The Journal of FINANCE
The Journal of THE AMERICAN FINANCE ASSOCIATION

THE JOURNAL OF FINANCE • VOL. LXXVI, NO. 1 • FEBRUARY 2021

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

HENGJIE AI, DANA KIKU, RUI LI, and JINCHENG TONG*

ABSTRACT

We develop a unified theory of dynamic contracting and assortative matching to explain firm dynamics. In our model, neither firms nor managers can commit to arrangements that yield lower payoffs than their outside options, which are microfounded by the equilibrium conditions in a matching market. The 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 investment, dividend payout, and CEO compensation.

IN THIS PAPER, WE PROVIDE a unified theory of limited commitment and assortative matching to account for several stylized features of the dynamics of firm investment, dividend payout, and managerial compensation. The contracting literature on limited commitment has developed a compelling theoretical framework describing 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, in existing models are specified exogenously, are difficult to measure empirically. To impose discipline and to provide a microfoundation 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

*Hengjie Ai is affiliated with the Carlson School of Management, University of Minnesota. Dana Kiku is at the University of Illinois at Urbana-Champaign. Rui Li is at the University of Massachusetts Boston. Jincheng Tong is at the University of Toronto. 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. The authors have no conflicts of interest to disclose as identified in *The Journal of Finance* disclosure policy.

Correspondence: Hengjie Ai, Carlson School of Management, University of Minnesota, 321 19th Ave South, Minneapolis, MN 55455; email: hengjie@umn.edu.

DOI: 10.1111/jofi.12980

and managers are jointly determined by the equilibrium conditions in a matching market, and the time-series dynamics and cross-sectional distribution of firm size, investment, and CEO pay are determined by the optimal contract.

We establish two theoretical results. First, we show that limited commitment on the manager side translates a power law in firm size into a power law in CEO compensation, and that the slope of the power law in CEO compensation is determined by the returns to scale of the matching technology. Second, we demonstrate that limited commitment on both the firm and the manager sides results in an inverse relationship between firm investment and firm size. We further show that our model is able to quantitatively account for the key features of the joint distribution of firm size, age, growth, and managerial compensation.

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's (1931) 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 that managerial pay does not respond to firm performance, which is clearly inconsistent with 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 microfoundation for outside options, we introduce a market in which managers and firms meet to form new productive relationships. A firm's organization capital 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 human capital, 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 increases the 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 separation is inefficient.

To derive the power law in CEO compensation, we first show that permanent productivity shocks result in a power law in firm size (see also Luttmer (2007)). Under the optimal contract, CEO compensation is a linear function of the running maximum of managers' outside options. Because the running maximum of a power law process obeys the same power law, the power law coefficient on CEO pay depends on that of firm size and the elasticity of managers' outside options with respect to size, which is determined by the returns to scale of the matching technology.

We show that both limited commitment constraints are responsible for an inverse relationship between firm growth and size. When a firm's value declines, the firm-side commitment constraint is likely to bind. To avoid further losses and a potential separation, small firms choose to defer dividends, reduce 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 managers' departure, it is optimal for large firms to decrease their investment and increase CEO compensation. Hence, consistent with the data, small firms in our model pay out less, invest more, and grow faster compared with large firms.

We calibrate our model and evaluate its quantitative implications that are central to the economic mechanism we propose. We show that our model is able to replicate the wedge in the right-tail characteristics of the empirical distributions of firm size and CEO compensation. We also show that the inverse relationship between investment and size and that between investment and age predicted by our model are quantitatively consistent with the data. Our model therefore improves significantly upon standard frictionless models in matching the life-cycle dynamics of firm growth.

We next present empirical evidence that supports a unique set of restrictions on the joint cross-sectional distribution and time-series dynamics of CEO compensation and firm investment implied by our model. In particular, we show that the CEO-pay-to-size ratio in our model is informative about the distance to the firm-side limited commitment constraint, and thus, firms with high CEO-pay-to-size ratios feature high investment rates. Consistent with the model's prediction, we find a strong positive correlation between CEO-pay-to-size ratios and investment rates in the cross section of publicly traded firms. In the time series, our model predicts that firms that have recently experienced a sequence of positive (negative) shocks that push them toward limited commitment constraints are likely to simultaneously increase (decrease) CEO pay and reduce (raise) the investment rate. Consistent with the model, we show that investment and CEO compensation in the data are history-dependent and feature-pronounced asymmetric responses to firm-specific productivity shocks.

The trade-off 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 the compensation of 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 with two-sided limited commitment in our model is closely related to Ai and Li (2015), Bolton, Wang, and Yang (2019), and Ai and Bhandari (2020). None of the above papers explicitly models a matching market and relate the power law of CEO compensation to the returns to scale of the matching technology.

Our paper also builds on both models of power law as well as 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) propose assortative matching models that link CEO compensation to firm size taking the size distribution as given. Gomez (2016) studies power law in wealth distribution in an asset pricing model with overlapping generations.

Our paper is also related to a 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 et al. (2012), Biais, Mariotti, and Villeneuve (2010), and Axelson and Bond (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 I. In Section II, 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 III. We calibrate our model and evaluate its quantitative implications in Section IV. Section V provides concluding remarks. Additional discussions and omitted proofs are included in an Internet Appendix.[2]

[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).

[2] The Internet Appendix is available in the online version of the article on *The Journal of Finance* website.

I. Model Setup

In this section, we set up an industry equilibrium model with heterogeneous firms and limited commitment. There are two types of agents in our model: firms and managers. A firm requires a manager to produce output. We define "corporation" as a productive relationship formed within a matched firm-manager pair. We use the term "firm" to refer to a generic firm that may or may not be matched with a manager.

A. Production Technology

Time is continuous and infinite. As in Atkeson and Kehoe (2005), the output of corporation j, denoted by y_j, is produced from organization capital (Z_j), physical capital (K_j), and labor (N_j) using a standard Cobb-Douglas production technology, $y_j = Z_j^{1-\nu}(K_j^{\alpha}N_j^{1-\alpha})^{\nu}$, where $\nu \in (0, 1)$ is the span-of-control parameter.[3] The operating profit of firm j is defined by

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

where MPK is the rental rate of physical capital and MPL is the equilibrium wage of unskilled workers. We assume that unskilled labor, N, and physical capital, K, are not corporation-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, that is, $\pi(Z_j) = AZ_j$, where A is the equilibrium marginal product of organization capital. Due to the CRS production technology, the optimal K_j and N_j are proportional to Z_j. Hence, in our model, these three variables are equivalent measures of firm size.

Organization capital, Z_j, is specific to a firm-manager pair. We assume that the accumulation of organization capital depends on the 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, $h(\frac{I}{Z})Z$, where $h(i) = i + \frac{1}{2}h_0 \cdot i^2$ with $h_0 > 0$.

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

B. Entry, Exit, and Matching Technology

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 denoted by $\bar{Z}$. Here, we assume for simplicity that the initial distributions are a point mass.[4] Newly arrived firms and newly arrived managers meet immediately in a directed matching market to form corporations, 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, organization capital of the firm and human capital of the manager evaporate.

