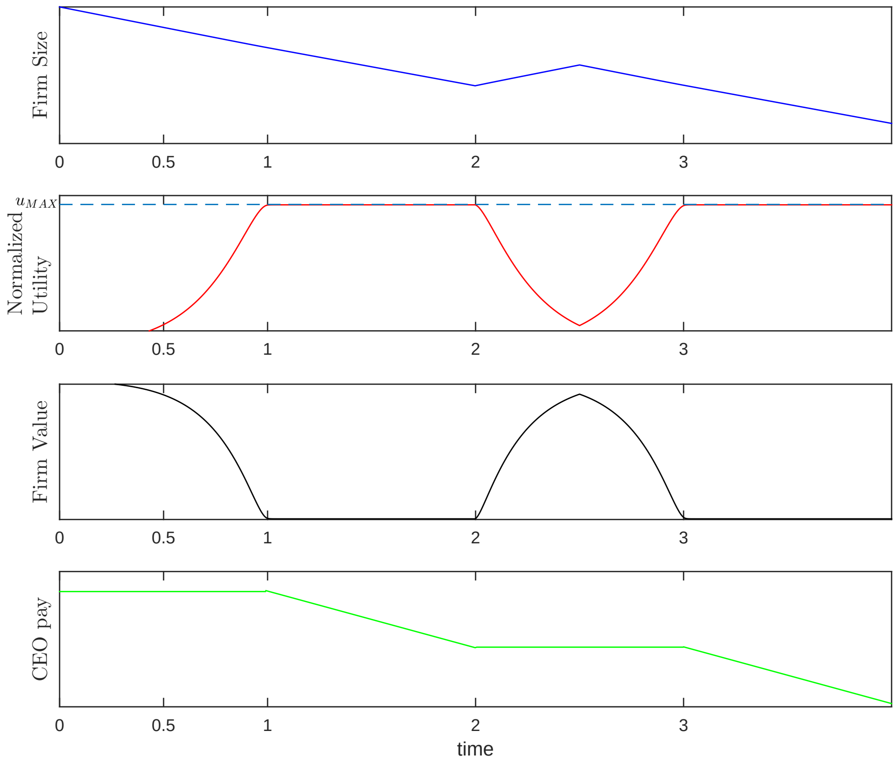


Figure 5: Sample Path of CEO Compensation: The Bankruptcy Constraint

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 bankruptcy point, u_{MAX}.

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 1 lowers capital stock of the firm (top panel). For $t < 1$, $u_t < u_{MAX}$ is in the interior (second panel). In this region, firm value is strictly positive (third panel) and managerial compensation is constant (bottom panel). At $t = 1$, u_t hits the boundary u_{MAX} and cannot increase further despite subsequent negative productivity shocks. For $t \in (1, 2)$, the firm continues to receive a sequence of negative productivity shocks and the total capital stock of the firm shrinks further (top panel). During this period, u_t stays at u_{MAX}, where the shareholder-side limited-commitment constraint binds, as shown in the second panel of Figure 5. The firm value remains at zero and does not cross over to the negative region due to reductions in managerial compensation, which keeps decreasing

[25] $\log K_t$ 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.

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. 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 6, we plot a sample path of a firm with u_0 close to the left boundary, u_{MIN}. At

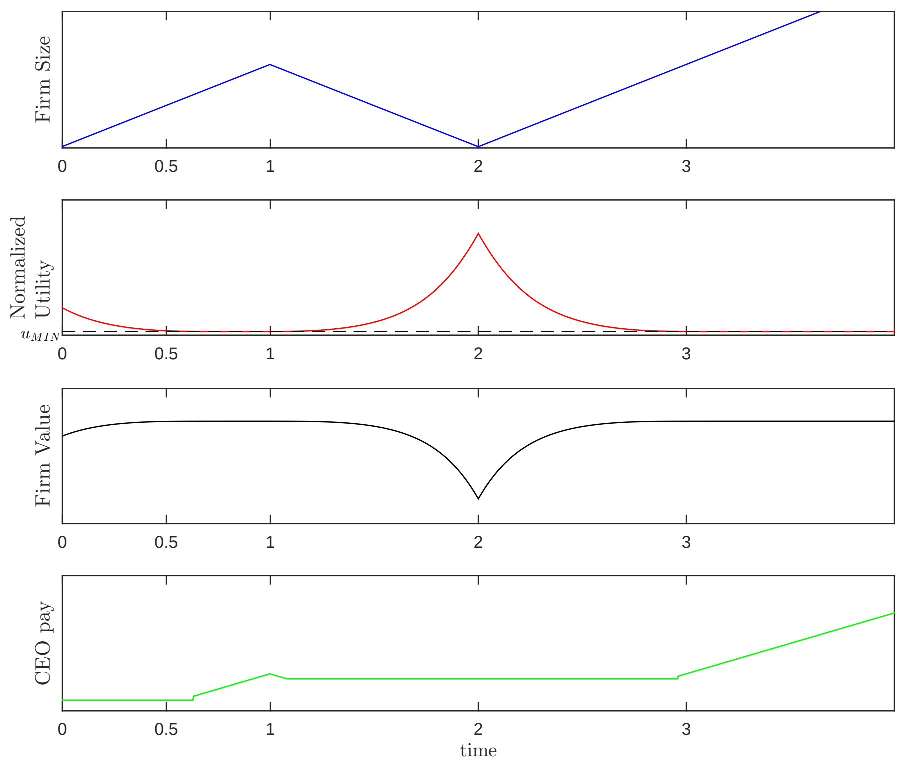


Figure 6: Sample Path of CEO Compensation: The Participation Constraint

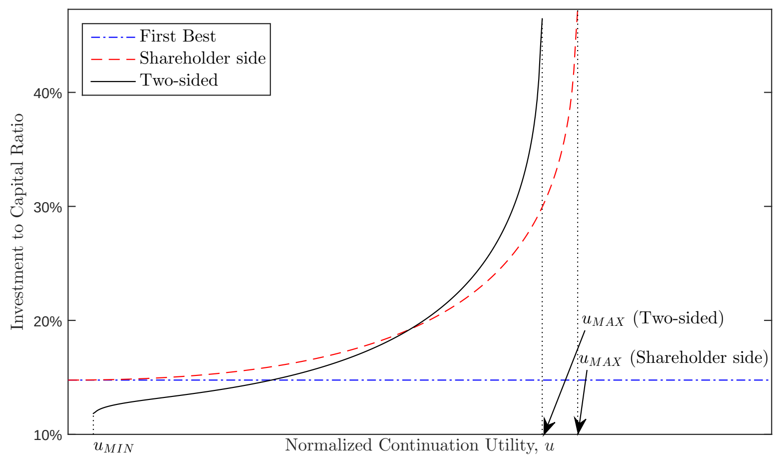
Figure 6 plots sample paths of firm size (top panel), normalized continuation utility (second panel), firm value (third panel), and log CEO pay (bottom panel) in the neighborhood of the binding participation constraint, u_{MIN}.

time 0, the firm starts from the interior of the normalized utility space, $u_{MIN} < u_0 < u_{MAX}$. A sequence of positive productivity shocks from time 0 to 0.5 increases 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 normalized firm value, $v(u_t)$, increase. At time 0.5, the normalized continuation utility reaches the left boundary, u_{MIN}, and the manager-side limited-commitment 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 4). During $t \in (1, 3)$, 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 manager-side limited-commitment constraint starts to bind again and, as a result, manager's compensation increases.

