A Unified Model of Firm Dynamics with Limited Commitment and Assortative Matching*

Hengjie Ai, Dana Kiku, Rui Li, and Jincheng Tong †

June 10, 2018

We develop a unified theory of dynamic contracting and assortative matching to explain firm dynamics. In our model, the joint distribution of firm size, investment and managerial compensation is determined by the optimal contract where neither firms nor managers can commit to arrangements that yield lower payoffs than their outside options. The outside options of both parties are micro-founded by the equilibrium conditions in a matching market. Our model endogenously generates power laws in firm size and CEO compensation and explains differences in their right tails. We also show that our model quantitatively accounts for many salient features of the time-series dynamics and the cross-sectional distribution of firm growth, dividend payout and CEO compensation.

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) and Jincheng Tong (tongx160@umn.edu) are affiliated with the Carlson School of Management, University of Minnesota, Dana Kiku (dka@illinois.edu) is at the University of Illinois at Urbana-Champaign, 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.

Disclosure Statement

Dear editors,

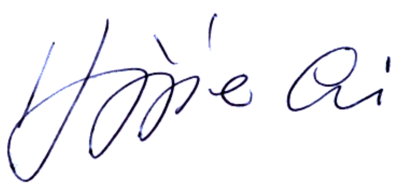
I have nothing to disclose.

Sincerely,

Hengjie Ai

Associate Professor of Finance

University of Minnesota



Disclosure Statement

Dear editors,

I have nothing to disclose.

Sincerely,

Dana Kiku

Assistant Professor

University of Illinois at Urbana-Champaign

Disclosure Statement

Dear editors,

I have nothing to disclose.

Sincerely,

Rui Li

Assistant Professor of Finance

University of Massachusetts Boston

1 Introduction

In this paper, we provide a unified theory of limited commitment and assortative matching that accounts for many stylized features of the time-series dynamics and the cross-sectional distribution of firm investment, dividend payout and managerial compensation. The contracting literature on limited commitment has developed a compelling theoretical framework of the dynamics of firm growth and managerial compensation (Albuquerque and Hopenhayn [2004], Harris and Holmstrom [1982]). Yet, the quantitative relevance of limited commitment remains unresolved because the outside options of firms and managers, which are specified exogenously in existing models, are hard to measure empirically. To impose discipline and provide a micro-foundation for the outside options we integrate an assortative matching framework into the theory of dynamic contracting with limited commitment. In our model, the outside options of firms and managers are jointly determined by the equilibrium conditions in a matching market, and the time series dynamics and the cross-section distribution of firm size, investment, and CEO pay are determined by the optimal contract.

To make the case for limited commitment, we first consider a continuous-time version of the organization capital model of Atkeson and Kehoe [2005]. Because technology shocks are i.i.d., without agency frictions, Gibrat [1931]'s law holds and the distribution of firm size follows a power law. However, contrary to the data, this model 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 of CEO compensation in the data. Motivated by the limitations of a frictionless model, we introduce two-sided limited commitment by assuming that neither firms nor managers can commit to contracts that result in lower payoffs than their outside options.

To provide a micro-foundation for firms' and managers' outside options, we introduce a market where managers and firms meet to form new productive relationships. Organization capital of a firm is jointly determined by managerial human capital and firm productivity. As in Gabaix and Landier [2008], the efficient outcome requires assortative matching. Once a productive match is formed, firms' investment, CEO compensation and dividend payout policies are determined by optimal contracting subject to limited commitment. Firms and managers may choose to voluntarily terminate their relationships and search for new ones, hence, their outside options are endogenously determined by the active matching market. Because the equilibrium matching rule is assortative, the outside options of firms and managers are increasing functions of their productivity and ability, respectively.

The optimal compensation contract in our model is determined by risk-sharing considerations and the equilibrium dynamics of the outside options. As in Harris and Holmstrom [1982], risk sharing implies that CEO compensation must be constant whenever limited commitment constraints are not binding. A sequence of positive shocks to human capital raises CEO's outside options and makes a constant compensation contract unsustainable because, as is, it eventually makes the manager voluntarily leave the firm. Because separation is associated with a (partial) loss of organization capital, the optimal compensation contract implements a minimal increase in CEO compensation necessary to prevent separation whenever the manager's limited commitment constraint binds. Similarly, a sequence of negative productivity shocks lowers the value of a firm and as the firm approaches its outside option, managerial compensation has to be reduced to curtail shareholders' incentive to abandon the firm because default is inefficient.

We 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 return to scale of the matching market technology similar to the implication of the Gabaix and Landier [2008] model.

Second, we show that two-sided limited commitment 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 firm-side 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.

Third, our model makes a set of unique time-series and cross-sectoral predictions. In particular, our model implies an asymmetric response of CEO compensation to productivity shocks — in small firms that are close to the firm-side limited commitment constraint, CEO compensation is sensitive to negative shocks, whereas in large firms that are close to the manager commitment constraint, managerial compensation is sensitive to positive shocks. Further, the degree of distortions due to agency frictions varies with the return to scale of the matching market technology. In particular, industries with a less decreasing returns to scale of the matching technology, feature high managerial compensation relative to firm size, a strong inverse relationship between firm investment and size, and a relatively fat tail of the

CEO-pay distribution.

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. Using a cross-section of industry portfolios, we also show that consistent with the model's predictions, industries that feature a less decreasing returns to scale of the matching technology feature a significantly higher ratio of CEO compensation to firm size and a significantly larger cross-sectional variation in investment rates compared with industries that have a significantly decreasing returns to scale in the matching market.

The tradeoff between risk sharing and limited commitment in our model builds on the earlier work of Kehoe and Levine [1993], Kocherlakota [1996], and Alvarez and Jermann [2000]. Albuquerque and Hopenhayn [2004] develop a model of firm dynamics based on limited commitment. Berk, Stanton, and Zechner [2010] solve for the optimal labor contract in a model with limited commitment and capital structure decisions. Eisfeldt and Papanikolaou [2013, 2014] emphasize that compensation of the key firm employees depends on their outside options. Rampini and Viswanathan [2010, 2013] study the implications of limited commitment for risk management and capital structure.[1] Cooley, Marimon, and Quadrini [2013] develop a model with two-sided limited commitment to study the increase in the size of the financial sector and in the compensation of financial executives. Lustig, Syverson, and Van Nieuwerburgh [2011] consider a model with limited commitment on the manager side and study the link between the inequality of CEO compensation and productivity growth. The optimal contracting aspect of our model is closely related to the continuous-time model of Ai and Li [2015], and Bolton, Wang, and Yang [2014]. None of the above papers explicitly model a matching market and relate the power law of CEO

[1] A broader literature that focuses on the implications of dynamic agency problems for firms' investment and financing decisions includes Quadrini [2004], Clementi and Hopenhayn [2006], and DeMarzo and Fishman [2007]. Limited commitment is also featured in Lorenzoni and Walentin [2007], Schmid [2008], Arellano, Bai, and Zhang [2012], and Li [2013].

compensation to the return to scale parameter of the matching technology.

A second strand of literature that our paper builds on is models of power law and models of assortative matching. Gabaix [2009] and Luttmer [2010] provide excellent surveys of the literature on power law and firm dynamics. The neoclassical model without frictions considered in our paper is essentially an interpretation of the model in Luttmer [2007]. Terviö [2008], and Gabaix and Landier [2008] are assortative matching models that link CEO compensation to firm size taking size distribution as given.

Our paper is also related to the large literature on agency frictions and managerial compensation. Edmans and Gabaix [2016] provide a comprehensive review of the earlier literature; more recent papers include Edmans, Gabaix, Sadzik, and Sannikov [2012], Biais, Mariotti, and Villeneuve [2010], and Bond and Axelson [2015].

The rest of the paper is organized as follows. We describe the setup of our model with limited commitment and assortative matching in Section 2. In Section 3, we consider a frictionless economy and discuss its implications and limitations. We characterize the optimal dynamic contract under limited commitment and its implications for firm investment and CEO compensation in Section 4. We calibrate our model and evaluate its quantitative implications in Section 5. Concluding remarks are provided in Section 6.

2 Model setup

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

2.1 Production Technology

Time is continuous and infinite. There is a continuum of firms indexed by j. As in Atkeson and Kehoe [2005], the output of firm j, denoted y_j, is produced from organization capital (Z_j), physical capital (K_j), and labor (N_j) using a standard Cobb-Douglas production technology: $y_j = Z_j^{1-\nu}\left(K_j^{\alpha}N_j^{1-\alpha}\right)^{\nu}$, where $\nu \in (0,1)$ is the span-of-control parameter.[2] The operating profit of firm j is defined as

$$\pi\left(Z_j\right) = \max\left\{Z_j^{1-\nu}\left(K_j^{\alpha}N_j^{1-\alpha}\right)^{\nu} - MPK \cdot K_j - MPL \cdot N_j\right\}, \tag{1}$$

[2]Organization capital is firm-specific knowledge that makes physical capital and labor more productive. Examples of organization capital include corporate culture, team work, firm-specific human capital, etc. (see Prescott and Visscher [1980]).

where MPK is the rental rate of physical capital and MPL is the equilibrium wage of unskilled workers. We assume that the unskilled labor, N, and physical capital, K, are not firm-specific and can be hired or rented in competitive markets. As a result, the constant returns to scale (CRS) of the production function implies that the profit function is linear in organization capital, i.e., $\pi(Z_j) = AZ_j$, where A is the equilibrium marginal product of organization capital. Due to the CRS production technology, the optimal choice of K_j and N_j is proportional to Z_j. Hence, in our model, all three are equivalent measures of firm size. For simplicity, we normalize the total supply of K and N in the economy to be one, although our model can be easily extended to allow for elastic supply of both factors.

The organization capital, Z_j, is specific to a matched firm-manager pair, and we assume that the accumulation of organization capital depends on manager's investment decisions:

$$dZ_{j,t} = Z_{j,t}\left[\left(i_{j,t} - \delta\right)dt + \sigma dB_{j,t}\right], \tag{2}$$

where $i_{j,t} = \frac{I_{j,t}}{Z_{j,t}}$ is the investment-to-organization capital ratio, δ is the depreciation rate of organization capital, $dB_{j,t}$ represents productivity shocks to organization capital, and σ is the sensitivity of Z with respect to Brownian motion shocks. As in Atkeson and Kehoe [2005], and Luttmer [2012], firm dynamics are driven by the accumulation of organization capital. The cost of investment in organization capital is specified by a standard quadratic adjustment cost: $h\left(\frac{I}{Z}\right)Z$, where $h(i) = i + \frac{1}{2}h_0 \cdot i^2$ with $h_0 > 0$.