The operating firm-manager pair may separate for exogenous reasons that 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. We then have

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

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 Cobb-Douglas technology specified above but are not allowed to access the credit market.[5] That is, idle managers can consume the cash flow from home production but are not allowed to enter into any financial contract for the purpose of risk sharing. Home production cannot be used to invest in human capital, and therefore, the law of motion of 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 thus their organization capital follows the same law of motion, that is,

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

where j is interpreted as the index of the firm.

[4] Because $Z_{j,t}$ follows a diffusion process after arrival, the steady-state distribution of $Z_{j,t}$ has a continuous density with a mode at $\bar{Z}$. The density is everywhere differentiable (and obeys a forward equation) except at $\bar{Z}$, as shown in Proposition 2 below. Note that as long as initial distributions have a bounded support, the right tails of the distributions of firm size and human capital are determined solely by their equilibrium dynamics and are not affected by the functional form of initial distributions.

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

At Poisson rate κ_M, idle firms and managers have an opportunity to meet in a directed matching market in which idle firms offer competitive contracts to managers to form corporations. We denote the stopping time at which an idle firm or an idle manager has an opportunity to go to the directed matching market by τ_M. If the contract is accepted, firms and managers become productive corporations. We assume that the initial level of organization capital of a newly formed corporation depends on the firm-specific organization capital, Y, and the manager-specific human capital, X, through a Cobb-Douglas function,

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

where τ denotes the stopping time at which a match is formed and the new firm starts operation.[6] After a match is formed, Z_t follows the dynamics described in equation (2). Thus, the level of the match-specific organization capital, $Z_{j,t}$, is determined by the manager-specific human capital and the firm-specific organization capital upon the match, after which it is subject to match-specific shocks (see equation (2)) until separation or exit. Our model therefore assumes in effect that manager-specific human capital and firm-specific organization capital are perfectly correlated once they are matched and become a corporation. The assumption of perfect correlation is made for tractability.

The parameter ψ can be interpreted as the returns to scale of the matching technology. Decreasing returns to scale (DRS), that is, $\psi < 1$ captures the idea that it is harder for organization capital to impact large firms than small firms. This assumption is consistent with empirical evidence in Baker and Hall (2004), who estimate the elasticity of the marginal product of CEO effort with respect to firm size to be approximately 0.4.

C. Preferences

A contract offered to a manager at time t, denoted by $\{\{C_s, i_s\}_{s=t}^{\tau_D \wedge \tau_S}\}$, specifies compensation of the manager and investment in organization capital as functions of the history of shock realizations. 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 if the firm is hit by a death shock or the firm-manager pair separates.[7] 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, by

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

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

[7] We specify a contract to be a sequence of managerial compensation and investment that are functions of the history of shocks. This specification is without loss of generality as long as both compensation and investment are publicly observable and fully contractible.

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 fraction λ 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 one in the event that $\tau_S < \tau_D$.

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 that a manager with human capital X can achieve in the directed matching market. Then $U^S(X)$ must satisfy

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

That is, an unemployed manager consumes home production 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 $\bar{U}(X_{\tau_M})$.

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

$$E_t\left[\int_t^{\tau_D\wedge\tau_S} e^{-r(s-t)}\left[AZ_s - C_s - h\left(\frac{I_s}{Z_s}\right)Z_s\right]ds + \mathbb{1}_{\{\tau_S<\tau_D\}}e^{-r(\tau_S-t)}V^S\left(\lambda Z_{\tau_S}\right)\right], \tag{8}$$

where $V^S(Y)$ denotes the value of an idle firm with organization capital Y. The term $AZ_s - C_s - h(\frac{I_s}{Z_s})Z_s$ is the dividend payment at time s, which we denote by D_s.[8]The term $V^S(\lambda Z_{\tau_S})$ 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

$$V^S(Y) = E_t\left[\left(\mathbb{1}_{\{\tau_M<\tau_D\}}e^{-r(\tau_M-t)}\bar{V}\left(Y_{\tau_M}\right)\right| Y_t = Y\right], \tag{9}$$

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

For tractability, our formulations in equations (6) and (7) normalize managers' utility to be homogeneous of degree 1 in consumption. 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 at the initiation of the contract. In our environment, the stock of organization capital and the promised utility constitute a state variable pair that is sufficient to summarize any information that is relevant for

[8] Our formulation allows dividends to be negative, which can be interpreted as equity issuance. The qualitative implications of the model are not affected if one imposes a nonnegativity constraint on dividends.

the design of the optimal contract. We let $V(Z, U)$ denote the value of a firm with organization capital Z and promised utility to its manager U.

D. Limited Commitment

We assume that both firms and managers can unilaterally choose to separate at any time and cannot commit not to do so ex ante. Upon separation, firms and managers stay 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,

$$U_t \geq U^S(\lambda Z_t). \tag{10}$$

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

$$V(Z_t, U_t) \geq V^S\big(\lambda Z_{\tau_S}\big). \tag{11}$$

Because both firm value and the manager's utility are strictly increasing, if one of the inequalities in equations (10) and (11) 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 (10) and (11) 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. Competition in the matching market therefore determines outside options of firms and managers of all types.

E. Matching Decisions

In the directed matching market, firms of all types offer competitive contracts to managers to form productive corporations. We construct equilibria in which the matching rule is a bijection, 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 by $X(Y)$. Clearly, $Y(\cdot)$ and $X(\cdot)$ are inverse functions of each other. We 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 a firm is given by Equation (5), if matched with a manager with human capital X, the value of the firm is $V(Y^{\psi_Y} X^{\psi_X}, U)$, 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,

$$X(Y) \in \arg\max_X V\big(Y^{\psi_Y} X^{\psi_X}, \bar{U}(X)\big). \tag{12}$$

That is, for each Y, manager type $X(Y)$ is chosen optimally to maximize firm value.

Given the function $\bar{U}(X)$, equation (12) determines the optimal matching rule for firms of all types. Market clearing requires that the function $\bar{U}(\cdot)$ be chosen so that the decision rule of the associated firms', $X(Y)$, ensures that the matching market clears, that is, all firms and managers pair into productive units as soon as they meet. As a result, firm value upon receiving an opportunity to match is given by

$$\bar{V}(Y) = V\big((Y)^{\psi_Y} X(Y)^{\psi_X}, \bar{U}(X(Y))\big). \tag{13}$$

F. 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[9]

$$dU = \left[\frac{\beta+\kappa}{1-\gamma}\left(U - C^{1-\gamma}U^{\gamma}\right) + \frac{1}{2}\gamma\frac{G(Z,U)^2\sigma^2}{U} + \frac{\kappa_S}{1-\gamma}U\left(1-\left(\frac{U^S(\lambda Z)}{U}\right)^{1-\gamma}\right)\right]dt$$
$$+\, G(Z,U)\sigma\, dB. \tag{14}$$

Formally, an equilibrium in our model consists of the following quantities: an optimal recursive contract $\{C(Z,U), I(Z,U), G(Z,U)\}$, an equilibrium matching rule, $Y(X)$, and the maximum values that firms and managers can achieve on the matching market, $\bar{U}(X)$ and $\bar{V}(Y)$, that satisfy the following conditions:

[9] 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 details on the derivation in Section IV of the Internet Appendix.