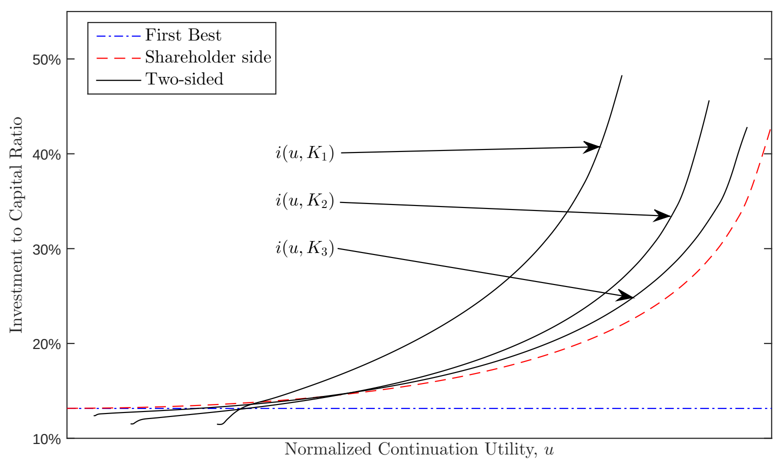
5.6 Investment Policy

The optimal investment-to-capital ratio monotonically increases with the normalized continuation utility under both model specifications. We plot the optimal investment rate as a function of the normalized utility in Figure 7. Because the optimal investment rate in Specification 2 depends both on the normalized utility and firm size, in Panel (b) we plot $i(u, K)$ for three levels of firm size: $K_1 < K_2 < K_3$. Consistent with the pattern in value functions, the dashed line, which represents the optimal investment-to-capital ratio for the case with shareholder side limited commitment only, converges to the first-best investment level as $u \to 0$. Also note that in Specification 2, as $K \to \infty$, the optimal investment rate converges to that in the case with the shareholder-side limited commitment only because managers' outside options become negligible compared to the size of the firm's cash flow.

Under both specifications, limited commitment generates a negative relationship between the normalized utility and rate of investment. As shown above, the optimal risk sharing implies that u and K are negatively correlated. Hence, small firms in our model invest more than large firms despite the constant return to scale technology. Intuitively, the shareholder-side limited commitment implies that managers in small firms are poorly diversified and therefore it is optimal for them to accelerate investment and grow out of the constraint. Managers in large firms are more likely to take their outside options. In response, large firms reduce investment to limit managers outside options and prevent them from default. Therefore, in the model with two-sided limited commitment, small firms over-invest and large



(a) Specification 1



(b) Specification 2

Figure 7: Investment Policy: Two-Sided Limited Commitment

Figure 7 plots the optimal investment rates ($\frac{I}{K}$) as a function of the state variables (u, K) for Specification 1 (top panel) and those for Specification 2 (bottom panel). The optimal investment rate in the first-best case is represented by the dash-dotted line and that in the case with one-sided limited commitment on shareholder side is depicted by the dashed line in both panels. Investment is increasing in normalized utility in both specifications. In the bottom panel, we plot investment policy for three different levels of capital stock, $K_1 < K_2 < K_3$. As the level of K increases, the optimal investment rate approaches that of the model with shareholder side limited commitment constraint only.

firms under-invest relative to the first-best model.

Our model provides an alternative explanation for the negative relationship between investment and firm size in the data in a way that is consistent with a power law in firm size. Although decreasing return to scale combined with adjustment cost and/or some form of agency frictions can also be consistent with the fact that small firms invest more than large firms, models with decreasing return to scale typically imply that in the long run firms are concentrated around their optimal sizes and the distribution of firm size is unlikely to have a fat tail. In our model, constant return to scale implies that large firms do not stop growing and generates a power law in firm size. At the same time, small firms invest more than large firms to mitigate limited commitment frictions.

6 Quantitative Results

In this section, we present the quantitative implications of our two-sided limited commitment model and discuss its ability to account for key characteristics of the empirical distribution of firm size, CEO compensation, investment and dividend policies. The cross-sectional data that we use consist of US non-financial firms and come from the Center for Research in Securities Prices (CRSP) and Compustat. They are sampled on the annual frequency and cover the period from 1992 till 2011. Our data set is standard and we refer to Appendix F for a detailed description of the data.

6.1 Calibration

We calibrate the two specifications of our two-sided limited commitment model that differ in terms of managers' outside options. We choose parameter values of Specification 1 to match a set of key aggregate moments. To facilitate the comparison across model specifications, in Specification 2, we keep all parameters the same except those that govern managers' outside options.

We follow the standards of the macroeconomics literature to calibrate preference and technology parameters; see, for example, Kydland and Prescott [40], King and Rebelo [37] and Rouwenhorst [54]. We choose a risk aversion (γ) of 2 and set the discount rate (β) that determines the risk-free interest rate at 4% per year.[26] The capital share parameter α is set at 0.36. We calibrate the exogenous firm death rate, κ, to be 5% per year and choose $\delta = 7\%$ that together with the exit rate imply a 12% effective annual depreciation rate of capital. The volatility parameter σ is set at 35% to match the average volatility of firms' sales growth in the data.

We calibrate the equilibrium marginal product of capital, $\mathbf{A}$, without explicitly specifying

[26] This allows our model to match the average return of risky and risk-free assets in the data, as in Kydland and Prescott [40].

the level of aggregate productivity, $\mathbf{z}$. Note that at steady state, aggregate investment must be related to aggregate capital stock by $\mathbf{I} = (\kappa + \delta)\,\mathbf{K}$. Because the aggregate labor supply is normalized to one, the output-to-investment ratio is given by:

$$\frac{\mathbf{Y}}{\mathbf{I}} = \frac{\mathbf{z}\mathbf{K}^{\alpha}}{(\kappa+\delta)\,\mathbf{K}} = \frac{1}{\kappa+\delta}\mathbf{z}\mathbf{K}^{\alpha-\mathbf{1}}. \tag{30}$$

Hence, $\mathbf{A}$ is proportional to the aggregate investment-to-output ratio and can be chosen to match the corresponding moment in the data, i.e.,

$$\mathbf{A} = \alpha\mathbf{z}\mathbf{K}^{\alpha-\mathbf{1}} = \alpha\,(\kappa+\delta)\,\frac{\mathbf{Y}}{\mathbf{I}}. \tag{31}$$

We set $\mathbf{A} = 0.231$, which implies an investment-to-output ratio of 18.7%.[27]

We choose capital adjustment cost parameter $\phi = 5$ to account for the average market-to-book ratio (Tobin's Q) in the data and set the initial promised utility, $\bar{U}$, at 0.0879 to match the observed CEO pay to capital ratio of young firms. Two additional moments that we target in our calibration comprise the cross-sectional mean and median of firm growth rates. The calibrated parameter values and the set of moments that they are based on are summarized in Tables 1 and 2.

We assume that the cost function for the creation of new firms is of constant elasticity:

$$H\,(K) = \frac{\psi_0}{1+\psi_1}K^{1+\psi_1}, \tag{32}$$

and we choose the parameters ψ_0 and ψ_1 so that the initial size of firms is normalized to one and the profit of setting up a new firm is zero. Specifically, $\psi_0 = 1.90$ and $\psi_1 = 2.18$.