2.2 Entry, Exit, and Matching Technology

Operating a firm requires a manager, who is the only type of agents in the economy that can efficiently build up organization capital for firms. A measure $\bar{e}$ of firms and a measure $\bar{e}$ of managers arrive in the economy per unit of time. The initial level of human capital of managers and the initial level of firm-specific organization capital are assumed to be the same and are denoted by $\bar{Z}$. Newly arrived firms and newly arrived managers meet immediately in a directed matching market to form productive firms, which we discuss shortly.

Firms and managers exit the economy if they receive an exogenous death shock that arrives at Poisson rate κ_D. For simplicity, we assume that within a match, the death shock of the firm and the death shock of the manager are perfectly correlated and we denote the stopping time associated with the firm's exit by τ_D. Once hit by the death shock, the organization capital of the firm and the human capital of the manager evaporate.

The operating firm-manager pair may separate for exogenous reasons, which occur at Poisson rate κ_S. Upon separation, the firm retains fraction $\lambda \in (0, 1)$ of organization capital as firm-specific organization capital and becomes an idle firm. The manager retains fraction

λ of organization capital as manager-specific human capital and becomes jobless. Formally, let τ_S denote the stopping time of separation, then:

$$Y_{\tau_S} = \lambda Z_{\tau_S}; \quad X_{\tau_S} = \lambda Z_{\tau_S},$$

where Y is the firm-specific organization capital of an idle firm and X is the human capital of an unemployed manager.

We assume that idle managers can produce home consumption goods according to the technology in Equation (2) but are not allowed access to the credit market.[3] Home production cannot be used to invest in human capital and therefore the law of motion of the human capital of an unemployed manager follows:

$$dX_{j,t} = X_{j,t} \left[-\delta dt + \sigma dB_{j,t} \right], \tag{3}$$

Idle firms neither produce cash flow nor invest in organization capital and as a result, their organization capital follows the same law of motion, i.e.,

$$dY_{j,t} = Y_{j,t} \left[-\delta dt + \sigma dB_{j,t} \right], \tag{4}$$

where j is interpreted as the index of the firm.

At Poisson rate κ_M, idle firms and managers (both newly arrived and extant) have an opportunity to meet in a directed matching market where idle firms offer competitive contracts to managers to form productive firms. We denote the stopping time at which an idle firm or an idle manager have an opportunity go to the directed matching market as τ_M. Once the contract is accepted, firms and managers pair together and become productive firms. We assume that the initial level of organization capital of a newly formed firm-manager pair depends on the firm-specific organization capital, Y, and the manager-specific human capital, X, through a Cobb-Douglas function:

$$Z_\tau = Y_\tau^{\psi_Y} X_\tau^{\psi_X}, \quad \text{with } \psi_Y, \psi_X \in (0,1) \text{ and } \psi \equiv \psi_Y + \psi_X \leq 1, \tag{5}$$

where τ denotes the stopping time at which a match is formed and the new firm starts operation.[4] Once a match is formed, Z_t follows the dynamics described in Equation (2). That is, the initial level of the match-specific organization capital, $Z_{j,t}$, is determined by the

[3] The assumption that idle managers can produce home consumption is inconsequential. It is made to ensure that consumption remains positive and hence utility is well defined under constant relative risk aversion (CRRA) preferences.

[4] In equilibrium, for idle managers and idle firms, $\tau = \tau_M$. That is, contracts are immediately accepted once matches are formed.

manager-specific human capital and the firm-specific organization capital upon the match, and is subject to match-specific shocks (see Equation (2)) until separation or exit. The parameter ψ can be interpreted as the return to scale of the matching technology. Decreasing return to scale, i.e., $\psi < 1$ captures the idea that it is harder for organization capital to make an impact in large firms than in small firms.

2.3 Preferences

A contract offered to a manager at time t, denoted $\left\{\{C_s, i_s\}_{s=t}^{\tau_D \wedge \tau_S}\right\}$, specifies compensation of the manager and investment in organization capital as functions of the history of the realization of the shocks. Here, we use the notation $\tau_D \wedge \tau_S \equiv \min\{\tau_D, \tau_S\}$, because the contractual relationship between a firm and a manager is terminated once the firm is hit by the death shock or the firm-manager pair separates. Given the contract, a manager evaluates his utility under the contract using standard CRRA preferences with a discount rate of r. Let $\kappa = \kappa_D + \kappa_S$ denote the rate of termination of the contractual relationship and define the continuation utility of a manager at time t, U_t, as:

$$U_t = \left\{ E_t \left[\int_t^{\tau_D \wedge \tau_S} (r+\kappa)\, e^{-r(s-t)} C_s^{1-\gamma} ds + \mathbb{1}_{\{\tau_S < \tau_D\}} e^{-r(\tau_S - t)} U^S \left(\lambda Z_{\tau_S}\right)^{1-\gamma} \right] \right\}^{\frac{1}{1-\gamma}}. \tag{6}$$

Here, $U^S(X)$ denotes the utility of an unemployed manager with human capital X. The term $U^S(\lambda Z_{\tau_S})$ reflects the fact that upon separation, the manager keeps λ fraction of the match-specific organization capital as his human capital, and $\mathbb{1}_{\{\tau_S<\tau_D\}}$ is the indicator function that takes a value of 1 in the event of $\tau_S < \tau_D$.

Recall that at the stopping time τ_M, an unemployed manager receives an opportunity to meet with a firm in the centralized directed matching market. Let $\bar{U}(X)$ denote the maximum utility a manager with human capital X can achieve in the directed matching market. Then, $U^S(X)$ must satisfy:

$$U^S(X) = \left\{ E_t \left[\int_t^{\tau_D \wedge \tau_M} (r+\kappa)\, e^{-r(s-t)} \left(AX_s\right)^{1-\gamma} ds + \mathbb{1}_{\{\tau_M < \tau_D\}} e^{-r(\tau_M - t)} \bar{U}\left(X_{\tau_M}\right)^{1-\gamma} \middle| X_t = X \right] \right\}^{\frac{1}{1-\gamma}}. \tag{7}$$

That is, an unemployed manager consumes home production of AX_s before receiving a matching opportunity at τ_M. At the stopping time τ_M, the manager pairs with an idle firm in the directed matching market to obtain utility $\bar{U}(X_{\tau_M})$.

Our formulations in Equations (6) and (7) normalize managers' utility to be homogeneous of degree one in consumption for tractability. As is standard in the dynamic contracting literature (for example, Thomas and Worrall [1988]), we use promised utility as a state variable and index contracts with the initial promised utility, U:

$$\left\{E_0\left[\int_0^{\tau_D\wedge\tau_S}(r+\kappa)\,e^{-rs}C_s^{1-\gamma}ds+\mathbb{1}_{\{\tau_S<\tau_D\}}e^{-r\tau_S}U^S\left(\lambda Z_{\tau_S}\right)^{1-\gamma}\right]\right\}^{\frac{1}{1-\gamma}}\geq U \tag{8}$$

We assume that shareholders of firms are well diversified and therefore risk-neutral with respect to idiosyncratic shocks. Given a contract, $\left\{\{C_s,i_s\}_{s=t}^{\tau_D\wedge\tau_S}\right\}$, the firm evaluates the present value of the cash flow at the same discount rate r:

$$E_t\left[\int_t^{\tau_D\wedge\tau_S}e^{-r(s-t)}\left[AZ_s-C_s-h\left(\frac{I_s}{Z_s}\right)Z_s\right]ds+\mathbb{1}_{\{\tau_S<\tau_D\}}e^{-r(\tau_S-t)}V^S\left(\lambda Z_{\tau_S}\right)\right], \tag{9}$$

where $V^S(Y)$ denotes the value of an idle firm with organization capital Y. The term $V^S\left(\lambda Z_{\tau_S}\right)$ reflects the fact that upon separation, the firm retains λ fraction of the match-specific organization as firm-specific organization capital. Because the firm does not produce any cash flow until paired with another manager, $V^S(Y)$ is determined by

$$V^S(Y)=E_t\left[\mathbb{1}_{\{\tau_M<\tau_D\}}e^{-r(\tau_M-t)}\bar{V}\left(Y_{\tau_M}\right)\middle|\,Y_t=Y\right], \tag{10}$$

where $\bar{V}(Y)$ is the maximum value that the firm can achieve on the directed matching market.

In our environment, the stock of organization capital and the promised utility constitute a pair of state variables that are sufficient to summarize any information that is relevant for the design of the optimal contract. We denote the value of a firm with organization capital Z and promised utility to its manager U as $V(Z,U)$.

2.4 Limited Commitment

We assume that both firms and managers can choose to unilaterally separate at any time and cannot commit not to do so ex ante. Upon separation, firms and managers become idle until they receive an opportunity to become part of a new productive firm in the directed matching market. Incentive compatibility requires that before separation, $t<\tau_S\wedge\tau_D$, the continuation utility of managers must be higher than what they can obtain upon separation:

$$U_t\geq U^S\left(\lambda Z_t\right). \tag{11}$$

Similarly, the continuation value of a firm must be higher than its outside option. That is, for all $t<\tau_S\wedge\tau_D$,

$$V\left(Z_t,U_t\right)\geq V^S\left(\lambda Z_{\tau_S}\right). \tag{12}$$

Because both firm value and manager utility are strictly increasing, if one of the inequalities in Equations (11) and (12) is strict, the firm's cash flow can always be reallocated to make both parties better off to prevent inefficient separation. Therefore, under the optimal contract, both Equations (11) and (12) must hold with equality at the time of endogenous separation.

In our model, endogenous separation may or may not happen in equilibrium depending on the specification of the optimal contract. However, matched firms and managers of all types exogenously separate with positive probability, and idle firms and managers continuously meet in the directed matching market. Therefore, competition in the matching market determines outside options of firms and managers of all types.

2.5 Matching Decisions

In the directed matching market, firms of all types offer competitive contracts to managers to form productive firms. We construct equilibriums in which the matching rule is one-to-one and onto, that is, any manager of type X is matched to a single type of firm $Y = Y(X)$, and any firm of type Y is matched to a single type of manager, which we denote $X(Y)$. Clearly, $Y(\cdot)$ and $X(\cdot)$ are inverse functions of each other. We will use $Y(X)$ and $X(Y)$ interchangeably to represent the equilibrium matching rule.