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 (8) subject to the constraint that the initial promised utility to the manager is at least U_τ and subject to the limited commitment constraints in equations (10) and (11).
3. Given the equilibrium outside option of managers, $\bar{U}(X)$, the equilibrium matching rule satisfies firms' optimality in equation (12).
4. The outside options of firms are determined by the optimal matching rule in equation (13).
5. The optimal matching rule implied by firm optimality, that is, equation (12), clears the matching market. Thus, for any Y, the rate of arrival of type Y firms equals the rate of arrival of type $X(Y)$ managers. In addition, the prices of capital and labor (*MPK* and *MPL*) clear the markets for physical capital and labor.

In general, both the distribution of idle managers and the distribution of idle firms should be part of equilibrium objects. Let $\Phi(X)$ and $\Psi(Y)$ denote the distributions of manager and firm types, respectively. Because a constant fraction of managers and firms flows to the matching market at each point in time, under the equilibrium matching rule, $Y(X)$, market clearing requires $\Phi(X) = \Psi(Y(X))$ for all X so that firms and managers instantaneously pair with each other. In our model, the dynamics of X and Y are completely symmetric. Therefore, for all X, $\Phi(X) = \Psi(X)$, and market clearing holds under the symmetric matching rule, $Y(X) = X$. In the rest of the paper, we focus on the equilibriums with a symmetric matching rule without explicitly solving for the distributions $\Phi(X)$ and $\Psi(Y)$.

G. 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

$$V_Z\big(X^{\psi_Y+\psi_X}, \bar{U}(X)\big)\psi_X X^{\psi_Y+\psi_X-1} + V_U\big(X^{\psi_Y+\psi_X}, \bar{U}(X)\big)\bar{U}'(X) = 0. \tag{15}$$

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

II. The First-Best Case

As a starting point of our analysis, we first consider the first-best case in which no separation occurs, that is, $\kappa_S = 0$, and firms and managers fully commit. Here, shareholders maximize the present value of the firm's cash flow subject to the constraint that the contract must provide the manager with an initial promised utility of at least U. The present value of cash flow can be written as

$$E_t\left[\int_t^{\tau_D} e^{-r(s-t)}\left(AZ_s - h\left(\frac{I_s}{Z_s}\right)Z_s\right)ds\right] - E_t\left[\int_t^{\tau_D} e^{-r(s-t)}C_s ds\right]. \quad (16)$$

Note that the participation constraint only affects 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 (16) 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.[10]

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 constant consumption for 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*

$$A > r + \delta + \kappa > \frac{-1 + \sqrt{1 + 2h_0 A}}{h_0}. \quad (17)$$

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

$$V(Z, U) = \bar{v}Z - \frac{1}{r+\kappa}U, \quad (18)$$

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

$$\hat{\imath} = \arg\max_{i<\hat{r}} \frac{A - h(i)}{\hat{r} - i} = \hat{r} - \sqrt{\hat{r}^2 - \frac{2}{h_0}(A - \hat{r})}, \quad (19)$$

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

PROOF: See Section I of the Internet Appendix.

The first term $\bar{v}Z = h'(\hat{\imath})\,Z$ in equation (18) is the firm value in the neoclassical model with CRS production function and quadratic adjustment costs. The

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

second term is the present value of the cost of managerial compensation. In the absence of aggregate uncertainty, perfect risk sharing implies constant managerial compensation, the present value of which is simply given by Gordon's (1959) formula: $\frac{1}{r+\kappa}U$.[11] 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 details on this calculation in Section V of the Internet Appendix.

The constant compensation policy, $C_t = U_0$ for all t, implies that the steady-state distribution of CEO compensation is the same as the initial entrance distribution of CEO pay. In particular, it collapses to a mass point at U_0 if all corporations start at the same initial condition. In addition, equation (19) 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.\negthinspace across firms and the distribution of firm size follows a power law as in Luttmer (2007), which we summarize 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 stationary distribution of firm size is given by

$$\phi(Z) = \begin{cases} \frac{1}{\sqrt{\left(\hat{\imath}-\delta-\frac{1}{2}\sigma^2\right)^2+2\kappa\sigma^2}} Z^{\xi-1} & Z \geq \bar{Z} \\ \frac{1}{\sqrt{\left(\hat{\imath}-\delta-\frac{1}{2}\sigma^2\right)^2+2\kappa\sigma^2}} Z^{\eta-1} & Z < \bar{Z}, \end{cases} \tag{20}$$

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

PROOF: See Section I of the Internet Appendix.

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 of 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 if all corporations start at the same initial condition of promised utility. Hence, the dynamics of managerial compensation under the optimal contract does not endogenously generate a fat tail in the distribution of CEO pay.

[11] Recall that we normalize the utility function of the manager so that the lifetime utility is measured in consumption units.

III. Limited Commitment

A. 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 CRS matching technology, that is, $\psi = 1$, because it significantly simplifies the optimal contracting analysis. We discuss the general case of DRS matching technology in the Internet Appendix. 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 (15) simplifies to a restriction on $v(u)$ at $\bar{u}$,

$$\psi_X v(\bar{u}) + (1 - \psi_X)v'(\bar{u})\bar{u} = 0. \tag{21}$$

A.1. Value Function and Dynamics of Continuation Utility

Due to the assumption of the CRS of 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 lifetime 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 loses $(1 - \lambda)$ 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

$$\left(r + \kappa_M + (1-\gamma)\delta + \frac{1}{2}(1-\gamma)\gamma\delta^2\right)u_{MIN}^{1-\gamma} = (r + \kappa_D)A^{1-\gamma} + \kappa_M\bar{u}^{1-\gamma}. \tag{22}$$

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

Similarly, because firms also have the option to voluntarily separate, $V(Z,U) \geq V^S(\lambda Z)$ should hold. Given equation (9), the monotonicity of the value function $v(u)$ implies that $V^S(Y) = v(u_{MAX})Y$ for some u_{MAX} that satisfies

$$\kappa_M v(\bar{u}) = (r + \kappa_D + \kappa_M + \delta)v(u_{MAX}). \tag{23}$$

That is, limited commitment on the firm side requires $u_t \leq u_{MAX}$ for all t. The following proposition characterizes the optimal contract and the equilibrium levels of $\bar{u}$, u_{MIN} and u_{MAX}.[12]

PROPOSITION 3: *(Optimal Contracting with Two-Sided Limited Commitment)*

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

$$0 = \max_{c,i,g} \left\{ \begin{array}{c} A - c - h(i) + v(u)(i - \delta - r - \kappa) \\ +uv'(u) \left[\begin{array}{c} \frac{r+\kappa}{1-\gamma}\left(1 - \left(\frac{c}{u}\right)^{1-\gamma}\right) - (i - \delta) + \frac{1}{2}\gamma g^2 \sigma^2 \\ + \frac{\kappa_S}{1-\gamma}\left(1 - \left(\frac{\lambda u_{MIN}}{u}\right)^{1-\gamma}\right) \end{array} \right] \\ + \frac{1}{2}u^2 v''(u)(g-1)^2\sigma^2 + \kappa_S \lambda v(u_{MAX}) \end{array} \right\} \tag{24}$$

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 (21). Given $\bar{u}$, u_{MIN} is determined by equation (22) and u_{MAX} is determined by equation (23).*
4. *The optimal compensation ratio, $c(u_t) = \frac{C_t}{Z_t}$, takes the form*

$$\ln c(u_t) = \ln C_0 - \ln Z_t + l_t^+ - l_t^-, \tag{25}$$

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

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

PROOF: See Section II of the Internet Appendix.