In Specification 2, we retain the same parameter values as in Specification 1, including the productivity parameter, $\mathbf{z}$.[28] This leaves us with only one additional parameter to calibrate: the managers' endowment of home labor, $\bar{n}$. Because $\bar{n}$ affects managers' outside options relative to their compensation, we choose $\bar{n} = 0.011$ so that the implied average CEO compensation-to-capital ratio is the same under the two model-specifications.[29]

We solve the model numerically, simulate it for 500 years, and discard the first 400 years of data. We aggregate simulated data from the continuous-time model to an annual frequency that corresponds to sampling frequency of the observed data. Our simulated sample consists of two million firms and, as such, can be treated as population.

[27]Note that we can then back out the associated productivity parameter $\mathbf{z}$.

[28]Note that in Specification 2, the death rate is endogenous, therefore, we can no longer use equation (31) to calibrate $\mathbf{A}$.

[29]We also verify that the profit of entrance is positive in Specification 2 under our calibration.

6.2 Size, Age, and Growth Dynamics

In this section, we discuss the cross-sectional distributions of key growth moments in the data and in the model. In what follows, we present average characteristics of quintile (or decile) portfolios sorted by firm size. We follow the standard sorting procedure in the data by assigning firms into portfolios according to their size using breakpoints based on the NYSE-traded firms. In the model, firms are sorted using breakpoints that are equally-spaced in log size. Portfolios are re-balanced at the annual frequency. We consider two measures of firm size in the data: the number of employees and gross capital. In the model, we employ a single size measure because, as discussed before, the two measures of size (capital and labor) are equivalent.[30]

We begin with the relationship between age and size. Figure 8 plots the median age across decile portfolios sorted on size. Panel (a) presents the data, and Panel (b) shows the implications of the two model specifications. Notice that the data features a nearly monotonic positive relationship between age and size. Smaller firms, on average, are significantly younger than larger firms. While a model with positive average growth rate is likely to generate a positive correlation between size and age, the monotonic pattern observed in the data is more challenging as shown in Panel (b).

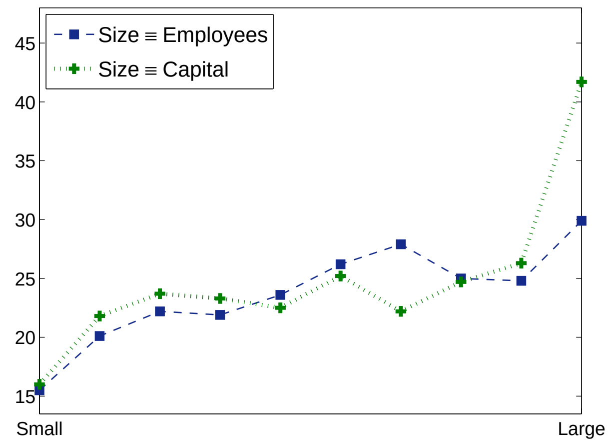
Although Specification 1 overall generates a positive correlation between size and age, the age-size relationship is not monotonic. Here, firms in the left tail of the distribution are on average older then those in the center of the distribution due to the absence of endogenous exit. Because growth rates are stochastic, the very small firms are not new comers but are those that have experienced long sequences of negative productivity shocks. Thus, without endogenous death, firms in the left and right tail of the size distribution are on average older.

Specification 2 of our model features endogenous death. In this version of the model, the home production technology is less scalable than the market technology and, therefore, is less affected by negative shocks to firm size. As a result, small firms optimally abandon market production and exit the economy. Hence, most of the small firms are new entrants and the age-size relationship is positive and monotonic.

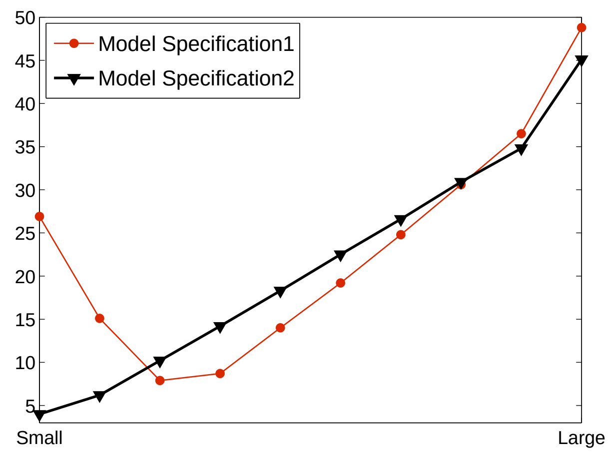
Despite its simplicity, our model captures the distribution of firm age quite well. Luttmer [45] documents that the median age of firms with more than $10,000$ employees in 2008 was about 75 years. In our model, this number is 93 for Specification 1 and 91 for Specification 2.[31] Luttmer [46] shows that to jointly account for the decline in volatility with firm size

[30]The sorting procedures that we use generate a meaningful and sizable heterogeneity in firm size. In particular, in the data, firms in the lowest and highest quintile portfolios on average account for about 4% and 65% of capital stock, respectively. In the model, the bottom and the top quintiles own about 3% and 60% of capital, on average.

[31]There are about 6 million employer firms in 2008 in the US and the largest 1000 firms with more than



(a) Data



(b) Model

Figure 8: Distribution of Age across Size

Figure 8 shows the average firm age across ten size-sorted portfolios. Size in the data is measured by either the number of firm employees or gross capital.

and the existence of young and large firms, one needs a mechanism where young and small firms grow faster than the population. Although our model does not account for the decline in volatility with respect to size, limited commitment does create an inverse relationship between firm growth and firm size (and age) as we show next.

Table 3 shows the average investment rate (defined as annual investment divided by

10,000 employees account for 27% of the total employment (Luttmer [45]). In our calibration, the median age of the largest firms that account for 27% of total employment of the economy is 93 years in Specification 1 and 91 years in Specification 2.

beginning-of-year capital stock) as a function of firm size. In the data, small firms invest at a higher rate of about 17% per year relative to large firms, which on average, invest at an about 9% rate. As the table shows, the difference in investment rates of large and small firms is strongly statistically significant. The right two columns show that both versions of our model are able to replicate the observed negative relationship between investment and firm size. The model-implied difference in investment rates of firms in the bottom and top quintiles is about 10%.[32]

Table 4 reports the average growth rates of size-sorted portfolios. Consistent with the previous literature (for example, Evans [23, 24] and Hall [31]) we find that in the data small firms grow at a significantly higher rate than large firms and this pattern is robust to different measures of firm size. On average, small firms grow by about 10% faster compared with large firms. Our model generates a similar cross-sectional variation in growth rates. The difference in average growth rates between small and large firms implied by the two model specifications is about 11-12%. As explained above, managers in small firms are poorly diversified due to the limited commitment on the shareholder side. As a result, small firms accelerate investment to grow out of the agency conflict. As firm size increases, managers' outside option becomes more attractive. Therefore, large firms have a strong incentive to reduce investment and curtail growth to limit managers' outside options.

Table 5 shows the fraction of dividend-paying firms in each size-sorted portfolio. In the data, large firms are much more likely to pay dividends to shareholders than small firms. In the bottom size quintile, on average, only one out of ten firms pays dividends. The fraction of dividend-paying firms increases to about 70-80% in the right tail of the size distribution. Our model similarly implies a monotonic increase in the fraction of dividend-paying firms with size. Small firms in the model use most of their resources for investment and, therefore, tend not to pay dividends. These firms start distributing profits to shareholders as they grow and become less constrained.