Consider an idle firm with organization capital Y. Because the initial organization capital of firm is given by Equation (5), if matched with a manager with human capital X, the value of the firm is $V\left(Y^{\psi_Y} X^{\psi_X}, U\right)$, where U is the initial promised utility to the manager. Because $\bar{U}(X)$ is the minimum utility that the firm must provide to managers with human capital X, the equilibrium matching rule must satisfy the optimality condition for firms:

$$X(Y) \in \arg\max_X V\left(Y^{\psi_Y} X^{\psi_X}, \bar{U}(X)\right). \tag{13}$$

That is, firms must choose the type of the manager optimally to maximize firm value. In equilibrium, firms' outside options are determined by the maximum value they can obtain in the directed matching market:

$$\bar{V}(Y) = V\left((Y)^{\psi_Y} X(Y)^{\psi_X}, \bar{U}(X(Y))\right). \tag{14}$$

2.6 Recursive Stationary Equilibrium

To build up the concept of recursive stationary equilibrium similar to the construction of Atkeson and Lucas [1992], we first describe a recursive procedure to specify the optimal contract within a match. Consider a productive firm initiated at time τ by matching a manager with human capital X_τ and an idle firm with organization capital Y_τ. The initial organization capital of the firm is $Z_\tau = Y_\tau^{\psi_Y} X_\tau^{\psi_X}$ and the initial utility promised to the manager is $U_\tau = \bar{U}(X_\tau)$.

Without loss of generality, the optimal compensation and investment policy can be specified by a two-step procedure. First, we specify compensation and investment, $C(Z,U)$ and $I(Z,U)$, as functions of the state variables (Z,U). Second, we determine the law of motion of Z_t by Equation (2) and the law of motion of U_t by specifying its sensitivity with respect to the Brownian motion shocks, $G(Z,U)$. Given $G(Z,U)$, Equation (6) implies that prior to the death shock and the separation shock, the law of motion of U is given by:[5]

$$\begin{aligned} dU &= \left[\frac{\beta+\kappa}{1-\gamma}\left(U - C^{1-\gamma}U^{\gamma}\right) + \frac{1}{2}\gamma\frac{G(Z,U)^2\sigma^2}{U} + \frac{\kappa_S}{1-\gamma}U\left(1-\left(\frac{U^S(\lambda U)}{U}\right)^{1-\gamma}\right)\right]dt \\ &\quad + G(Z,U)\,\sigma dB. \end{aligned} \tag{15}$$

Formally, an equilibrium in our model consists of the following quantities: an optimal recursive contract $\{C(Z,U), I(Z,U), G(Z,U)\}$, and an equilibrium matching rule, $Y(X)$, that satisfy the following conditions:

1. Given the equilibrium matching rule $Y(X)$, the initial condition of a firm is set by: $Z_\tau = [Y(X_\tau)]^{\psi_Y} X_\tau^{\psi_X}$.

2. The initial condition of a contract offered at time τ is determined by Z_τ and $U_\tau = \bar{U}(X_\tau)$. Given the equilibrium outside options of managers and firms, $\bar{U}(X)$ and $\bar{V}(Y)$, respectively, the optimal contract offered at time τ maximizes firm value in Equation (9) subject to the participation constraint, Equation (8), and the limited commitment constraints in Equations (11) and (12), with the understanding that $t = \tau$ in Equations (8) and (9).

3. Given the equilibrium outside option of managers, $\bar{U}(X)$, the equilibrium matching rule satisfies firms' optimality in Equation (13).

4. The outside options of firms are determined by the optimal matching rule in Equation (14).

[5]This formulation is similar to the representation in Sannikov [2008], except that we use a monotonic transformation so that utility is measured in consumption units. We provide the details of the derivation in Appendix D.

5. Together with the equilibrium matching rule, managers' equilibrium outside option, $\bar{U}(X)$, clears the matching market. That is, all firms and managers match instantaneously upon birth or separation. In addition, the prices of capital and labor (MPK and MPL) clear the markets for physical capital and labor.

2.7 Assortative Matching Rule

We conjecture and later verify that the equilibrium features assortative matching. That is, firms and managers are matched according to their ranking in their respective population. Intuitively, more productive firms are matched with managers with higher human capital. Because the laws of motion of firm organization capital and manager human capital are identical and their initial conditions are both normalized to $\bar{Z}$, symmetry and market clearing imply a simple assortative matching rule: $Y(X) = X$ and $X(Y) = Y$.

Under the above matching rule, the optimality condition can be written as:

$$V_Z\left(X^{\psi_Y+\psi_X}, \bar{U}(X)\right)\psi_X X^{\psi_Y+\psi_X-1} + V_U\left(X^{\psi_Y+\psi_X}, \bar{U}(X)\right)\bar{U}'(X) = 0. \tag{16}$$

Given the functional form of $V(Z, U)$, the above equation determines the slope of managers' outside options, $\bar{U}'(X)$.

3 The First-Best Case

As a starting point of our analysis, we first consider the first-best case in which no separation occurs, i.e., $\kappa_S = 0$, and firms and managers fully commit. Here, shareholders maximize the present value of firm's cash flow subject to the manager's participation constraint in Equation (8). The present value of cash flow can be written as:

$$E_t\left[\int_t^{\tau_D} e^{-r(s-t)}\left(AZ_s - h\Big(\frac{I_s}{Z_s}\Big)Z_s\right)ds\right] - E_t\left[\int_t^{\tau_D} e^{-r(s-t)}C_s ds\right]. \tag{17}$$

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.[6]

The firm-value maximization problem in the first step is standard as in Hayashi [1982]. 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: $C_s = U_t$ for all $s \geq t$. 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:*

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

The following proposition summarizes the solution to the firm's problem.

Proposition 1. *The First-Best Case*

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

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

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

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

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

Proof. See Appendix A. □

The first term $\bar{v}Z = h'(\hat{\imath})Z$ in Equation (19) is the firm value in the neoclassical model with constant returns to scale production function and quadratic 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 present value of which is simply given by the Gordon [1959]'s formula: $\frac{1}{r+\kappa}U$.[7] Finally, the equilibrium marginal product of organization capital, A, and the total amount of organization capital, $\mathbf{Z}$, are jointly determined by $\mathbf{Z} = \frac{\bar{Z}}{\kappa+\delta-\hat{\imath}}$. We provide the details of this calculation in Appendix E.

[6]This procedure is not possible in the case with agency frictions because limited commitment imposes restrictions on the joint dynamics of C_t and K_t.

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

Equation (20) in Proposition 1 reveals that the investment-to-capital ratio in the first-best 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 [2007], which is summarized in the following proposition.

Proposition 2. *Power Law of Firm Size*

Given firms' initial size, $Z_0 = \bar{Z}$, and their optimal investment policy, $\hat{\imath}$, the density of the firm size distribution is given by:

$$\phi(Z) = \begin{cases} \frac{1}{\sqrt{\left(\hat{\imath}-\delta-\frac{1}{2}\sigma^2\right)^2+2\kappa\sigma^2}} Z^{\xi-1} & Z \geq \bar{Z} \\ \frac{1}{\sqrt{\left(\hat{\imath}-\delta-\frac{1}{2}\sigma^2\right)^2+2\kappa\sigma^2}} Z^{\eta-1} & Z < \bar{Z}, \end{cases} \tag{21}$$

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. □

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 fails to account for the observed fat tail in CEO pay, which we document below.

4 Limited Commitment

4.1 Optimal Contracting with CRS Matching Technology

In this section, we discuss our benchmark model with two-sided limited commitment and assortative matching. We focus on the case of constant return to scale matching technology, i.e., we set $\psi = 1$, because it significantly simplifies the optimal contracting analysis. We will discuss the general case of DRS matching technology in subsequent sections. Under the assumption of $\psi = 1$, the outside option of the manager is given by $\bar{U}(Z) = \bar{u}Z$, for some constant $\bar{u}$. The value function and policy functions in this case are homogeneous:

$$V(Z,U) = v\left(\frac{U}{Z}\right)Z, \quad C(Z,U) = c\left(\frac{U}{Z}\right)Z, \quad I(Z,U) = i\left(\frac{U}{Z}\right)Z, \quad G(Z,U) = g\left(\frac{U}{Z}\right)Z,$$

for some normalized value function $v(\cdot)$, and policy functions $c(\cdot)$, $i(\cdot)$, and $g(\cdot)$. Let $u = \frac{U}{Z}$ denote the normalized utility. Given the homogeneity property, the optimal matching rule in Equation (16) simplifies to a restriction on $v(u)$ at $\bar{u}$:

$$\psi_X v(\bar{u}) + (1-\psi_X)\, v'(\bar{u})\,\bar{u} = 0. \tag{22}$$

4.1.1 Value Function and Dynamics of Continuation Utility

Due to the constant returns to scale assumption on the matching technology, the normalized continuation utility, u_t, is the single state variable that summarizes the history of shocks. Note that because U is the life-time utility promised to the manager and Z is firm size, the normalized utility $u = \frac{U}{Z}$ can be interpreted as managers' equity share in the firm. Note also that due to two-sided limited commitment, u must be bounded from above and below. First, managers may always choose to leave the firm and find a new match. Upon separation, a manager looses $(1-\lambda)$ of human capital; therefore, the maximum amount of utility the manager can obtain by voluntary separation is $U^S(\lambda Z)$. Equation (7) implies that $U^S(X) = u_{MIN}X$, where u_{MIN} is related to $\bar{u}$ by:

$$\left(r + \kappa_M + (1-\gamma)\delta + \frac{1}{2}(1-\gamma)\gamma\delta^2\right) u_{MIN}^{1-\gamma} = (r+\kappa_D)\, A^{1-\gamma} + \kappa_M \bar{u}^{1-\gamma} \tag{23}$$

The manager chooses to stay with the firm if and only if $u_t \geq u_{MIN}$.

Similarly, because firms have the option to voluntarily separate as well, $V(Z,U) \geq V^S(\lambda Z)$ should hold. Given Equation (10), the monotonicity of the value function $v(u)$ implies that $V^S(Y) = v(u_{MAX})\,Y$ for some u_{MAX} that satisfies

$$\kappa_M v(\bar{u}) = (r + \kappa_D + \kappa_M + \delta)\, v(u_{MAX})\,. \tag{24}$$

That is, limited commitment on the firm side requires $u_t \leq u_{MAX}$ for all t. The following proposition provides a characterization of the optimal contract and the equilibrium levels of $\bar{u}$, u_{MIN} and u_{MAX}.[8]

Proposition 3. *Optimal Contracting with Two-Sided Limited Commitment*

[8] A recent paper by Bolton, Wang, and Yang [2014] shows how similar contracts can be implemented by corporate liquidity and risk management policies.