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 managers implies that a smaller fraction of cash flow is retained by the firm and thus 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 (23) and 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 (21), 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.

[12] A recent paper by Bolton, Wang, and Yang (2019) shows that similar contracts can be implemented by corporate liquidity and risk management policies.

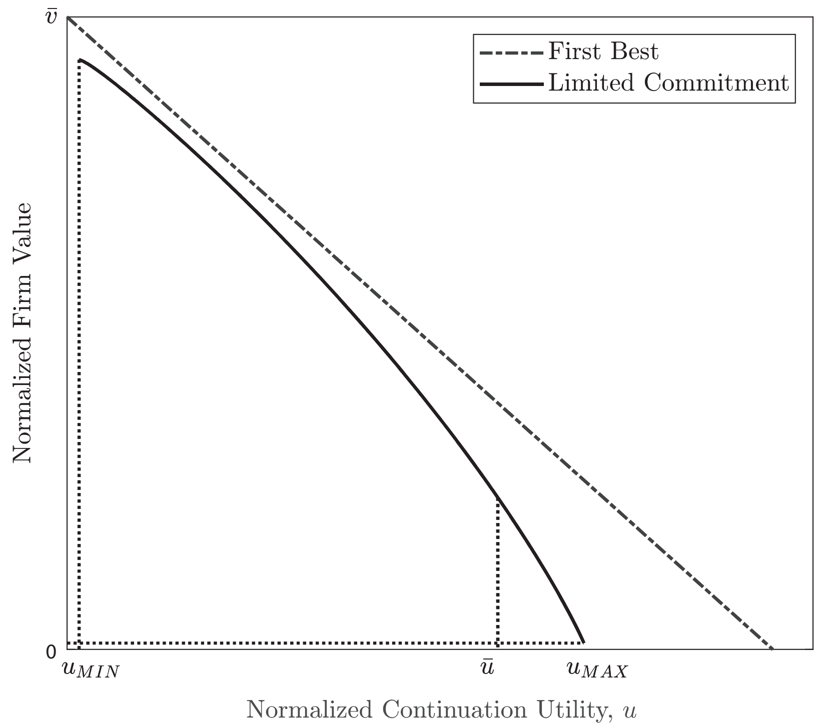


Figure 1. Normalized value function. This figure plots the drift of $\ln u$ (top panel) and the diffusion of $\ln u$ (bottom panel) under the optimal contract.

Third, a firm-manager match starts at $\bar{u}$ and follows a diffusion process implied by the optimal contract afterward. 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,

$$d \ln u_t = \mu_u(u_t)dt + \sigma_u(u_t)dB_t, \tag{26}$$

where the expressions for $\mu_u(u)$ and $\sigma_u(u)$ are given in Section A of the Internet Appendix. Note that $\ln u_t$ is determined by two factors, the drift (dt term) and the diffusion (dB_t term), which intuitively represent the effect of age and size, respectively. The law of motion, $dZ_t = Z_t[(i_t - \delta)dt + \sigma dB_t]$, implies that the cross-sectional variation in firm size is driven by cross-sectional differences in i_t or dB_t. As we show below, quantitatively, the standard deviation of dB_t-shocks is much higher than the heterogeneity in i_t, and thus, the

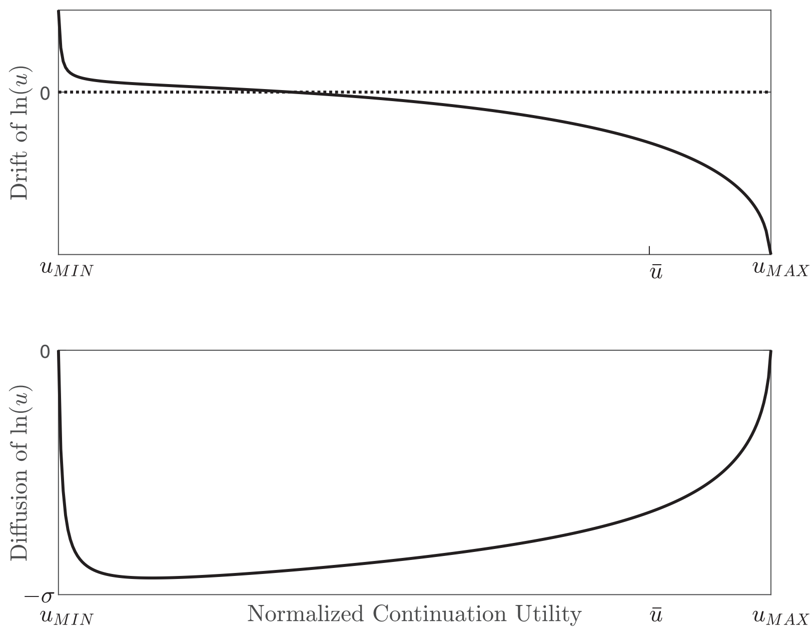


Figure 2. Dynamics of the normalized continuation utility. This figure plots the drift of $\ln u$ (top panel) and the diffusion of $\ln u$ (bottom panel) under the optimal contract.

cross-sectional variation in firm size is governed largely by dB_t shocks. Hence, the dB_t term in equation (26) reflects the effect of size while the drift determines the effect of age on u_t.

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 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, tend to converge toward 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 that $\sigma_u(u) = -\sigma$ because continuation utility remains constant as Z_t varies 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 zero 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}.

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 that managerial compensation 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.

A.2. Dynamics of CEO Compensation

Part 4 of Proposition 3 characterizes 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 leaving the firm 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, $\{l_t^+, l_t^-\}_{t=0}^{\infty}$.