Empirically, small firms are more likely to fail and exit the market than large firms. As Table 6 shows, on average, exit rates of small firms are about twice as high compared with large firms. The model predictions vary depending on the model specification. Under Specification 1, firm death rate does not depend on size because this version of the model has no endogenous exit. Under Specification 2, similar to the data, small firms exit more frequently relative to large firms. However, in the data, exit rates monotonically decrease with size, while in our second specification, death rates of medium and large firms are virtually identical. Note that for simplicity, we assume that, upon entrance, all firms have the same initial condition. Specification 2 of our model can easily generate a smooth variation of exit

[32]We do not present t-statistics of the differences in various moments in the model columns because the reported model statistics represent population moments.

rates with size if we specify a smooth entrance density of new firms.[33]

In summary, our model captures the stylized empirical features of firm growth dynamics. We now turn to the implications for power law in firm size, dividend payout, and CEO pay.

6.3 Power Laws in Size, Dividends, and CEO Compensation

As discussed above, our model provides a unified explanation of power-law behavior of the right tail of firm size, CEO compensation and dividend payout. We first discuss our empirical estimates of the power laws and then compare the quantitative implications of our model to the data.

Following Luttmer [44] and Gabaix [26], we use the following parametrization of power law. The distribution of random variable X obeys a power law if its density is of the form:

$$f(x) \propto x^{-(1+\xi)},$$

for some constant $\xi > 0$. The parameter ξ is called the power-law exponent. The complementary cumulative distribution function of X is given by:

$$P(X > x) \propto x^{-\xi}.$$

That is, the complementary distribution of a power-law variable is log linear with slope $-\xi$.

It has been shown in the literature that firm size follows a power-law distribution (for example, Axtell [7], Gabaix [26], and Luttmer [44]). We confirm this evidence and show that empirical distributions of CEO compensation and dividends are also fat-tailed. Year-by-year estimates of the power-law coefficient for firm size measured by either the total number of employees or market capitalization, dividend payout, and CEO compensation are presented in Table 7. The table reports the estimates and the corresponding p-values of the Kolmogorov-Smirnov goodness-of-fit test constructed via bootstrap. The details of the estimation procedure are provided in Appendix G.

On average, the estimate of the power-law coefficient of firm size is about 1.26 when size is measured by the number of employees and about 1.09 if size is measured by market capitalization. The latter is very close to the estimates obtained using Census data. For example, Luttmer [44] reports a power law estimate of 1.07; similar estimates are reported in Gabaix and Landier [27]. Overall, the goodness-of-fit test does not reject the power-

[33] We do not present the dependence of exit rates on firm age. However, given that in Specification 2, age is a monotone function of size, this version of the model is also consistent with the negative relationship between exit rates and age observed in the data.

law null especially in market capitalization, for which the p-value is above the conventional five-percent level in all but two sample years.[34]

The power-law coefficient of dividends is remarkably close to that of size, particularly market capitalization. In contrast, CEO compensation is characterized by a much larger power-law coefficient of about 2.1. That is, dividend payout and firm size seem to feature similar behavior in the right tail, whereas the right tail of CEO compensation is significantly thinner.

Similar to the data, our calibrated model produces a power law in firm size and dividends with a slope close to one. In particular, the tail slope of firm size and dividend payout is 1.09 under Specification 1 and 1.04 under Specification 2.[35] Figure 9 provides a visual comparison of the tail behavior of the model-implied distribution and the empirical distribution constructed using one year of the sample data. The horizontal axis in the figure

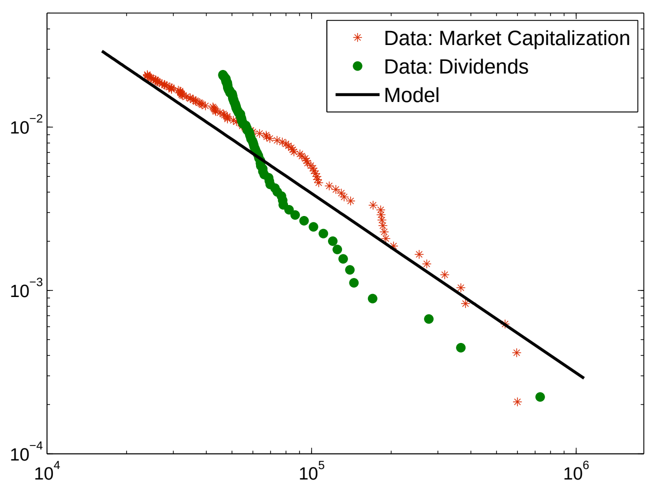


Figure 9: The Right Tail of the Distribution of Market Capitalization and Dividends

Figure 9 plots the right tail of the distribution of market capitalization and dividends (using 2000 data) and the slope implied by the model.

represents market capitalization (and re-scaled dividends) and the vertical axis shows the complementary cumulative distribution function, both are equally spaced on the log scale. In the data, market capitalization is represented by stars and dividends are represented by circles. The solid line is the power law in both size and dividends generated by Specification

[34] We also find a significant evidence of a power law in gross capital with the average estimate of the exponent of about 1.4 and average p-value of 0.32. This evidence is not reported due to space constraints.

[35] Specification 2 produces a slightly fatter tail of the distribution of firm size relative to Specification 1 because the manager-side limited commitment constraint is less likely to bind in large firms to hold back their growth.

1 of our model.[36] Under power law, the log-log plot is a straight line with a slope equal to the negative exponent. As the figure shows, overall, the model-implied slope matches the slope observed in the data quite well.

As shown in Section 4, limited commitment on the manager side implies that CEO compensation in large firms is a linear function of the running maximum of K^ν and, therefore, obeys the same power law as K^ν. That is, the optimal contract under limited commitment makes a power law in firm size translate into a power law in CEO compensation. Note that the power law in CEO pay is a limiting result that applies to firms in the right tail of the size distribution. Thus, our power law results for CEO compensation remain valid in the model with two-sided limited commitment.

In Figure 10, we show the complementary cumulative distribution function of CEO compensation in the data and in our model. The power law in CEO pay generated by Specification 1 of our model is represented by the solid line with a slope of (negative) 1.09, which is the same as that of firm size. Specification 2 is represented by the dashed line with a power-law coefficient of $\frac{1.04}{0.36} = 2.89$. The observed CEO compensation data (using data of 2000) is represented by circles. As the figure shows, the power-law exponent in the data is in-between the slopes implied by the two model specifications.

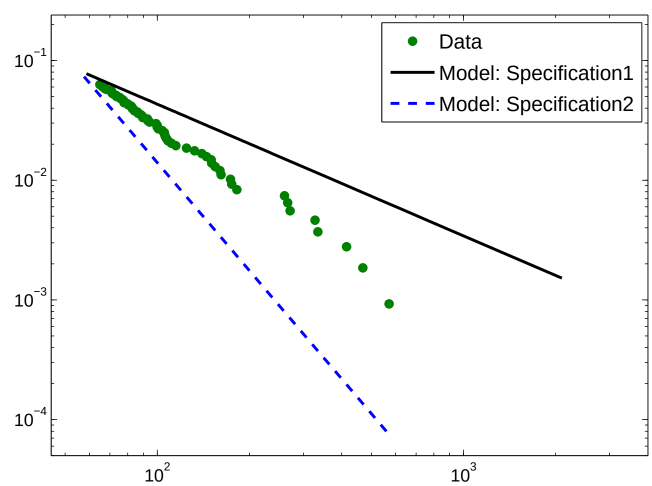


Figure 10: The Right Tail of the Distribution of CEO Compensation

Figure 10 plots the right tail of the distribution of CEO compensation in the data (using 2000 data) and the slope implied by the two model specifications.