1. *The normalized value function is decreasing and concave and satisfies the following ODE on* (u_{MIN}, u_{MAX})*:*

$$0 = \max_{c,i,g} \left\{ \begin{array}{c} A - c - h(i) + v(u)(i - \delta - r - \kappa) \\ +uv'(u) \left[\begin{array}{c} \frac{r+\kappa}{1-\gamma}\left(1 - \left(\frac{c}{u}\right)^{1-\gamma}\right) - (i-\delta) + \frac{1}{2}\gamma g^2 \sigma^2 \\ +\frac{\kappa_S}{1-\gamma}\left(1 - \left(\frac{\lambda u_{MIN}}{u}\right)^{1-\gamma}\right) \end{array} \right] \\ +\frac{1}{2}u^2 v''(u)(g-1)^2 \sigma^2 + \kappa_S \lambda v(u_{MAX}) \end{array} \right\} \tag{25}$$

with boundary conditions $\lim_{u \to u_{MIN}} v''(u) = \lim_{u \to u_{MAX}} v''(u) = -\infty$.

2. *Under the optimal contract,* $u_t \in [u_{MIN}, u_{MAX}]$ *and is decreasing in productivity shocks.*

3. *The initial normalized utility* $\bar{u}$ *satisfies Equation (22). Given* $\bar{u}$*,* u_{MIN} *is determined by Equation (23) and* u_{MAX} *is determined by Equation (24).*

4. *The optimal compensation ratio,* $c(u_t) = \frac{C_t}{Z_t}$*, takes the following form:*

$$\ln c(u_t) = \ln C_0 - \ln Z_t + l_t^+ - l_t^-, \tag{26}$$

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

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

Proof. See Appendix B. □

We plot the normalized value function for the first-best case (dashed line) and that for the two-sided limited commitment case (solid line) in Figure 1. First, note that agency frictions lower the Pareto frontier and as a result, $v(u)$ in our model is below the first-best value function in its entire domain. The normalized value function is monotonically decreasing and concave on $[u_{MIN}, u_{MAX}]$. It is decreasing because a higher promised utility to the managers implies that a smaller fraction of cash flow is retained by the firm and, therefore, a lower present value of the cash flow is delivered to firms' shareholders. The normalized value function is concave due to agency frictions and risk aversion of the manager.

Second, Equation (24) and the monotonicity of the value function $v(u)$ imply that the manager's outside option $\bar{u}$ is in the interior of (u_{MIN}, u_{MAX}), as shown in Figure 1. In addition, by the optimal matching condition, in Equation (22), at $\bar{u}$, $v(\bar{u}) = -\frac{(1-\psi_X)}{\psi_X} v'(\bar{u})\bar{u} > 0$. Hence, the value function is strictly positive in its domain.

Figure 1: Normalized Value Function

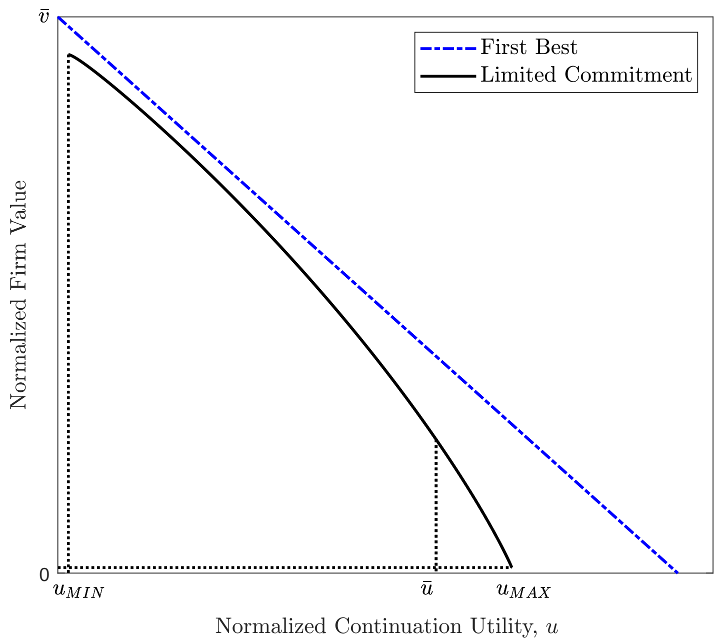


Figure 1 plots the normalized value function for the first best case (dashed line) and that for the case of limited commitment (solid line).

Third, a firm-manager match starts at $\bar{u}$ and follows a diffusion process implied by the optimal contract afterwards. Under the optimal contract, the normalized utility u_t travels within the bounded interval $[u_{MIN}, u_{MAX}]$ until the match dissolves.

In the model with CRS matching technology, normalized utility, u_t, is the only state variable that determines firms' investment and CEO compensation policies. Therefore, observable firm characteristics that are correlated with u_t, such as firm size and age, contain information about the cross-sectional variation in investment and CEO compensation. To understand the predictive power of firm size and age, consider the law of motion of u_t prior to separation:

$$d \ln u_t = \mu_u (u_t) \, dt + \sigma_u (u_t) \, dB_t, \tag{27}$$

where the expressions for $\mu_u (u)$ and $\sigma_u (u)$ are given in Appendix B.1. Intuitively, $\ln u_t$ is determined by two factors, the drift (dt term) and the diffusion (dB_t term). The drift determines the effect of age on u_t, and the diffusion term reflects the impact of size of u_t because firm size is mainly affected by the dB_t shocks.

We plot the drift term, $\mu_u (u)$, in the top panel of Figure 2 and the diffusion term, $\sigma_u (u)$, in the bottom panel. To understand how firm age is correlated with u_t, note that $\mu_u (u)$ is

Figure 2: Dynamics of the Normalized Continuation Utility

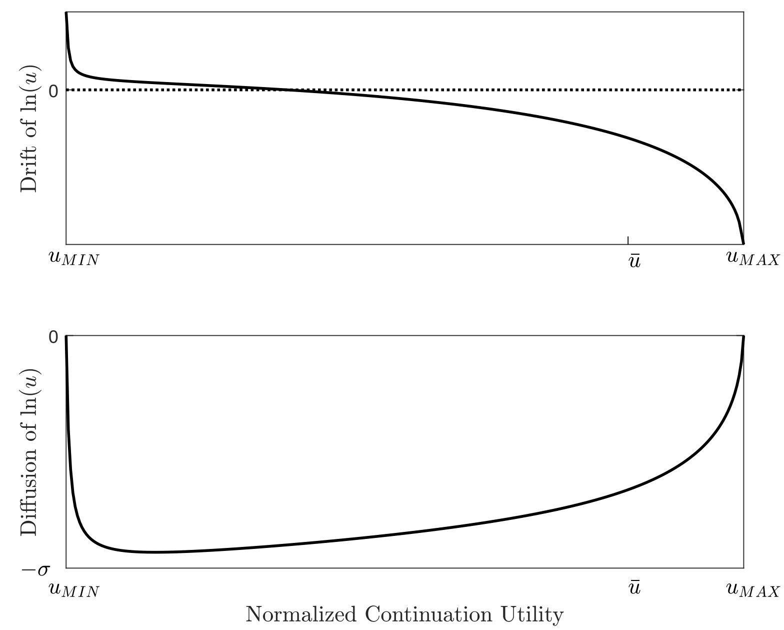


Figure 2 plots the drift of $\ln u$ (top panel) and the diffusion of $\ln u$ (bottom panel) under the optimal contract.

positive at u_{MIN}, monotonically decreases with u, and becomes negative at u_{MAX}. Thus, u_t is mean reverting — young firms enter at $\bar{u}$ and, as they grow older, they tend to converge towards the steady state where $\mu_u(u)$ crosses zero.

The diffusion coefficient of $\ln u$ is above $-\sigma$ and below zero. Note that $\ln u_t = \ln U_t - \ln Z_t$. Perfect risk sharing (first-best) implies $\sigma_u(u) = -\sigma$ because continuation utility remains constant as Z_t moves with productivity shocks. In contrast, no risk sharing corresponds to $\sigma_u(u) = 0$ because continuation utility moves one for one with productivity shocks. A diffusion coefficient between $-\sigma$ and 0 is the consequence of imperfect risk sharing — a positive productivity shock raises Z_t and U_t simultaneously, but continuation utility U_t is less sensitive to shocks. The U-shaped diffusion coefficient of u_t indicates that risk sharing is poor when limited-commitment constraints are binding at u_{MIN} and u_{MAX}, respectively.

The fact that $\sigma_u(u)$ always stays below zero implies a negative correlation between firm size and managers' equity share in the firm, u_t. Negative shocks to organization capital reduce firm size and push u_t closer to u_{MAX} because risk sharing requires managerial compensation to be less sensitive to productivity shocks than firms' cash flow. This feature of the optimal contract determines the dynamics of CEO compensation and implies a negative relationship between firm size and investment, which we discuss below.

4.1.2 Dynamics of CEO Compensation

Part 4 of Proposition 3 provides a characterization of managerial compensation under the optimal contract. Intuitively, optimal risk sharing requires that CEO compensation remains constant whenever none of the commitment constraints binds, increases by the minimum amount to keep the manager from defaulting when the manager-side limited commitment constraint binds, and decreases by a necessary minimum amount to prevent the firm from shutting down when the firm-side limited commitment constraint binds. Formally, the logarithm of the compensation-to-capital ratio, $\ln c(u_t)$ can be obtained from $\ln C_0 - \ln Z_t$ by imposing a two-sided regulator, $\left\{l_t^+, l_t^-\right\}_{t=0}^{\infty}$.

To illustrate the dynamics of the optimal CEO compensation contract, in Figure 3, we plot a sample path of a firm starting from a promised utility close to the manager-side limited commitment constraint represented by u_{MIN}.[9] The top panel is the realization of the log size of the firm, $\ln Z_t$. The second panel is the path of the normalized utility, u_t, and the third panel is the trajectory of the log managerial compensation, $\ln C_t$. At time 0, the firm starts from the interior of the normalized utility space, $u_{MIN} < \bar{u} < u_{MAX}$. A sequence of positive productivity shocks from time 0 to 1 increases organization capital of the firm (top panel). For $t < 1$, $u_t > u_{MIN}$ is in the interior (second panel) and the manager's compensation is constant (bottom panel). At time 1, the normalized continuation utility reaches the left boundary, u_{MIN}, and the manager-sided limited commitment constraint binds. Further realizations of positive productivity shocks from $t = 1$ to $t = 2$ translate directly into an increase in CEO compensation (bottom panel), but the normalized continuation utility (second panel) remains constant. At time $t = 2$, the firm starts to experience a sequence of negative productivity shocks. The stock of organization capital, Z_t, declines and the normalized utility $u_t = \frac{U_t}{Z_t}$ increases because risk sharing implies that the continuation utility U_t is less sensitive to shocks than Z_t. During the period $t \in (2, 4)$, u_t stays in the interior of $[u_{MIN}, u_{MAX}]$ and the manager's consumption stays constant. At time $t = 3$, 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 reaches its previous running maximum at $t = 4$, at which time, the manager-side limited commitment constraint starts to bind again, and manager's compensation has to increase (bottom panel).