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}.[13] 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 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 shown in Figure 3, in the region in which 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

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

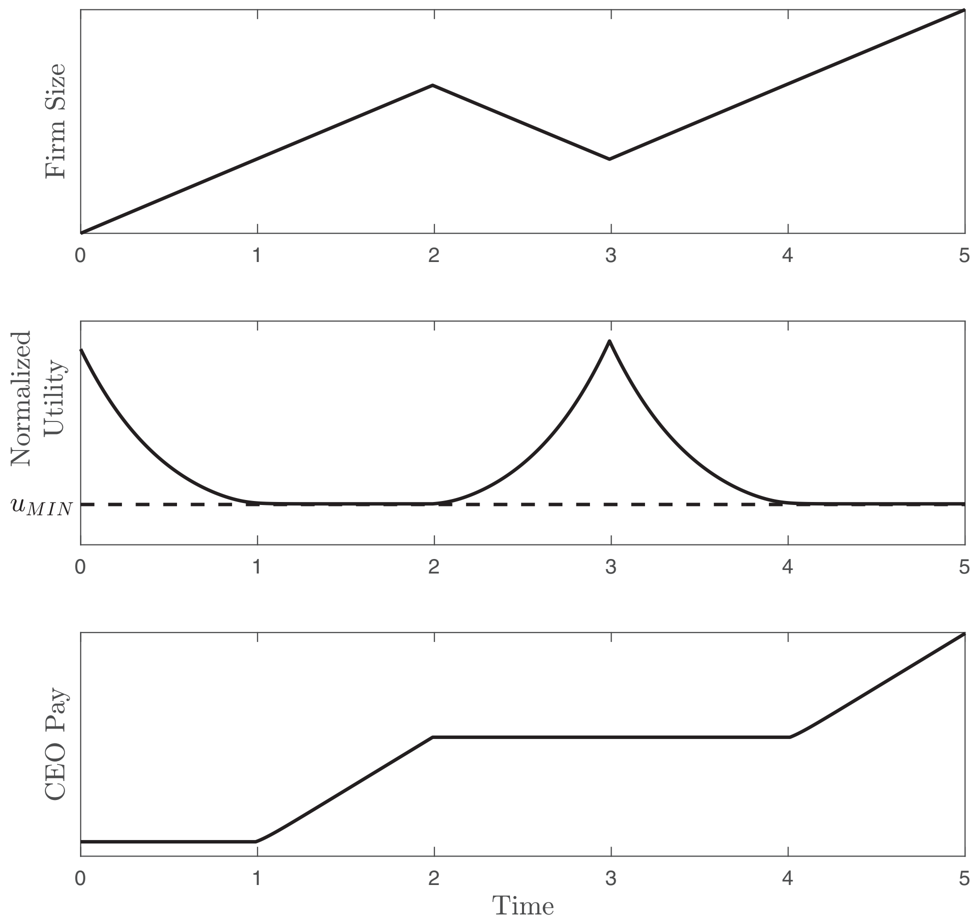


Figure 3. A sample path of CEO compensation. This figure plots a sample path of log firm size (top panel), the implied sample path for the normalized continuation utility (second panel), and the implied sample path for log CEO pay (bottom panel) in the neighborhood of the manager-side limited commitment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

is the key mechanism for the power law of CEO compensation: the running maximum of a power law process follows the same power law. We formalize and generalize this observation in Section III.B below.

A.3. Investment

We plot optimal investment as a function of the normalized promised utility in Figure 4. Consistent with Proposition 3, optimal investment, $i(u)$, increases with the normalized continuation utility. Under the optimal contract, the endogenous correlation between size, age, and normalized utility translates into an endogenous correlation between size, age, and investment. Small

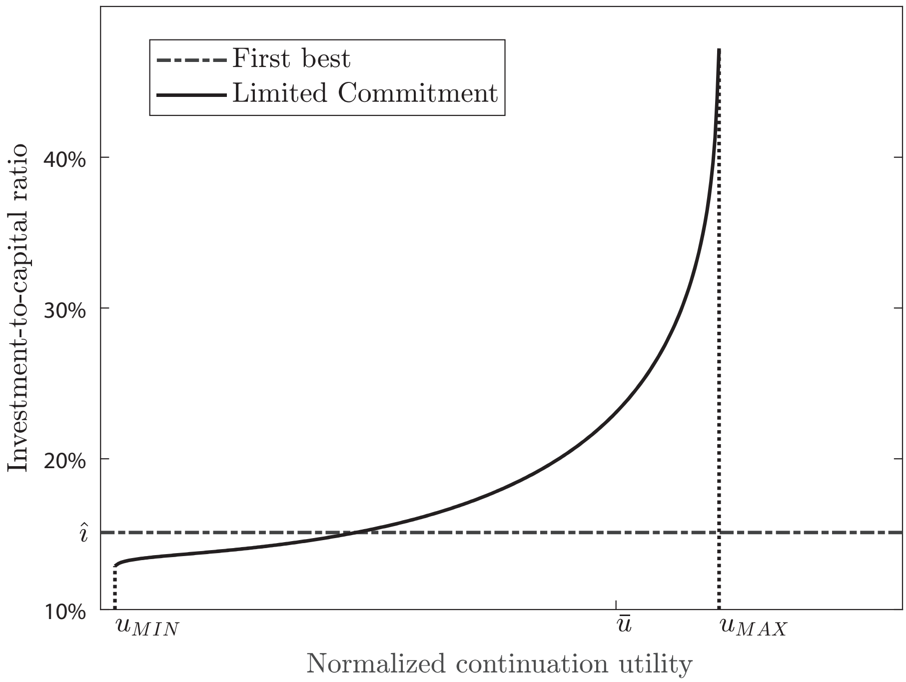


Figure 4. Investment policy. This figure plots the optimal investment policy in the first-best case (dash-dotted line) and that under limited commitment (solid line).

firms that experience a sequence of negative productivity shocks move toward 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. Thus, firm size is negatively correlated with u and, in turn, with firm investment rates. Further, in our model, new firms start at $\bar{u}$, where the investment rate is relatively high, and over time slowly converge to the interior, where investment rates are relatively low. Consequently, older firms invest at a lower level compared with younger firms.

Both the firm-side and the manager-side limited commitment constraints are important in matching the negative relationship between investment and age, as well as that between 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.

Note that a common mechanism that generates the inverse relationship between investment and size is a DRS production function. However, models with a DRS production function generally imply that firms converge to the optimal size in the long run, and therefore, are inconsistent with the empirical evidence of the fat-tailed distribution of firm size.

B. 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 on CEO pay depends on the ratio of the power law slope of firm size, ξ, and the returns to scale of the matching technology, ψ. In general, under DRS, $\psi < 1$, 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 CRS 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 IV. First, we assume that firms and managers experience the exogenous separation shock only once in their lifetime. 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*

$$h\left(i\right) = \left\{ \begin{matrix} i & if\, 0 \le i \le \iota \\ \infty & if\, i > \iota \end{matrix}, \right)$$

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

$$A > r + \kappa + \delta > \iota > \delta, \tag{27}$$

and

$$\frac{A - \iota}{r + \kappa + \delta - \iota} - \frac{\psi\gamma}{(r + \kappa)(\varsigma_1 - 1)}\varsigma_1^{\frac{\psi\gamma - 1}{\psi(\gamma - 1)}}(\varsigma_1 - (1 - \gamma))^{\frac{-\psi\gamma}{\psi(\gamma - 1)}}\bar{\bar{u}}^{\frac{1}{\psi}} \ge 1, \tag{28}$$

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 (IA.8) in Section III.A of the Internet Appendix.