In Specification 1, because managers upon default are allowed to participate in the

[36]We do not plot Specification 2 because the two model lines would be very close to each other and hard to distinguish.

competitive labor market, capital is perfectly scalable and managers' outside options are proportional to firm size. As a result, the power law in CEO pay is the same as the power law in firm size. Under Specification 2, managers are excluded from labor market after default, and capital is no longer scalable. Therefore, managers' outside options are proportional to K^{α} and the power law in CEO pay is $\frac{1}{\alpha}$ times of the power law in firm size. As shown in Figure 10, Specification 1 overstates the thickness of the right tail of CEO pay while Specification 2 understates it. Note that we could always choose the parameters of the home production technology to match exactly the power law in CEO compensation data. We do not entertain this approach here as we think that the choice of the home production technology should be ultimately guided by micro-evidence that is not available. Our objective is not to match the data estimates exactly but rather to show that limited commitment on the manager side goes a long way in explaining the right-tail behavior of the CEO pay that we document in the data.

6.4 CEO Compensation and Firm Size

The cross-sectional distribution of CEO compensation is characterized by several stylized features. While the level of managerial pay increases with firm size, the increase in less than proportional. Therefore, CEO's of small firms have a significantly larger stake in their companies relative to managers of large firms. Table 8 shows the variation in the CEO pay-to-capital ratio across size-sorted portfolios. In the data, the median ratio falls from 5.9% for small firms to 0.4% for large firms when size is measured by the number of firm employees, and from 7.3% to 0.1% when size is measured by firm capital. The model features a similar cross-sectional pattern – CEO compensation scaled by firm size declines significantly as size increases under both specifications. In the model, the negative relationship between CEO pay-to-capital ratio and firm size is an implication of (imperfect) risk sharing. Risk sharing reduces sensitivity of CEO compensation to productivity shocks. Therefore, as firms become large, CEO compensation as a fraction of firm size declines.

Similar to the previous literature (see Gabaix [26]), we find that in the data, the average elasticity of CEO compensation to size is about one-third and is very similar across alternative measures of firm size. Table 9 shows the elasticities of managerial compensation to market capitalization, the number of firm employees and capital. The elasticities are estimated in a panel regression of CEO pay on lagged firm size (both measured in logs), controlling for firm and time fixed effects. In the model, the magnitude of size elasticity of CEO compensation is 0.25 and 0.05 under Specifications 1 and 2, respectively. Note that the model-implied elasticities are smaller than in the data because limited commitment is the sole mechanism that generates a positive relationship between CEO pay and firm size. Other agency frictions (for example, moral hazard) are likely to increase the response of managerial compensation to size and help account for the magnitudes observed in the data.

While managerial compensation in the model generally increases with firm size, the elasticity of CEO compensation to size is not uniform – it is mostly driven by small and large firms and is virtually zero for medium-size firms. In fact, a V-shape of size elasticity of CEO pay is one of the signature features of our two-sided limited commitment model. In the model, managerial compensation varies with firm size either due to a binding shareholder constraint or a binding manager constraint. Because small firms tend to be close to the exit threshold, and large firms tend to face a binding manager-side constraint, CEO compensation is more sensitive to size for very small and very large firms. We find that in the data, the dependance of managerial pay on size also tends to be stronger for firms in the left and right tails of size distribution.

Table 10 shows variation of size elasticity of CEO pay across quintile portfolios sorted on firm size.[37] As the table shows, the model-implied elasticity is close to zero for medium-size firms, and is quantitatively large for firms in the bottom and the top size portfolios. Under Specification 1, size elasticity of CEO compensation is equal to 0.42, 0.00 and 0.63 for firms in the bottom, middle and top portfolios, respectively. Under Specification 2, only small and large firms feature non-trivial elasticities of 0.08 and 0.17, respectively. In the data, the elasticity is likewise V-shaped. For example, when firm size is measured by market capitalization, the sensitivity of CEO pay to size falls from 0.29 (SE=0.04) for small firms to 0.17 (SE=0.08) for medium-size firms and increases to 0.38 (SE=0.04) for large firms. The cross-sectional pattern in size elasticities is similar if firm size is measured by the number of employees or firm capital.

The V-shaped elasticity of CEO pay with respect to firm size is consistent with our empirical evidence on power law. The power law coefficient for firm size in the data is about 1.1. A uniform elasticity of CEO pay with respect to firm size of $\frac{1}{3}$ would imply a power-law exponent of CEO compensation of $1.1 \times 3 = 3.3$. However, the distribution of CEO compensation in the data has a fatter tail, with a power law coefficient of about 2.1, which is consistent with a higher than average size elasticity of CEO pay in large firms.

Our model has important implications for the time-series dynamics of CEO compensation in response to variation in firm size. Note that while the high sensitivity of CEO pay to size is a feature shared by both small and large firms, its mechanism and, therefore, the dependance of managerial compensation on firm history are quite different. CEO compensation of small firms is sensitive to negative productivity shocks because small firms are more likely to run

[37] To ensure that there are enough data to estimate firm fixed effects, we exclude firms with less than ten observations in a given portfolio.

into a binding limited commitment constraint on the shareholder side. For small firms, a decline in firm size relative to its running minimum is likely to lead to a decline in CEO compensation. In contrast, managerial compensation of large firms is sensitive to positive productivity shocks. For large firms, a sizable increase in firm size (an increase relative to its running maximum) is likely to result in an increase in CEO compensation.[38] We test these model-implied dynamics using panel regression analysis. We elaborate on the model predictions further below after we specify our regression and introduce relevant notation.

We run the following panel regression:

$$\Delta c_{i,t} \sim \left\{ \Delta k_{i,t-1},\ \Delta^{Min-} k_{i,t} \cdot I^{Small}_{i,t},\ \Delta^{Min+} k_{i,t} \cdot I^{Small}_{i,t},\ \Delta^{Max-} k_{i,t} \cdot I^{Large}_{i,t},\ \Delta^{Max+} k_{i,t} \cdot I^{Large}_{i,t} \right\}$$

We regress the log-growth of CEO compensation ($\Delta c_{i,t} \equiv \log \frac{C_{i,t}}{C_{i,t-1}}$) on the log-growth of firm size ($\Delta k_{i,t-1} \equiv \log \frac{K_{i,t-1}}{K_{i,t-2}}$), and four interaction terms that correspond to opposite changes in size realized by small and large firms. For each firm i, we compute the change in size at time t relative to its (previous) running minimum and maximum: $\Delta^{Min} k_{i,t} \equiv \log \frac{K_{i,t}}{K^{Min}_{i,t-1}}$ and $\Delta^{Max} k_{i,t} \equiv \log \frac{K_{i,t}}{K^{Max}_{i,t-1}}$, where $K^{Min}_{i,t-1}$ and $K^{Max}_{i,t-1}$ are respectively minimum and maximum of firm size observed in the five years prior to year t. We focus separately on negative and positive changes in firm size: $\Delta^{Min-} k_{i,t}$ and $\Delta^{Max-} k_{i,t}$ represent declines, and $\Delta^{Min+} k_{i,t}$ and $\Delta^{Max+} k_{i,t}$ represent increases in firms size relative to the running minimum and running maximum, respectively. $I^{Small}_{i,t}$ and $I^{Large}_{i,t}$ are size dummies that select firms either in the bottom or the top decile of size distribution at the beginning of year t. We use market capitalization to measure size in the data and control for firm and time fixed effects. The regression coefficients are estimated using annual data but we use monthly series to obtain more accurate measures of running minimum and maximum of firm size. Standard errors are clustered by firm and time to ensure robustness of our inference to the cross-sectional dependence and serial correlation in residuals.