As Figure 3 shows, in the region where the limited commitment constraint on the manager side binds, CEO compensation under the optimal contract behaves like a linear function of the running maximum of firm size. This is the key mechanism for the power law of CEO compensation: the running maximum of a power law process follows the same power law.

[9] The dynamics of managerial compensation close to the firm-side limited commitment constraint represented by u_{MAX} follows a similar pattern.

Figure 3: A Sample Path of CEO Compensation

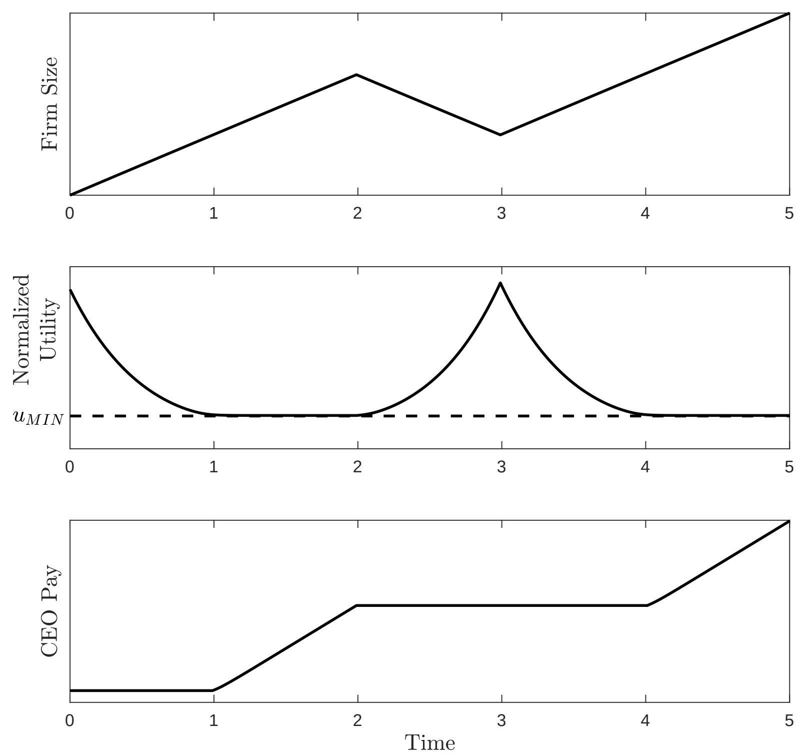


Figure 3 plots a sample path of log firm size (top panel), and the implied sample path for the normalized continuation utility (second panel) and that for log CEO pay (bottom panel) in the neighborhood of the manager-side limited commitment constraint, u_{MIN}.

We formalize and generalize this observation in Section 4.2 below.

4.1.3 Investment

As discussed in part 5 of Proposition 3 and displayed in Figure 4, the optimal investment, $i(u)$, increases with normalized continuation utility. Note that only firms in the interior invest at a rate similar to the investment rate in the first-best economy ($\hat{\imath}$). Small firms that have experienced a sequence of negative productivity shocks move towards the firm-side limited commitment constraint u_{MAX} where risk sharing is poor. To improve risk sharing and grow out of the constraint, small firms accelerate investment. Similarly, risk sharing deteriorates as firms grow large and approach the manager-side limited commitment constraint u_{MIN}. In this region, it is optimal for firms to reduce investment in organization capital to limit managers' outside options.

Figure 4: Investment Policy

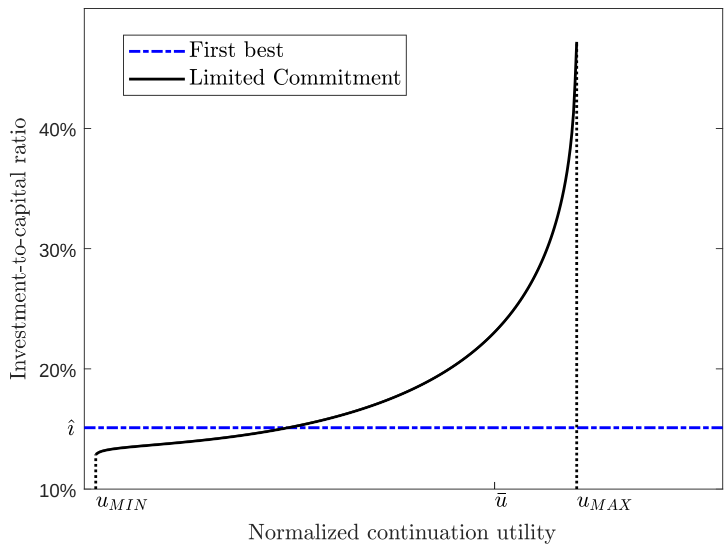


Figure 4 plots the optimal investment policy in the first best case (dash-dotted line) and that under limited commitment (solid line).

Under the optimal contract, the endogenous correlation between size, age, and normalized utilty translates into an endogenous correlation between size, age and investment. In our model, young firms start at $\bar{u}$, where investment rate is relatively high, and over time, slowly converge to the interior, where investment rates are low. In addition, size is negatively correlated with u and therefore with firm investment rates. In our model, both the firm-side and the manager-side limited commitment constraints are important in matching the negative relationship between investment and age, and investment and size in the data. However, as Figure 4 shows, quantitatively, the impact of limited commitment on investment is most significant as firms get close to u_{MAX}, where the firm-side limited commitment constraint binds. Note also that two-sided limited commitment implies that the endogenous state variable u_t travels between u_{MIN} and u_{MAX} under the optimal contract and therefore, the negative relationship between investment and size and that between investment and age persist in the long-run.

Finally, we note that a common device to generate the inverse relationship between investment and size is a decreasing returns to scale production function. However, models with decreasing returns to scale production function generally imply that firms converge to an optimal size in the long run and therefore are inconsistent with the empirical evidence of the fat-tailed size distribution.

4.2 Power Law of CEO Compensation

In this section, we provide an explicit characterization of the power law in CEO compensation implied by our model. Our key result is that the power law coefficient of CEO pay depends on the ratio between the power law slope of firm size, ξ, and the return to scale of the matching technology, ψ. In general, under decreasing returns to scale, $\psi < 1$, the distribution of CEO pay obtains a power law as long as the distribution of firm size does but with a thinner tail.

To illustrate the general relationship between the power law in CEO pay and firm size, we relax the constant returns to scale assumption of the matching technology. To derive a sharp theoretical result and illustrate its intuition, we make several simplifying assumptions, all of which will be relaxed in the quantitative exercise in Section 5. First, we assume that firms and managers experience the exogenous separation shock only once in their life-time. After separation, they meet in the matching market and sign a contract with full commitment. Because the optimal contract with full commitment has a simple solution, this assumption allows us to solve for the outside option of managers, $U^S(X)$, in closed form. For simplicity, we also assume that $\kappa_M = \infty$, that is, once separated, firms and managers immediately find an opportunity to match with each other.

Second, we assume that the adjustment cost function takes the following form:

Assumption 2. *The adjustment cost function satisfies*

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

where $\iota > 0$ is a parameter that determines the upper bound of investment and satisfies the following condition:

$$A > r + \kappa + \delta > \iota > \delta, \tag{28}$$

and

$$\frac{A-\iota}{r+\kappa+\delta-\iota} - \frac{\psi\gamma}{(r+\kappa)(\varsigma_1 - 1)}\varsigma_1^{\frac{\psi\gamma-1}{\psi(\gamma-1)}}(\varsigma_1 - (1-\gamma))^{\frac{-\psi\gamma}{\psi(\gamma-1)}}\bar{\bar{u}}^{\frac{1}{\psi}} \ge 1, \tag{29}$$

where

$$\varsigma_1 = \sqrt{\left(\frac{\iota-\delta}{\sigma^2} - \frac{1}{2}\right)^2 + \frac{2(\kappa+r)}{\sigma^2}} - \left(\frac{\iota-\delta}{\sigma^2} - \frac{1}{2}\right)$$

and $\bar{\bar{u}}$ is defined by Equation (C.2) *in Appendix C.1.*

That is, we assume that the marginal cost of investment is one if $\frac{I}{K} \le \iota$, and is infinite if $\frac{I}{K} > \iota$. Under Assumption 2, firms' optimal policy is to always invest at the maximum rate ι. Inequality (28) imposes a lower bound on the marginal product of capital, A, and an upper bound on the investment rate ι. This guarantees that investment is profitable and the present value of cash flow is finite. Condition (29) is a restriction on the magnitude of agency frictions. As we show in Appendix C.1, the outside option of the manager is given by $\bar{\bar{u}}X^{\psi}$. Inequality (29) implies that the manager's outside option is not too large, so that it is always optimal to invest at the maximum investment rate ι even when the limited commitment constraint is binding.

Finally, we assume that firms can fully commit. Because managerial compensation increases only when the manager-side limited commitment constraint binds, firm-side limited commitment does not affect the right tail of the distribution of CEO pay. Assuming full commitment on the firm side simplifies our analysis and allows us to focus on the key agency friction that determines the power law of CEO pay. We summarize our results below.

Proposition 4. *Power Law in CEO Compensation*

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

$$C_t = \max\left\{\hat{c}\max_{0\le s\le t} Z_s^{\psi},\ C_0\right\}, \tag{30}$$

where the constant $\hat{c}$ is defined in Equation (C.5) *in Appendix C.2. The optimal investment-to-capital ratio is constant: $I_t = \iota K_t \quad for\ all\ t$.*

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

Proof. See Appendix C. □

In the model with limited commitment on the manager side, the compensation contract is downward rigid, as in Harris and Holmstrom [1982]. Compensation has to increase to match the manager's outside option whenever the 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.[10]

Under our assumptions, 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 Z follows a power law with slope coefficient

[10] See also Lustig, Syverson, and Van Nieuwerburgh [2011], Grochulski and Zhang [2011], and Miao and Zhang [forthcoming].

ξ, the distribution of Z^{ψ} obeys a power law with slope coefficient $\frac{\xi}{\psi}$. By part 1 of Proposition 3, managerial compensation is a linear function of the running maximum of Z_t^{ψ}. 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 $\frac{\xi}{\psi}$. 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.

It is also 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 $C_t = \hat{c} \max_{0\leq s\leq t} Z_s^{\psi}$ and $D_t = AZ_t - I_t - C_t = AZ_t - \iota Z_t - \hat{c} \max_{0\leq s\leq t} Z_s^{\psi}$. Because $\psi \leq 1$, it follows that

$$(A - \iota)\, Z_t - \hat{c} \max_{0\leq s\leq t} Z_s \leq D_t \leq (A - \iota)\, Z_t.$$

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

4.3 Testable Implications

In this section, we briefly summarize testable implications for firms' investment, size, and CEO compensation that are unique to our model of limited commitment with assortative matching markets.