That is, we assume that the marginal cost of investment is equal to 1 if $\frac{I}{Z} \leq \iota$, and is infinite if $\frac{I}{Z} > \iota$. Under Assumption 2, firms' optimal policy is to always invest at the maximum rate ι. Inequality (27) 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 (28) is a restriction on the magnitude of agency frictions. As we show in Section III.A of the Internet Appendix, the outside option of the manager is given by $\bar{u}X^{\psi}$. Inequality (28) 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. In Proposition 4, we formalize the relationship between the power law coefficient on CEO pay and that on firm size. To introduce our notation, for any stochastic process $\{X_t : t \geq 0\}$, let $\check{X}_t = \sup_{0\leq s\leq t} X_s$ denote its running maximum until time t and let $\widehat{X}_t = \inf_{0\leq s\leq t} X_s$ denote its running minimum until time t.

PROPOSITION 4: *(Power Law in CEO Compensation):*

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

$$C_t = \max\left\{\hat{c}\check{Z}_t^{\psi},\ C_0\right\}, \tag{29}$$

   where the constant $\hat{c}$ is defined in equation (IA.11) in Section B of the Internet Appendix. The optimal investment-to-capital ratio is constant: $I_t = \iota Z_t$ for all t.
2. *The right tail of the steady-state distribution of CEO compensation obeys a power law with a slope coefficient of $\frac{\xi}{\psi}$, where ξ is defined in Proposition 2.*

PROOF: See Section III of the Internet Appendix.

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.[14]

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 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 that firm size is large enough we have that the limited commitment constraint for managers binds at least once in the past, then $C_t = \hat{c}\check{Z}_s^{\psi}$ and $D_t = AZ_t - I_t - C_t = AZ_t - \iota Z_t - \hat{c}\check{Z}_t^{\psi}$. Because $\psi \leq 1$, it follows that

$$(A - \iota)Z_t - \hat{c}\check{Z}_t \leq D_t \leq (A - \iota)Z_t.$$

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

IV. Quantitative Results

As discussed above, 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.

A. Data Description and Calibration of Parameters

To calibrate the model and evaluate its quantitative implications, we use a panel of U.S. nonfinancial firms from CRSP and Compustat. We measure executive compensation using total compensation from the 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.[15] For each firm, we collect market capitalization, the number of firm

[14] See also Lustig, Syverson, and Van Nieuwerburgh (2011), Grochulski and Zhang (2011), and Miao and Zhang (2015).

[15] Executive compensation measured in the data does not necessarily correspond to CEO consumption in our model. We follow the related literature and do not make a distinction between

Table I
Calibrated Parameters

This table presents the calibrated parameter values chosen to target the set of moments listed in Table II. The rest of the parameters are calibrated following the literature, in particular, the interest rate $r = 4\%$, the death rate $\kappa_D = 5\%$, the depreciation rate $\delta = 7\%$, and the span of control $\nu = 0.85$.

Description	Notation	Value
Marginal product of capital	A	0.40
Adjustment cost	h_0	5
Volatility	σ	35%
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

employees, the book value of firm assets, the gross value of property, plant and equipment as a proxy for capital, capital expenditure as a proxy for investment, and the amount of common dividends. Firm age is the number of years since the firm's founding date using the Field-Ritter data set (Field and Karpoff (2002), and Loughran and Ritter (2004)).[16] 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 values using the Consumer Price Index compiled by the Bureau of Labor Statistics. The data are sampled at the annual frequency and cover the period from 1992 till 2016.[17]

Tables I and II summarize the calibrated parameter values and the set of moments targeted in calibration. The parameters of our model can be divided into two groups. The first group is fairly standard and can be calibrated by following existing literature. We choose risk aversion of 2. We set the discount rate r to 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 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.[18] We set the span of control

consumption and compensation (e.g., Gabaix and Landier (2008)). For the construction of alternative measures of executive compensation, see Clementi and Cooley (2010).

[16] We thank Jay R. Ritter for making the data publicly available.

[17] We limit our sample to the post-1992 period because coverage of the Execucomp data set begins in 1992.

[18] We choose the depreciation rate similar to the one in Eisfeldt and Papanikolaou (2013), and within the range of existing estimates in the R&D literature. See, for example, Schankerman and Pakes (1986), Lev and Sougiannis (1996), Nadiri and Prucha (1996), Ballister, Mannuel, and Livnat (2003), and Bernstein and Mamuneas (2006).

Table II
Targeted Moments

This table shows the set of moments targeted in calibrating the parameters listed in Table I. We report data statistics and the corresponding moments implied by the calibrated model.

Moments	Data	Model
Average Tobin's Q	1.67	1.45
Median sales growth	11.1%	8.0%
Volatility of sales growth	37.1%	32.1%
CEO pay/Capital of young firms	0.091	0.098
CEO median time between jobs	0.73	0.75
Average CEO departure rate	10%	10%
Average decrease in CEO pay upon separation	40%	44.4%
Relative slope of power law	0.48	0.45

parameter to $\nu = 0.85$ as in Atkeson and Kehoe (2005) and consistent with the estimate in Jovanovic and Rousseau (2001).[19]

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 the 40% decline in CEO income after forced departure reported in Nielsen (2017).[20] In our model, CEO turnover is inefficient because it results in a loss of human capital and CEO income, and hence voluntary departures do not occur. We therefore calibrate λ to match the empirical evidence on forced CEO turnover.[21]

As shown in Section II, the marginal product of capital, A, determines firms' incentives to invest in organization capital and therefore firms' equilibrium growth rate. 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. Due to CRS, the economy is scale invariant. We therefore normalize total labor supply to one and choose $\bar{Z}$ so that the implied $A = 0.40$. This allows our model to match the median sales growth rate of 11.1% in the data.[22] The volatility parameter σ is set at 35% to match the average volatility of firms' sales growth in the data. We set the capital adjustment cost

[19] The span of control parameter is typically calibrated in the 0.6 to 0.95 range. See, for example, Atkeson and Kehoe (2007), Gollin (2008), and Hsieh and Klenow (2009).

[20] In the data and in the model, income decline is measured by the decline in average income in the five years after the turnover event, regardless of whether the CEO has found a new employment opportunity after the turnover.

[21] Forced turnover is defined as departures for which press reports state that the CEO is fired, forced out, or resigns due to policy differences. See Parrino (1997) and Nielsen (2017) for further details.

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

parameter to $h_0 = 5$, so that with $A = 0.40$, our model matches an average Tobin's Q of 1.67 in our sample.[23]

Finally, we calibrate the returns 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 (21)). We choose $\frac{\psi_X}{\psi} = 0.89$ to account for the 0.091 average CEO-pay-to-size ratio of young firms, which are defined as firms of less than five years of age.