As discussed above, CEO compensation is sensitive to size for firms that are approaching one of the two commitment constraints. Small firms have to cut down managerial compensation if firm value is at the risk of falling below zero. Thus, the model predicts a positive coefficient on the first interaction term $\Delta^{Min-} k_{i,t} \cdot I^{Small}_{i,t}$. Large firms have to offer higher compensation to retain their managers if they continue to grow fast. Thus, we expect to see a positive coefficient on the last term $\Delta^{Max+} k_{i,t} \cdot I^{Large}_{i,t}$.

Panel A of Table 11 presents the estimates of our panel regression. First, notice that in

[38] As we show in Section 4, the running minimum and the running maximum of firm size are sufficient statistics for the optimal contract in models with one-sided limited commitment. In our full model with two-sided limited commitment, they can no longer summarize firm history completely but they do remain highly informative of binding limited-commitment constraints.

the model, the coefficients on $\Delta^{Min-}k_{i,t} \cdot I_{i,t}^{Small}$ and $\Delta^{Max+}k_{i,t} \cdot I_{i,t}^{Large}$ are indeed large and positive: 0.53 and 0.76, respectively. For small firms that are declining and large firms that are growing, CEO compensation is highly sensitive to changes in firm size. The coefficients on the other two interaction terms are close to zero – CEO compensation does not change for small firms that experience a positive shock and large firms that realize a negative shock. These firms are moving away from either shareholder or manager constraints and have no need to adjust managerial compensation. Also, the coefficient on the leading term ($\Delta k_{i,t-1}$) is small, of about 0.05, due to the inelastic response of CEO compensation to changes in firm value for medium-size firms. The reported numbers in "Model" column correspond to Specification 1. They are computed using a large panel of simulated data and, as such, represent population values. Specification 2 produces a similar cross-sectional pattern in elasticities and, therefore, is not reported.

"Data" column of Panel A shows the corresponding estimates in the data. Consistent with the model, size elasticity of CEO compensation in the data has a V-shaped pattern – it is higher for firms in the left and right tails and lower for firms in the middle of the distribution. Small firms (especially those with weak performance) and large firms (especially those with superior performance) feature significantly higher elasticities compared with the rest of the market. The estimates on the first and the last interaction terms, $\Delta^{Min-}k_{i,t} \cdot I_{i,t}^{Small}$ and $\Delta^{Max+}k_{i,t} \cdot I_{i,t}^{Large}$, are 1.02 and 0.78, respectively. Again, the data estimates are generally higher than the model-implied parameters, which as mentioned above is expected because in the data, CEO compensation is likely to change with firm size due to agency frictions above and beyond limited commitment.

Our evidence is similar if we estimate elasticities by running a panel regression in levels. In Panel B of Table 11, we regress the log-level of CEO compensation ($c_{i,t}$) on its lag ($c_{i,t-1}$), the log of the running maximum of firm size ($k_{i,t-1}^{Max}$) and the four interaction terms. Notice that in the model, CEO compensation is more persistent than in the data because binding limited-commitment constraints are the only channel of variation in managerial pay. Importantly, the estimates in Panel B confirm the V-shaped pattern in size elasticities – in the model and in the data, under-performing small firms and out-performing large firms are characterized by significantly higher sensitivities of CEO compensation to size relative to their counterparts and relative to medium-size firms.

7 Conclusion

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

References

[1] Ai, Hengjie, 2012, Aggregation in Economies with Homogenous Decision Rules, *Working paper, University of Minnesota.*

[2] Ai, Hengjie, and Rui Li, 2015, Investment and CEO Compensation under Limited Commitment, *Journal of Financial Economics.* 116, 452–472.

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

[4] Alvarez, Fernando, and Urban J. Jermann, 2000, Efficiency, Equilibrium, and Asset Pricing with Risk of Default, *Econometrica* 68(4), 775–797.

[5] Arellano, Cristina, Yan Bai, and Jing Zhang, 2012, Firm Dynamics and Financial Development, *Journal of Monetary Economics* 59(6), 533–549.

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

[7] Axtell, Robert L., 2001, Zipf Distribution of U.S. Firm Sizes, *Science* 293, 1818–1820.

[8] Baker, George P., Michael C. Jensen, and Kevin J. Murphy, 1988, Compensation and Incentives: Practice vs. Theory, *The Journal of Finance* 43(3), 593–616.

[9] Berk, Jonathan B., Richard Stanton, and Joseph Zechner, 2010, Human Capital, Bankruptcy, and Capital Structure, *Journal of Finance* 65.

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

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

[12] Bolton, Patrick, Neng Wang, and Jinqiang Yang, 2014, A Theory of Liquidity and Risk Management Based on the Inalienability of Risky Human Capital, *Working Paper, Columbia University.*

[13] Bond, Philip, and Ulf Axelson, 2015, Wall Street Occupations, *Journal of Finance* 70(5), 1949–1996.

[14] Clementi, Gian, and Hugo Hopenhayn, 2006, A Theory of Financing Constraints and Firm Dynamics, *Quarterly Journal of Economics* 121, 229–265.

[15] Cooley, Thomas, Ramon Marimon, and Vincenzo Quadrini, 2013, Risky Investments with Limited Commitment, *NBER Working Paper No. 19594.*

[16] DeMarzo, Peter, and Michael Fishman, 2007, Agency and Optimal Investment Dynamics, *Review of Financial Studies* 20, 151–188.

[17] DeMarzo, Peter, Michael Fishman, Zhiguo He, and Neng Wang, 2012, Dynamic Agency and q Theory of Investment, *Journal of Finance, Forthcoming* 67, 2295–2340.

[18] DeMarzo, Peter, and Yuliy Sannikov, 2006, Optimal Security Design and Dynamic Capital Structure in a Continuous-Time Agency Model, *Journal of Finance* 61, 2681–2724.

[19] Edmans, Alex, and Xavier Gabaix, 2016, Executive Compensation: A Modern Primer, *Journal of Economic Literature* forthcoming.

[20] Edmans, Alex, Xavier Gabaix, Tomasz Sadzik, and Yuliy Sannikov, 2012, Dynamic CEO Compensation, *Journal of Finance* 67.

[21] Eisfeldt, Andrea, and Dimitris Papanikolaou, 2013, Organization Capital and the Cross-Section of Expected Returns, *The Journal of Finance* LXVIII(4), 1365–1406.

[22] Eisfeldt, Andrea, and Dimitris Papanikolaou, 2014, The Value and Ownership of Intangible Capital, *American Economic Review: Papers and Proceedings* 104(5), 1–8.

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

[24] Evans, David, 1987b, Tests of Alternative Theories of Firm Growth, *Journal of Political Economy* 95, 657–674.

[25] Frydman, Carola, and Raven Saks, 2010, Executive Compensation: A New View from a Long-Run Perspective, 1936-2005, *Review of Financial Studies* 23, 2099–2138.

[26] Gabaix, Xavier, 2009, Power Laws in Economics and Finance, *Annual Review of Economics* 1, 255–294.