First, our model implies that the response of CEO compensation to firm size is history dependent. The pattern of history dependence is best illustrated in Equation (26) of Proposition 3 — CEO pay of small firms is sensitive to negative productivity shocks because of the firm-side limited commitment constraint. Similarly, CEO pay of large firms is sensitive to positive shocks because for them, the manager-side limited commitment constraint is likely to bind.

Second, our model ascribes the power law of CEO pay to the power law of firm size through the return to scale parameter of the matching technology, ψ. Decreasing return to scale, i.e. $\psi < 1$, implies that it is harder for organization capital to make an impact in large firms than in small firms. In industries where ψ is small, managers who have accumulated a large amount of human capital with their current employer incur substantial losses if they choose to separate and match with a new firm due to the decreasing returns to scale in the formation of new organization capital. Hence, managers in low-ψ industries have strong incentives to commit to risk sharing contracts. As a result, in industries with significantly decreasing returns to scale of the matching technology, the level of CEO pay relative to firm size is expected to be low, and the distribution of CEO pay has a thin tail.

Third, because the inverse relationship between investment and size in our model is due to the presence of agency frictions, and because the manager-side limited commitment constraint is more likely to bind in industries where ψ is close to one, the inverse relationship between investment and firm size should be stronger in high-ψ industries than in industries with significantly decreasing returns to scale of the matching technology. We now turn to the quantitative implications of our model and formally test them in the data.

5 Quantitative Results

As discussed in the previous section, qualitatively, our model is able to generate power laws in firm size and CEO compensation and a negative relationship between firm size and investment. In this section, we explore the quantitative implications of our model and its ability to account for the joint empirical distribution of firm size, investment, CEO compensation and dividend policies.

5.1 Data Description and Calibration of Parameters

To calibrate the model and evaluate its quantitative implications, we use a panel of US non-financial firms from the Center for Research in Securities Prices (CRSP) and Compustat. We measure executive compensation by the total compensation figure from ExecuComp database, which comprises salary, bonuses, the value of restricted stock granted, the Black-Scholes-based value of options granted and long-term incentive payouts. For each firm we collect market capitalization, the number of firm employees, the book value of firm assets, the gross value of property, plant and equipment to measure capital, capital expenditure to measure investment, and the amount of common dividends. We measure firm age by the number of years since the firm's founding date. Firm exit rates are computed using Compustat deletion series that account for acquisitions and mergers, bankruptcy, liquidation, reverse acquisition and leverage buyout. All nominal quantities are converted to real using the consumer price index compiled by the Bureau of Labor Statistics. The data are sampled on the annual frequency and cover the period from 1992 till 2016.

Tables 1 and 2 summarize the calibrated parameter values and a set of moments targeted in calibration. The parameters of our model can be divided into two groups. The first group of parameters is fairly standard and can be calibrated by following the existing literature.

We choose risk aversion of 2. We set the discount rate r to be 4% per year to match the average return of risky and risk-free assets in the data, as in Kydland and Prescott [1982]. We calibrate the exogenous firm death rate, κ_D, to be 5% per year to match the average exit rate in the data. We choose $\delta = 7\%$ so that together with the exit rate, they imply a 12% effective annual depreciation rate of organization capital.

The second group of parameters is largely specific to our model. We choose the separation rate $\kappa_S = 5\%$ to match an average CEO departure rate of 10% reported by Fee and Hadlock [2004] (see also Fee, Hadlock, and Pierce [2018]). We choose $\kappa_M = 0.41$ to match CEOs' median time between jobs of 267 days as reported by Fee and Hadlock [2004]. We set $\lambda = 85\%$ to match an average income decline of 40% of CEOs after forced departure reported in Nielsen [2017].

As shown in Section 3, the marginal product of capital, A, determines firms' incentives to invest in organization capital and, therefore, the equilibrium growth rate of firms. Fixing the total supply of physical capital and labor, A is decreasing in the total stock of organization capital of the economy. The stock of organization capital depends on the initial size, $\bar{Z}$, and the initial entry rate, $\bar{e}$. We normalize $\bar{e}$ so that the total measure of operating firms is one in the steady state, and choose $\bar{Z}$ so that the implied $A = 0.21$, which allows our model to match the median sales growth rate of 3.2% in the data.[11] The volatility parameter σ is set at 35% to match the average volatility of firms' sales growth in the data. We set the span of control parameter $\nu = 0.85$ as in Atkeson and Kehoe [2005]. We choose capital adjustment cost parameter $h_0 = 5$, so that with $A = 0.21$, our model matches an average Tobin's Q of 1.67 in our sample.[12]

Finally, we calibrate the return to scale of the matching technology as follows. As shown in Proposition 4, $\psi = \psi_X + \psi_Y$, determines the relative magnitude of power laws in firm size and CEO pay. According to our estimates, which we discuss in detail below, the power law slope of CEO pay is about 2.28 and that of market capitalization is about 1.10. Therefore, we set $\psi = \frac{1.10}{2.28} = 0.48$ to match the relative magnitude of the power law slopes. The parameter ψ_X determines the equity share of the manager in a newly matched firm-manager pair (see Equation (22)). We choose $\frac{\psi_X}{\psi} = 0.89$ to account for the average CEO pay to size ratio of 0.091 of young firms, which are defined as firms of less than five years of age.

We solve the model numerically and aggregate simulated data from the continuous-time model to an annual frequency, which corresponds to sampling frequency of the observed data. Our simulated sample consists of two million firms and, as such, can be treated as population.

[11] The choice of normalization does not affect the model's implication for the relative size of firms and the distribution of firm size.

[12] We provide details of the calculation of the marginal product of organization capital and Tobin's Q in Appendix E.

5.2 Basic Statistics of Firm Dynamics

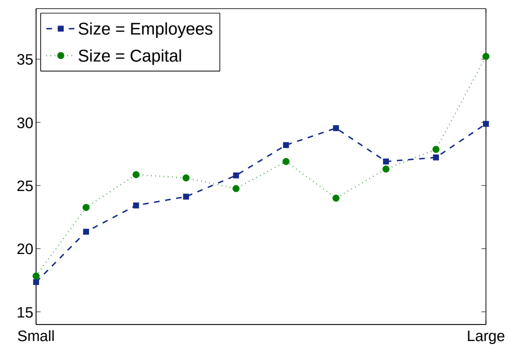
In this section, we show that our model is able to replicate key stylized features of the joint cross-sectional distribution of firm size, age, investment, CEO compensation and dividend payout. To this end, we construct a cross-section of size sorted portfolios. We follow the standard sorting procedure 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 benchmark measures of firm size in the data: gross capital and the number of firm employees. Recall that in the model, both capital and labor are proportional to organization capital and, therefore, are equivalent measures of size.

First, consistent with the data, our model implies a monotonic relationship between firm size and age. Figure 5 shows the observed and the model-implied variation of the median firm age across size-sorted portfolios (Panels (a) and (b), respectively). Note that the monotonic relationship between age and size is a stronger requirement on the model than a positive statistical correlation between size and age. Any model with a positive expected growth features a positive correlation between size and age but not necessarily a monotonic relationship between the two. For instance, as shown in Proposition 2, in the first-best economy without agency frictions, the distribution of firm size is fat not only in the right tail but also the left tail. As such, the first-best economy features a J-shaped relationship between age and size because the very small firms in this setting are old firms that have experienced a long history of negative productivity shocks.

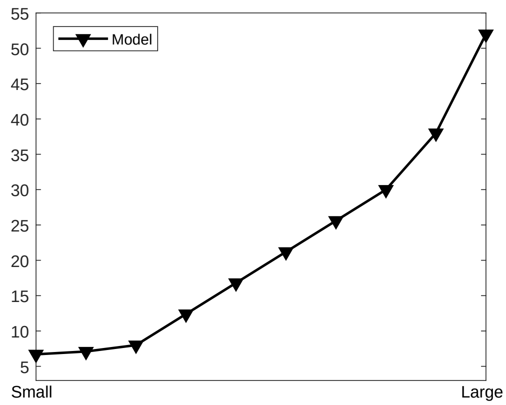
The decreasing returns to scale of the matching technology is the key assumption that allows our model to account for the monotonic positive relationship between firm size and age. The DRS matching technology implies that as the match-specific organization capital decreases, the outside options of firms and managers become more attractive relative to what they can produce within their match. Therefore, firms that continue to experience negative shocks find it optimal to separate with their managers. The ensuing endogenous separation rules out the fat left tail of the firm distribution. Thus, the smallest firms in our model are typically young firms that yet have not had time to grow (not old firms that have experienced a long sequence of negative shocks). In addition, this same feature of the model implies that small firms have a higher exit rate of about 8.9% per year than large firms that exit at a rate of 5% per year all due to exogenous death. This implications is also consistent with the data — empirically, the exit rate of small firms is about twice that of large firms.

Despite its simplicity, our model fares well with the empirical distribution of firm age. Luttmer [2010] documents that the median age of firms with more than $10,000$ employees

Figure 5: Distribution of Age across Size



(a) Data



(b) Model

Figure 5 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. Firm age on the vertical axis is measured in years.

in 2008 was about 75 years. In our benchmark model, this number is 92.[13] Luttmer [2012] shows that to jointly account for the decline in volatility with firm size 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 helps our model account for the joint distribution of firm age and size.

Second, our model quantitatively accounts for the negative relationship between firm size and investment observed in the data. Table 3 shows average investment rates (defined

[13]There are about 6 million employer firms in 2008 in the US and the largest 1000 firms with more than 10,000 employees account for 27% of the total employment (Luttmer [2010]). In our calibration, the median age of the largest firms that account for 27% of total employment of the economy is 92 years.

as annual investment divided by the beginning-of-year capital stock) of quintile portfolios sorted by 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 a rate of 9%. As the table shows, the difference in investment rates of large and small firms is strongly statistically significant. Our model matches well the large cross-sectional dispersion observed in the data — the model-implied difference in average investment rates of firms in the bottom and top quintile portfolios is 8%.[14] As explained in Section 4.1.3, the inverse relationship between size and investment is due to limited commitment. Managers in small firms are poorly diversified due to limited commitment on the firm side. Therefore, small firms optimally choose to accelerate investment to grow out of the agency conflict. As firm size increases, managers' outside options become more attractive prompting large firms to reduce investment to limit managers' incentives to walk away.

The inverse relationship between size and growth is also reflected in the cross-sectional variation in firms' dividend payout policies. In the data, large firms are much more likely to pay dividends to shareholders while small firms tend to retain earnings for investments. Table 4 shows the fraction of dividend-paying firms in each size-sorted portfolio in the data and in the model. 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 captures well the monotonic increase in the likelihood of dividend payments with size.