Note that the returns to scale parameter ψ measure the elasticity of CEO pay with respect to firm size among large firms and is the key parameter that determines the power law slope of CEO pay relative to that of firm size. Empirically, Baker and Hall (2004) estimate the elasticity of the marginal product of CEO effort with respect to firm size to be approximately 0.4. Although we do not target their moment in calibration, our calibrated value of ψ is remarkably close to the estimate obtained by Baker and Hall (2004).

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

B. Basic Statistics of Firm Dynamics

In this section, we show that our model is able to replicate key stylized features of the joint distribution of firm size, age, investment, CEO compensation, and dividend payout in the data. To this end, we construct a cross section of size sorted portfolios using two measures of firm size: gross capital and the number of firm employees.[24] 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.[25]

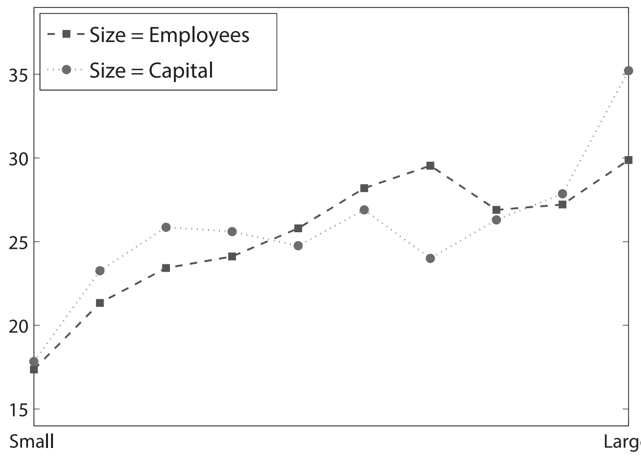
First, consistent with the data, our model implies a monotonic relationship between firm size and age. Figure 5 depicts the observed and the model-implied

[23] We provide details on the calculation of the marginal product of organization capital and Tobin's Q in Section V of the Internet Appendix.

[24] Recall that in the model, both capital and labor are proportional to organization capital, and therefore, are equivalent measures of size. To avoid measurement errors, we rely on the observable measures of firm size. As we show in Section VII of the Internet Appendix, our evidence is robust to instead using the Eisfeldt and Papanikolaou (2013) proxy for organization capital.

[25] We use log-size spaced portfolios to ensure that we separate very large firms from the rest, and hence, to ensure that the distribution of firms across portfolios based on simulated data is consistent with the observed data.

Panel A. Data



Panel B. Model

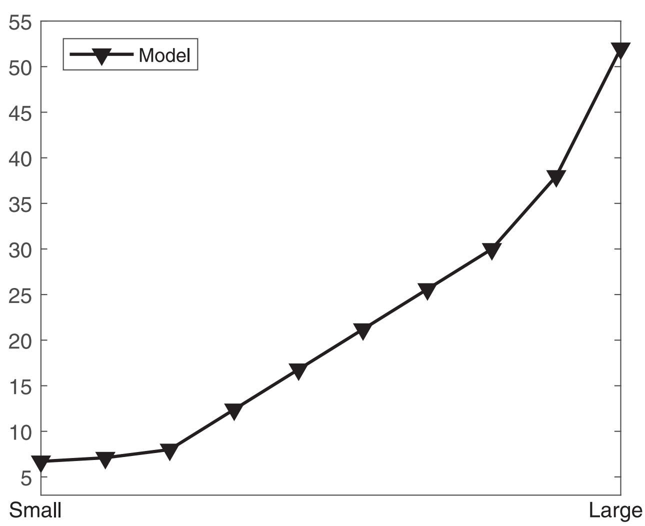


Figure 5. Distribution of age across size. This figure plots the average firm age across 10 size-sorted portfolios. Size in the data is measured by either the number of firm employees or gross capital. Firm age on the vertical axis is measured in years.

variation of the median firm age across size-sorted portfolios (Panels A and B, respectively).

Second, our model is able to quantitatively account for the negative relationship between firm size and investment observed in the data. As explained in

Table III
Investment Rates

This table 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.

	Data		
	Size≡Employees	Size≡Capital	Model
Small	0.153	0.182	0.171
2	0.108	0.140	0.152
3	0.093	0.125	0.128
4	0.084	0.107	0.121
Large	0.092	0.088	0.121
Large−Small	−0.061	−0.094	−0.050
	(−12.32)	(−5.65)	

Section III.A.3, a key implication of our two-sided limited commitment model is the inverse relationship between investment rate and firm size. Large firms are reluctant to invest in organization capital because doing so raises managers' outside options and hence their incentive to leave the firm. In contrast, small firms have an incentive to invest more in order to grow out of the firm-side limited commitment constraint. Table III compares the quantitative implications of the model with the sample statistics. It presents the average investment rates (defined as annual investment divided by the beginning-of-year capital stock) of quintile portfolios sorted by size.[26] 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 around 5%.[27,28] Further, in the data and consistently in the model, younger firms invest at a higher rate than old firms. Table IV presents the estimates from a panel regression of investment rates

[26] Because organization capital is not directly observable, we measure investment rates using available data on physical capital. Note that in our model, organization capital and physical capital are proportional to each other, and thus, feature the same growth rates. In Section VII of the Internet Appendix, we consider an alternative proxy for organization capital and show that our empirical evidence is robust.

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

[28] The inverse relationship between size and investment is also reflected in the cross-sectional variation in firms' dividend payout policies. In the data as well as in the model, large firms are much more likely to pay dividends to shareholders, while small firms tend to retain earnings to fund investments. This evidence is reported in the Internet Appendix.

Table IV

Firm Investment and Age

This table presents the estimates from a panel regression of log investment rates ($\ln(I_{i,t}/K_{i,t-1})$) on the log of firm size measured by capital ($\ln(K_{i,t-1})$) and the log of firm age ($\ln(Age_{i,t-1})$) both in the data and in the model. In the data, we run a panel regression with firm and time fixed effects and a set of controls, and we report the estimates and corresponding t-statistics based on double-clustered standard errors (in parentheses). The set of controls includes lagged values of the investment rate, the change in market capitalization, Tobin's Q, and the cash-flow to assets ratio. The model statistics represent population numbers that are computed using a large panel of simulated data.

Regressors	**Data**	**Model**
$\ln(K_{i,t-1})$	−0.291 (−13.80)	−0.101
$\ln(Age_{i,t-1})$	−0.142 (−4.40)	−0.053

Table V

Elasticity of CEO Compensation to Firm Size

This table 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 market capitalization, the number of firm employees, or gross capital. In Data columns, 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.

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

on lagged size and age and confirms that both size and age contain significant information about the variation in investment rates.[29]

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 V 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

[29] The literature provides mixed evidence on the relationship between size and average growth. For example, Haltiwanger, Jarmin, and Miranda (2013) argue that after controlling for age, the effect of size on growth is not significant. It is straightforward to address this issue in our framework by incorporating fixed operating costs. Even though small firms would still invest at a higher rate (because of limited commitment constraints), due to fixed operating costs, they would be more likely to exit relative to large firms, and hence, the overall effect of size on average growth rates would be ambiguous.