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

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

[29] Gordon, Myron, 1959, Dividends, Earnings and Stock Prices, *Review of Economics and Statistics* 41, 99–105.

[30] Grochulski, Borys, and Yuzhe Zhang, 2011, Optimal risk sharing and borrowing constraints in a continuous-time model with limited commitment, *Journal of Economic Theory* 146, 2356–2388.

[31] Hall, Bronwyn, 1987, The reslationship between firm size and firm growth in the US manufacturing sector, *The Journal of Industrial Economics* 35, 583–606.

[32] Harris, Milton, and Bengt Holmstrom, 1982, A Theory of Wage Dynamics, *Review of Economic Studies* 49, 315–333.

[33] Harrison, Michael, 1985, *Brownian Motion and Stochastic Flow Systems.* (John Wiley and Sons).

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

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

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

[37] King, Robert G., and Sergio T. Rebelo, 1999, Resuscitating real business cycles, in *Handbook of Macroeconomics* (Elsevier,).

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

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

[40] Kydland, Finn E., and Edward C. Prescott, 1982, Time to Build and Aggregate Fluctuations, *Econometrica* 50.

[41] Li, Shuyun May, 2013, Optimal lending contracts with long run borrowing constraints, *Journal of Economic Dynamics and Control* 37(5), 964–983.

[42] Lorenzoni, Guido, and Karl Walentin, 2007, Financial Frictions, Investment and Tobin's q, *NBER Working Paper No. 13092.*

[43] Lustig, Hanno, Chad Syverson, and Stijn Van Nieuwerburgh, 2011, Technological Change and the Growing Inequality in Managerial Compensation, *Journal of Financial Economics* 99, 601–627.

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

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

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

[47] Miao, Jianjun, and Yuzhe Zhang, forthcoming, A duality approach to continuous-time contracting problems with limited commitment, *Journal of Economic Theory.*

[48] Newey, Whitney, and Kenneth West, 1987, A simple, positive semi-definite, heteroscedasticity and autocorrelation consistent covariance matrix, *Econometrica* 55, 703–708.

[49] Quadrini, Vincenzo, 2004, Investment and Liquidation in Renegotiation-Proof Contracts with Moral Hazard, *Journal of Monetary Economics* 51, 713–751.

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

[51] Rampini, Adriano, and S. Viswanathan, 2013, Collateral and capital structure, *Journal of Financial Economics* 109, 466–492.

[52] Roberts, David R., 1956, A General Theory of Executive Compensation Based on Statistically Tested Propositions, *The Quaterly Journal of Economics* 70(2), 270–294.

[53] Rogerson, William, 1985, Repeated Moral Hazard, *Econometrica* 53, 69–76.

[54] Rouwenhorst, K. Geert, 1995, Asset pricing implications of equilibrium business cycle models, in *Frontiers of Business Cycle Research* (Princeton University Press,).

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

[56] Schmid, Lukas, 2008, A Quantitative Dynamic Agency Model of Financing Constraints, *Working paper, Duke University.*

[57] Terviö, Marko, 2008, The Difference That CEOs Make: An Assignment Model Approach, *American Economic Review* 98, 642–668.

Table 1

Calibrated Parameters

Description	Notation	Value
Capital share	α	0.36
Marginal product of capital	$\mathbf{A}$	0.231
Adjustment cost	ϕ	5
Volatility	σ	35%
Firm death rate	κ	5%
Effective depreciation	$\gamma + \delta$	12%
Manager initial utility	$\bar{U}$	0.0879

Table 1 presents the calibrated parameter values chosen to target the set of moments listed in Table 2.

Table 2

Targeted Moments

Moments	Data	Specification 1	Specification 2
Capital share	36%	36%	36%
Investment/Output	19.4%	18.7%	20.6%
Average Tobin's Q	1.64	1.81	1.76
Median sales growth	3.9%	2.8%	3.9%
Average sales growth	9.7%	9.6%	10.7%
Volatility of sales growth	37%	42%	40.6%
Average firm exit rate	3.6%	5%	6.5%
Capital depreciation	10%	12%	12.4%
CEO pay/Capital of young firms	0.082	0.073	0.071

Table 2 shows the set of moments targeted in calibrating parameters listed in Table 1. We report data statistics and the corresponding moments implied by the two model specifications. The empirical estimates of the capital share and capital depreciation are taken from Kydland and Prescott [40], and the investment-to-output ratio is computed using the National Income and Product Accounts data available on the Bureau of Economic Analysis website. All other moments are constructed using our sample data. In computing the CEO pay to capital ratio for young firms we consider firms that are less than five years old.

Table 3
Investment Rates

	Data		Model	Model
	Size≡Employees	Size≡Capital	Specification1	Specification2
Small	0.151	0.187	0.226	0.196
2	0.109	0.143	0.170	0.125
3	0.094	0.127	0.124	0.114
4	0.084	0.109	0.112	0.114
Large	0.094	0.088	0.112	0.118
Large−Small	−0.057	−0.099	−0.114	−0.079
	(−9.95)	(−5.19)		

Table 3 presents the average investment-to-capital ratio of size-sorted portfolios in the data and in the model. Size in the data is measured by either the number of firm employees or gross capital. T-statistics for the difference between large and small firms based on the Newey and West [48] estimator with four lags are reported in parentheses.

Table 4

Growth Rates

	Data		Model	Model
	Size≡Employees	Size≡Capital	Specification1	Specification2
Small	0.120	0.152	0.168	0.166
2	0.079	0.103	0.114	0.060
3	0.056	0.082	0.059	0.045
4	0.039	0.064	0.045	0.047
Large	0.032	0.033	0.045	0.051
Large−Small	−0.088	−0.118	−0.123	−0.115
	(−10.50)	(−4.14)		

Table 4 presents the average capital growth of size-sorted portfolios in the data and in the model. Size in the data is measured by either the number of firm employees or gross capital. T-statistics for the difference between large and small firms based on the Newey and West [48] estimator with four lags are reported in parentheses.

Table 5

Fraction of Dividend-Paying Firms

	Data		Model	Model
	Size≡Employees	Size≡Capital	Specification1	Specification2
Small	0.11	0.11	0.04	0.07
2	0.33	0.36	0.13	0.70
3	0.44	0.47	0.82	1.00
4	0.52	0.63	0.94	1.00
Large	0.71	0.81	0.94	1.00
Large−Small	0.60	0.71	0.90	0.93
	(26.19)	(34.19)		

Table 5 presents the average fraction of dividend-paying firms for size-sorted portfolios in the data and in the model. Size in the data is measured by either the number of firm employees or gross capital. T-statistics for the difference between large and small firms based on the Newey and West [48] estimator with four lags are reported in parentheses.

Table 6
Firm Exit Rates

	Data		Model	Model
	Size≡Employees	Size≡Capital	Specification1	Specification2
Small	0.052	0.051	0.050	0.089
2	0.055	0.050	0.050	0.050
3	0.042	0.042	0.050	0.050
4	0.037	0.033	0.050	0.050
Large	0.022	0.026	0.050	0.050
Large−Small	−0.031	−0.026	0.000	−0.039
	(−5.55)	(−5.31)		

Table 6 presents the average firm exit rate for size-sorted portfolios in the data and in the model. Size in the data is measured by either the number of firm employees or gross capital. T-statistics for the difference between large and small firms based on the Newey and West [48] estimator with four lags are reported in parentheses.