Third, the elasticity of CEO pay with respect to firm size is positive but less than one, as predicted by our model with limited risk sharing. Table 5 shows the empirical and the model-impled elasticities of managerial compensation with respect to size estimated in a panel regression of CEO pay on lagged firm size (both measured in logs), controlling for firm and time fixed effects. As the table shows, our calibrated model implies an elasticity of about 0.25, which is similar to the data estimates. In the data, the size elasticity of CEO compensation is about one-third and is overall quite robust to various measure of firm size (consistent with the earlier evidence in Gabaix [2009]).

Note that while the level of managerial pay increases with firm size, the ratio of CEO pay to firm size declines with size due to limited risk sharing. Table 6 presents the variation of the CEO pay-to-capital ratio across size-sorted portfolios. In the data, the median ratio falls from 6.2% for small firms to 0.4% for large firms when size is measured by the number of firm employees, and from 7.6% to 0.1% when size is measured by firm capital. The model features a similar cross-sectional pattern.

[14] We do not present t-statistics of the difference in the model-implied moments because the reported model statistics represent population moments.

5.3 Power Laws

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 present our empirical estimates of the power laws and then compare the quantitative implications of our model to the data.

Following Luttmer [2007] and Gabaix [2009], 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:

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

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

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

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

It has been shown in the literature that firm size follows a power-law distribution (for example, Axtell [2001], Gabaix [2009], and Luttmer [2007]). We confirm this evidence and show that the empirical distributions of CEO compensation and dividends are also fat-tailed. We estimate the power-law coefficients year by year and present time-series averages of the estimated parameters in Table 7. The table also reports sample averages of 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 F.

On average, the estimate of the power-law coefficient of firm size is about 1.2 when size is measured by the number of employees, 1.5 when size is measured by gross capital, and about 1.1 when size is measured by market capitalization. The latter is very close to the estimates obtained using Census data. For example, Luttmer [2007] reports a power law estimate of 1.07; similar estimates are reported in Gabaix and Landier [2008]. Overall, the goodness-of-fit test does not reject the power-law null — with just few exceptions, year-by-year p-values are above the conventional five-percent level for all measures of firm size. Notice that the power-law coefficient of dividends is very close to that of firm size, particularly of market capitalization. In contrast, CEO compensation is characterized by a much larger power-law coefficient of about 2.3. 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.

Consistent with the data, our calibrated model produces a power law in firm size and

Figure 6: The Right Tails of Firm Size and CEO Compensation

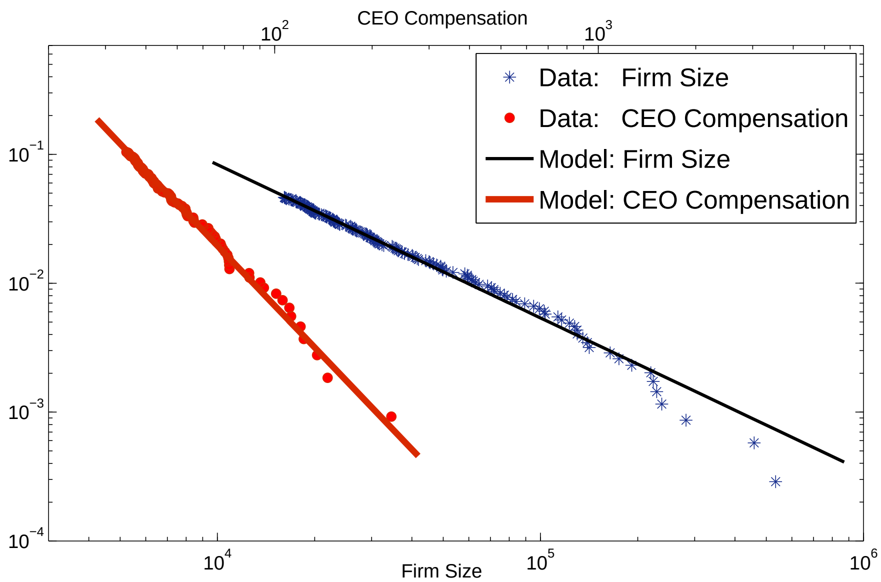


Figure 6 plots the right tails of the distributions of firm size and CEO compensation (using 2006 data) and the corresponding slopes implied by the model. Firm size in the data is measured by market capitalization. The top and bottom horizontal axes represent CEO compensation and firm size, respectively. The vertical axis shows the complementary cumulative distribution function. The horizontal and vertical axes are on a logarithmic scale.

dividends with a slope close to one. In particular, the model-implied tail slope of firm size and dividend payments is 1.09. Recall that our calibration implies that the tail of the CEO-pay distribution is about half that of firm size. Hence, the model-implied exponent of the power law in managerial compensation is about 2.3. Figure 6 provides a visual comparison of the tail behavior of the model-implied distribution and the empirical distribution constructed using a representative sample year. The top and bottom horizontal axes represent CEO compensation and firm size, respectively, and the vertical axis shows the complementary cumulative distribution function, both are equally spaced on the log scale. Under power law, the log-log plot is a straight line with a slope equal to the negative exponent. In the data, the firm-size distribution is represented by stars and CEO-compensation distribution is represented by circles. The solid thin and thick lines are the model-implied power laws in firm size and managerial compensation, respectively. As the figure shows, the model matches well the tail slopes observed in the data.

As shown in Section 4.2, limited commitment on the manager side implies that CEO compensation in large firms is a linear function of the running maximum of Z^{ψ} and, therefore, obeys the same power law as Z^{ψ}. That is, the optimal contract under limited commitment translates the power law in firm size 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 stated in Proposition 4 remain valid in our baseline model with two-sided limited commitment.

5.4 Dynamics of CEO Compensation

An important feature that distinguishes our model from the static setup of Gabaix and Landier [2008] is its implications for the time-series dynamics of CEO pay. As explained in Section 4.3, a unique prediction of our model is the asymmetric response of CEO compensation to productivity shocks — in small firms, CEO compensation is sensitive to negative shocks, whereas in large firms, managerial compensation is sensitive to positive shocks.

To test these implications in the data, we run the following panel regression:

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

We regress the log-growth of CEO compensation ($\Delta c_{i,t} \equiv \ln \frac{C_{i,t}}{C_{i,t-1}}$) on the log-growth of firm size ($\Delta k_{i,t-1} \equiv \ln \frac{K_{i,t-1}}{K_{i,t-2}}$), 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: $\Delta^{Min} k_{i,t} \equiv \ln \frac{K_{i,t}}{K_{i,t-1}^{Min}}$ and $\Delta^{Max} k_{i,t} \equiv \ln \frac{K_{i,t}}{K_{i,t-1}^{Max}}$, where $K_{i,t-1}^{Min}$ and $K_{i,t-1}^{Max}$ 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: $\Delta^{Min-} k_{i,t}$ 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. $I_{i,t}^{Small}$ and $I_{i,t}^{Large}$ 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.

In our model, small firms have to cut down managerial compensation if they continue to receive negative productivity shocks and their size keeps declining. Thus, the model predicts a positive coefficient on the first interaction term $\Delta^{Min-} k_{i,t} \cdot I_{i,t}^{Small}$. Large firms in the model have to offer higher compensation to retain their managers if they continue to experience positive productivity shocks and keep growing. Thus, we expect to see a positive coefficient on the last term $\Delta^{Max+}k_{i,t} \cdot I_{i,t}^{Large}$.

Panel A of Table 8 presents the estimates of our panel regression. First, notice that in 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: 0.52 and 0.63, 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 the firm- or manager-side constraint and have no need to adjust managerial compensation. Also, the coefficient on the leading term ($\Delta k_{i,t-1}$) is relatively small, of about 0.05, due to the inelastic response of CEO compensation to changes in firm value for medium-size firms.

"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 0.28 and 1.01, respectively. The data estimates are generally higher than their model counterparts, 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 8, 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. 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.

5.5 Cross-Industry Analysis

As discussed in Section 4.3, our model has a unique set of cross-sectoral implications. In particular, our model predicts that the impact of agency frictions increases with the return to scale of the matching technology. Hence, industries with less decreasing returns to scale of the matching technology, feature high managerial compensation relative to firm size and a strong inverse relationship between firm investment and size, whereas industries with significantly decreasing returns to scale of the matching technology (i.e., $\psi \ll 1$) are characterized by a relatively low ratio of CEO pay to size and a relatively thin right tail of the distribution of CEO compensation.

In this section, we evaluate the cross-sectoral implications of the model using a set of five industry portfolios. We first sort all firms into five industries: "Consumer", "Manufacturing", "Hitech", "Health", and "Others", and compute their average CEO compensation.[15] As Table 9 shows, we find a sizable dispersion in the average CEO pay to size ratio across the five industries. In particular, keeping firm size equal, CEOs in "Hitech" and "Health" sectors typically earn three to four times more relative as their peers in other industries. We confirm below that the gap in managerial compensation between the two groups is highly statistically significant. According to our model, this evidence suggests that organization capital in "Hitech" and "Health" sectors is highly valuable in the sense that it can be productively employed in another match if managers decide to voluntarily separate with their current employers. That is, "Hitech" and "Health" industries are likely to have a relatively high return to scale of the matching technology. In contrast, industries like "Consumer" and "Manufacturing" are likely to feature a significantly decreasing returns to scale of the matching technology.

Guided the model's intuition, we group the five industries into two categories: "High-ψ" group that consists of "Hitech" and "Health" sectors, and "Low-ψ" group that consists of the remaining industries.[16] We next examine if the two industry groups feature any discernible differences in their compensation and investment policies as implied by our model.

First, consistent with the model's predictions and our classification of industries, we find that in "High-ψ" group, the power-law coefficient of CEO pay is closer to that of firm size, whereas CEO compensation in "Low-ψ" group has a considerably thinner tail compared to the distribution of firm size. For each industry group, we estimate the power-law exponents for firm size measured by the number of employees and CEO compensation year by year over the 1992-2016 sample period, and in Table 10 we report time-series averages on the estimated coefficients. Recall that the ratio of power law in firm size to power law in CEO pay measures the degree of returns to scale of the matching technology, ψ (see Proposition 4). As the table shows, "High-ψ" group indeed features a higher ψ compared with "Low-ψ" group. In particular, the implied returns to scale of the matching technology of "Low-ψ" and "High-ψ" industries are about 0.47 and 0.75, respectively.

We next examine CEO compensation and investment policies in the two industry groups and compare them with the implications of our model calibrated under two alternative assumptions for the return to scale of the matching technology: $\psi = 0.47$ and $\psi = 0.75$ that

[15]We use the Fama-French industry definition that is available on Kenneth French website.