Table 7

Estimates of the Power-Law Parameters

Year	Employees		Market Cap		Dividends		CEO Pay	
	ζ	p-value	ζ	p-value	ζ	p-value	ζ	p-value
1992	1.89	0.57	1.14	0.46	1.00	0.16	1.60	0.00
1993	1.34	0.07	1.20	0.12	1.00	0.20	2.54	0.30
1994	1.12	0.01	1.10	0.11	1.05	0.07	2.21	0.09
1995	1.16	0.00	1.02	0.34	1.15	0.03	2.38	0.88
1996	1.54	0.43	1.03	0.25	1.33	0.10	2.02	0.13
1997	1.37	0.08	1.03	0.36	1.27	0.22	1.56	0.17
1998	1.00	0.09	1.00	0.14	1.15	0.35	1.59	0.04
1999	1.01	0.09	1.10	0.05	1.13	0.30	1.47	0.20
2000	1.06	0.04	1.01	0.22	1.12	0.49	1.90	0.69
2001	1.02	0.07	1.01	0.45	1.10	0.84	1.88	0.69
2002	1.29	0.02	1.02	0.54	1.18	0.71	1.56	0.00
2003	1.12	0.00	1.07	0.51	1.02	0.57	3.12	0.62
2004	1.34	0.07	1.13	0.12	1.09	0.55	3.14	0.73
2005	1.29	0.04	1.15	0.12	1.06	0.42	1.50	0.00
2006	1.26	0.06	1.21	0.20	1.01	0.16	2.55	0.71
2007	1.33	0.37	1.17	0.04	1.01	0.36	2.07	0.22
2008	1.32	0.40	1.15	0.44	1.04	0.65	2.34	0.53
2009	1.28	0.02	1.05	0.13	1.08	0.36	2.06	0.06
2010	1.21	0.12	1.01	0.38	1.07	0.33	2.21	0.08
2011	1.32	0.09	1.19	0.21	1.09	0.43	2.36	0.00
Average	1.26	0.13	1.09	0.26	1.10	0.36	2.10	0.31

Table 7 presents the estimates of the exponent of the power-law distribution (ζ) for the number of firm employees, market capitalization, dividends and CEO compensation, and p-values of the Kolmogorov-Smirnov goodness-of-fit test.

Table 8

CEO Pay-to-Capital Ratio

	Data		Model	Model
	Size≡Employees	Size≡Capital	Specification1	Specification2
Small	0.059	0.073	0.138	0.088
2	0.021	0.019	0.087	0.042
3	0.013	0.009	0.036	0.017
4	0.008	0.005	0.017	0.007
Large	0.004	0.001	0.016	0.002
Large−Small	−0.055	−0.072	−0.122	−0.086
	(−14.27)	(−14.78)		

Table 8 presents the median ratio of CEO compensation to gross capital for size-sorted portfolios in the data and in the model. Size in the data is measured by either the number of firm employees or gross capital. T-statistics for the difference between large and small firms based on the Newey and West [48] estimator with four lags are reported in parentheses.

Table 9

Elasticity of CEO Compensation to Firm Size

	Data			Model	Model
	Market Cap	Employees	Capital	Specification1	Specification2
Elasticity	0.33	0.31	0.28	0.25	0.05
	[0.019]	[0.022]	[0.023]		
R^2	0.27	0.20	0.20	0.28	0.14

Table 9 shows the elasticity of CEO compensation to firm size. Elasticities are estimated in a panel regression of CEO pay on lagged firm size (both measured in logs), controlling for firm and time fixed effects. Size in the data is measured by either market capitalization, the number of firm employees or gross capital. In "Data" panel, we report the estimated elasticities, standard errors clustered by firm and time (in brackets), and regression R^2's. The model statistics represent population numbers that are computed using a large panel of simulated data.

Table 10

Elasticity of CEO Compensation to Size Conditional on Firm Size

	Data						Model	Model
	Market Cap		Employees		Capital		Specification1	Specification2
Small	0.29	[0.04]	0.41	[0.04]	0.34	[0.04]	0.42	0.08
2	0.25	[0.05]	0.12	[0.09]	0.16	[0.10]	0.02	0.00
3	0.17	[0.08]	0.23	[0.12]	0.19	[0.11]	0.00	0.00
4	0.27	[0.05]	0.04	[0.11]	-0.05	[0.08]	0.03	0.00
Large	0.38	[0.04]	0.19	[0.06]	0.26	[0.06]	0.63	0.17

Table 10 shows variation in size elasticities of CEO compensation across size-sorted portfolios. Elasticities are estimated in a panel regression of CEO pay on lagged firm size (both measured in logs), controlling for firm and time fixed effects. Size in the data is measured by either market capitalization, the number of firm employees or gross capital. In "Data" panel, we report the estimated elasticities and clustered standard errors (in brackets). The model statistics represent population numbers that are computed using a large panel of simulated data.

Table 11

Dynamics of CEO Compensation

	Panel A: $Y = \Delta c_t$			**Panel B**: $Y = c_t$		
Regressors	**Data**		**Model**	**Data**		**Model**
Δk_{t-1}	0.119	(7.2)	0.049			
k_{t-1}^{Max}				0.312	(8.6)	0.048
c_{t-1}				0.173	(9.1)	0.928
$\Delta^{Min-} k_t \cdot I_t^{Small}$	1.020	(3.3)	0.530	0.975	(2.7)	0.485
$\Delta^{Min+} k_t \cdot I_t^{Small}$	0.115	(1.8)	-0.028	-0.160	(-1.5)	-0.015
$\Delta^{Max-} k_t \cdot I_t^{Large}$	0.183	(4.3)	-0.015	0.070	(1.8)	0.000
$\Delta^{Max+} k_t \cdot I_t^{Large}$	0.776	(2.0)	0.755	1.558	(5.3)	0.647

Table 11 presents the dynamics of CEO compensation in the data and in the model. Panel A shows the estimates from a panel regression of the log-growth of CEO compensation ($\Delta c_t \equiv \log \frac{C_t}{C_{t-1}}$) on the log-growth of firm size ($\Delta k_{t-1} \equiv \log \frac{K_{t-1}}{K_{t-2}}$), and four interaction terms that correspond to opposite changes in size realized by small and large firms.[†] For each firm, we compute the change in size at time t relative to its running minimum and maximum: $\Delta^{Min} k_t \equiv \log \frac{K_t}{K_{t-1}^{Min}}$ and $\Delta^{Max} k_t \equiv \log \frac{K_t}{K_{t-1}^{Max}}$, where K_{t-1}^{Min} and K_{t-1}^{Max} are respectively minimum and maximum of firm size observed in the five years prior to year t. We focus separately on negative and positive changes in firm size: $\Delta^{Min-} k_t$ and $\Delta^{Max-} k_t$ represent declines, and $\Delta^{Min+} k_t$ and $\Delta^{Max+} k_t$ represent increases in firms size. I_t^{Small} and I_t^{Large} are size dummies that select firms in the bottom and the top decile of size distribution, respectively. In Panel B, we consider a specification in levels, where we regress the log-level of CEO compensation (c_t) on its lag (c_{t-1}), the log of the running maximum of firm size (k_{t-1}^{Max}) and the four interaction terms. In the data, we run a panel regression with firm and time fixed effects and report the estimates and the corresponding t-statistics (in parentheses). Size in the data is measured by market capitalization. The model statistics represent population numbers that are computed using a large panel of simulated data under Model Specification 1.

[†] To simplify the notation, we omit the firm subscript that should appear on all variables.