[16]We combine the five industries into two groups to gain statistical power, which is particularly important in analyzing the tail behavior of the empirical distributions of firm size and CEO compensation.

correspond to our empirical estimates. To evaluate the cross-sectional variation in policy decisions, we sort firms in each industry group into three portfolios based on firm size.[17] To save space, we report average rates only for the bottom and top tercile portfolios, i.e., small and large firms.

Table 11 shows that in the data, CEO compensation to firm size ratio differs significantly in the two industry groups. The difference in the average CEO pay to size ratio between "High-ψ" and "Low-ψ" sectors is 3% with a t-statistic of 8.9. Further, consistent with the model's implications, the cross-sectional dispersion in managers' equity share is significantly larger in "High-ψ" group compared with "Low-ψ" group.

Table 12 presents the average investment rates and their cross-sectional dispersion in the two sectors. First, notice that on average, firms in "High-ψ" industry group invest at a higher rate compared with "Low-ψ" industries. The difference in investment rates is about 2.4% per annum with a t-statistics of 2.3. Second, consistent with the model's predictions, "High-ψ" group features a significantly larger cross-sectional variation in investment rates relative to "Low-ψ" group. On average, compared with large firms, small firms invest by 6.3% and 9% more in "Low-ψ" and "High-ψ" sectors, respectively, and the difference in small-minus-large spreads between the two groups is strongly statistically significant. Intuitively, the higher the magnitude of the return to scale of the matching technology, the more severe agency frictions and hence, the stronger the negative relationship between firm investment and size. Our calibrated model matches well the observed dispersion in investment rates, in particular, the large-minus-small spread in investment rates implied by the model increases in magnitude from -6.7% for "Low-ψ" sector to -8.8% for "High-ψ" sector.

6 Conclusion

We integrate the assortative matching model into the dynamic contracting theory with limited commitment to provide a unified theory of managerial compensation dynamics and labor market mobility. We use continuous-time tools to characterize the optimal dynamic contract and the implied distribution of firm size and CEO compensation. We show that our model generates a rich set of predictions that are consistent with empirical evidence.

Several possible extensions of our model may provide promising directions for future research. First, modeling endogenous separation and "on-the-job search" may allow our model to capture the history dependence of labor market separations. Second, for tractability, the specification of human capital dynamics in our model is quite stylized. A more general

[17] In Tables 11-12, we use gross capital as a measure of firm size; the evidence is similar if size is measured by the number of employees.

setting in which the outside options of managers are not perfectly correlated with their human capital may provide further insights about CEO compensation. Finally, it would be interesting to explore the implications of other agency frictions, such as moral hazard and hidden savings for firm dynamics.

References

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

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

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

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

Atkeson, Andrew, and Patrick J. Kehoe, 2005, Modeling and Measuring Intangible Capital, *The Journal of Political Economy* 113.

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

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

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

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

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.*

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

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

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

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

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

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

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

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

Fee, C. Edward, and Charles J. Hadlock, 2004, Management turnover across the corporate hierarchy, *Journal of Accounting and Economics* 37, 3–38.

Fee, C. Edward, Charles J. Hadlock, and Joshua R. Pierce, 2018, New evidence on managerial labor markets: An analysis of CEO retreads, *Journal of Corporate Finance* 48, 428–441.

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

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

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

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

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.

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

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

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

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

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

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

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

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.

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

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

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

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

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

Nielsen, Kasper Meisner, 2017, Personal Costs of Executive Turnovers, *Working paper, Hong Kong University of Science and Technology.*

Prescott, Edward C., and Michael Visscher, 1980, Organization Capital, *The Journal of Political Economy* 88.

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

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

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

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

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

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

Thomas, Jonathan, and Tim Worrall, 1988, Self-enforcing Wage Contracts, *The Review of Economic Studies* 55, 541–554.

Table 1

Calibrated Parameters

Description	Notation	Value
Marginal product of capital	A	0.21
Adjustment cost	h_0	5
Volatility	σ	35%
Depreciation rate	δ	7%
Death rate	κ_D	5%
Separation rate	κ_S	5%
CEO job finding rate	κ_M	41%
Separation loss	$1-\lambda$	15%
Return to scale	ψ	0.48
Manager share in match	$\frac{\psi_x}{\psi}$	0.89

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

Table 2
Targeted Moments

Moments	Data	Model
Investment/Output	17.2%	18.6%
Average Tobin's Q	1.67	1.67
Median sales growth	3.2%	3.2%
Average sales growth	10.1%	10.7%
Volatility of sales growth	37.1%	40.6%
CEO-pay/Capital of young firms	0.091	0.082
CEO median time between jobs	0.73	0.73
Average CEO departure rate	10%	10%
Relative slope of power law	0.48	0.48

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 calibrated model.

Table 3

Investment Rates

	Data		Model
	Size≡Employees	Size≡Capital	
Small	0.153	0.182	0.196
2	0.108	0.140	0.125
3	0.093	0.125	0.114
4	0.084	0.107	0.114
Large	0.092	0.088	0.118
Large−Small	−0.061	−0.094	−0.078
	(−12.32)	(−5.65)	

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 [1987] estimator with four lags are reported in parentheses.

Table 4

Fraction of Dividend-Paying Firms

	Data		Model
	Size≡Employees	Size≡Capital	
Small	0.12	0.12	0.10
2	0.34	0.37	0.25
3	0.45	0.48	0.81
4	0.53	0.62	0.98
Large	0.71	0.82	1.00
Large−Small	0.59	0.70	0.90
	(32.68)	(38.53)	

Table 4 presents the average fraction of dividend-paying firms 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 [1987] estimator with four lags are reported in parentheses.

Table 5

Elasticity of CEO Compensation to Firm Size

	Data			Model
	Market Cap	Employees	Capital	
Elasticity	0.36	0.36	0.33	0.25
	[0.015]	[0.022]	[0.012]	

Table 5 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 and standard errors clustered by firm and time (in brackets). The model statistics represent population numbers that are computed using a large panel of simulated data.

Table 6

CEO Pay-to-Firm Size Ratio

	Data		Model
	Size≡Employees	Size≡Capital	
Small	0.062	0.076	0.088
2	0.021	0.020	0.042
3	0.014	0.010	0.017
4	0.009	0.005	0.007
Large	0.004	0.001	0.002
Large−Small	−0.058	−0.075	−0.086
	(−20.66)	(−20.31)	

Table 6 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 [1987] estimator with four lags are reported in parentheses.

Table 7

Estimates of the Power-Law Parameters

	$\hat{\zeta}$	p-value
Employees	1.21	0.14
Market Cap	1.10	0.31
Gross Capital	1.50	0.40
Dividends	1.12	0.32
CEO Pay	2.28	0.39

Table 7 presents the estimates of the exponent of the power-law distribution (ζ) for the number of firm employees, market capitalization, gross capital, dividends and CEO compensation. The table reports time-series averages of the parameters estimated year-by-year in the 1992-2016 sample. Correspondingly, the reported p-values are time-series averages of year-by-year p-values of the Kolmogorov-Smirnov goodness-of-fit test.

Table 8
Dynamics of CEO Compensation

Regressors	Panel A: $Y = \Delta c_t$ Data	Panel A: $Y = \Delta c_t$ Model	Panel B: $Y = c_t$ Data	Panel B: $Y = c_t$ Model
Δk_{t-1}	0.106 (6.0)	0.046		
k_{t-1}^{Max}			0.178 (8.8)	0.046
c_{t-1}			0.265 (17.2)	0.824
$\Delta^{Min-}k_t \cdot I_t^{Small}$	0.280 (2.8)	0.520	0.320 (4.3)	0.486
$\Delta^{Min+}k_t \cdot I_t^{Small}$	0.021 (1.6)	−0.020	−0.047 (−2.9)	−0.009
$\Delta^{Max-}k_t \cdot I_t^{Large}$	0.044 (2.4)	−0.015	−0.019 (−0.3)	0.000
$\Delta^{Max+}k_t \cdot I_t^{Large}$	1.013 (2.1)	0.633	1.660 (3.7)	0.626

Table 8 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 deciles 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.

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

Table 9
Industry Sort

Industry	C/K	Group
Consumer	0.0136	Low-ψ
Manufacturing	0.0045	Low-ψ
HiTec	0.0380	High-ψ
Health	0.0420	High-ψ
Other	0.0171	Low-ψ

Table 9 presents the median ratio of CEO compensation to gross capital for five industry portfolios and the corresponding classification of the industry sectors into two groups, "Low-ψ" and "High-ψ".

Table 10

Industry Sort: Power Law

	Power Law ($\hat{\zeta}$)		Implied ψ
Industry Group	Employees	CEO Pay	
Low-ψ	1.12	2.36	0.47
High-ψ	1.64	2.17	0.75

Table 10 presents the estimates of the exponent of the power-law distribution (ζ) for the number of firm employees and CEO compensation in "Low-ψ" and "High-ψ" industry groups. The table reports time-series averages of the parameters estimated year-by-year in the 1992-2016 sample. The right panel shows the implied degree of returns to scale of the matching technology (ψ).

Table 11

Industry Sort: CEO Pay-to-Firm Size Ratio

	Firm Size				Model
Industry Group	All	Small	Large	Large−Small	Large−Small
Low-ψ	0.009	0.032	0.002	−0.030	−0.070
High-ψ	0.039	0.094	0.005	−0.089	−0.098
High−Low	0.030			−0.058	−0.028
	(8.92)			(−8.01)	

Table 11 presents the median ratio of CEO compensation to gross capital in "Low-ψ" and "High-ψ" industry groups. Small and large firms represent the bottom and top size-sorted tercile portfolios, respectively. Size in the data is measured by gross capital. T-statistics for the difference between "High-ψ" and "Low-ψ" industry groups based on the Newey and West [1987] estimator with four lags are reported in parentheses. The model-implied spreads are presented in "Model" column.

Table 12

Industry Sort: Investment Rates

	Firm Size				Model
Industry Group	All	Small	Large	Large−Small	Large−Small
Low-ψ	0.091	0.149	0.086	−0.063	−0.067
High-ψ	0.115	0.198	0.107	−0.090	−0.088
High−Low	0.024			−0.027	−0.021
	(2.26)			(−3.59)	

Table 12 presents the average investment-to-capital ratio in "Low-ψ" and "High-ψ" industry groups. Small and large firms represent the bottom and top size-sorted tercile portfolios, respectively. Size in the data is measured by gross capital. T-statistics for the difference between "High-ψ" and "Low-ψ" industry groups based on the Newey and West [1987] estimator with four lags are reported in parentheses. The model-implied spreads are presented in "Model" column.