Table VI

CEO Pay to Firm Size Ratio

This table 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.

	Data		
	Size≡Employees	Size≡Capital	Model
Small	0.062	0.076	0.084
2	0.021	0.020	0.070
3	0.014	0.010	0.040
4	0.009	0.005	0.020
Large	0.004	0.001	0.006
Large−Small	−0.058	−0.075	−0.078
	(−20.66)	(−20.31)	

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, as shown in Table VI.

C. 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. We then compare the quantitative implications of our model to the data.

Following Luttmer (2007) and Gabaix (2009), we use the following parametrization of the 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 ζ 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 $-\zeta$.

The literature shows that firm size follows a power-law distribution (e.g., 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 VII. The table also reports sample averages of the corresponding p-values of the Kolmogorov-Smirnov goodness-of-fit test constructed via bootstrap. We provide details on the estimation procedure in Section VI of the Internet Appendix.

Table VII

Estimates of the Power-Law Parameters

This table presents 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-over-year in the 1992 to 2016 sample. Similarly, the reported p-values are time-series averages of year-over-year p-values of the Kolmogorov-Smirnov goodness-of-fit test.

	$\hat{\zeta}$	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

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 estimates obtained using Census data. For example, Luttmer (2007) reports a power law estimate of 1.06; similar estimates are reported in Gabaix and Landier (2008). Overall, the goodness-of-fit test does not reject the power-law null—with just a few exceptions, year-by-year p-values are above the conventional 5% level for all measures of firm size. Notice that the power-law coefficient on dividends is very close to that on firm size, particularly 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 appear 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 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 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 the 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 well matches the tail slopes observed in the data.

As shown in Section III.B, limited commitment on the manager side implies that CEO compensation in large firms is a linear function of the running

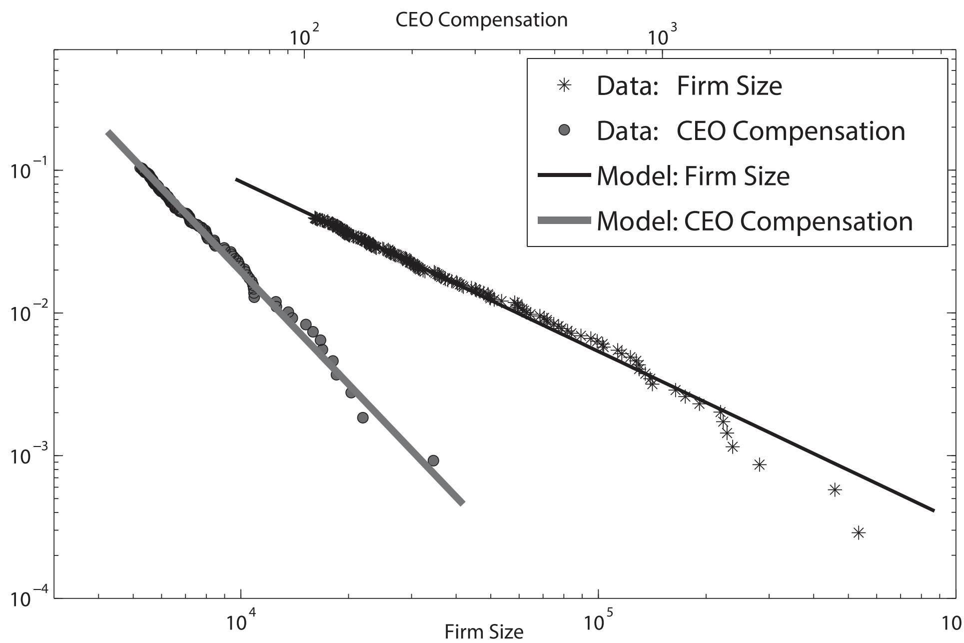


Figure 6. The right tails of firm size and CEO compensation. This figure plots the right tails of the distributions of firm size and CEO compensation (using 2006 data) and the corresponding slopes implied by the model. Firm size in the data is measured by market capitalization. The top and bottom horizontal 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.

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.

Note that our model simultaneously matches the power laws of firm size and CEO pay and the elasticity of CEO pay with respect to firm size, which is challenging for the static assortative matching model of Gabaix and Landier (2008). It is straightforward to show that given a power law of firm size of about 1.1 and elasticity of CEO pay with respect to firm size of one-third, the Gabaix and Landier (2008) model implies a power law of CEO pay of $1.1 \div \frac{1}{3} = 3.3$, which is significantly larger than the sample estimate. Our model is able to reconcile the empirical evidence. In our setting, the elasticity of CEO pay with respect to firm size is large in the tails, where the limited commitment constraints are likely to bind (of about 0.48), and due to risk sharing is much smaller in the interior. The former allows our model to match the tail slopes of both CEO pay and firm size, while the latter ensures that our model is able to match a relatively small average elasticity of CEO pay with respect to firm size of about one-third. The heterogeneous elasticity of CEO pay with respect to firm size is also consistent with the empirical evidence that we present next.

D. Joint Distribution of Investment and CEO Compensation

Our model imposes several restrictions on the joint behavior of CEO pay and firm investment. First, in the cross section, firms with low normalized utility promise a low fraction of future cash flow to managers. These are firms that have a low CEO-pay-to-size ratio, are likely to face a binding constraint on the manager side, and hence invest at a relatively low rate to limit managers' outside options. In contrast, high-u firms feature a high ratio of CEO compensation to firm size, are likely to run into the firm-side limited commitment constraint, and therefore invest at a higher than average rate. Thus, in the cross section, managerial share and investment are positively correlated.

Second, in the time series, our model implies that the response of investment and CEO pay to firm-level shocks is history dependent. As explained in Section III.A, both investment and CEO pay are more likely to respond to positive productivity shocks if they move the firm above its previous running maximum, and are more likely to respond to negative shocks if they make the firm fall below its previous running minimum. In this section, we evaluate the cross-sectional and time-series implications of our model in the data.

Table VIII presents the joint cross-sectional distribution of investment and CEO compensation in the data (Panel A) and in the model (Panel B). It reports average investment rates across 3×3 portfolios sorted first on firm size and then on CEO-pay-to-size ratio. According to our model, controlling for size, managerial share helps identify the tightness of limited commitment constraints. Firms with high CEO-pay-to-size ratio are more likely to encounter a binding constraint on the firm side and therefore invest at a higher rate than low CEO-pay-to-size ratio firms that are likely to run into a binding manager-side constraint. As Table VIII shows, the positive relation between the investment rate and the CEO-pay-to-size ratio implied by the model holds in the data. First, notice that consistent with the evidence discussed above, the investment rate decreases monotonically with firm size. Further, controlling for firm size, firms with high CEO-pay-to-size ratio invest at a significantly higher rate compared with firms that have a relatively low managerial share.[30]

To evaluate the dynamic implications of our model for CEO pay and investment, we test their response to variables that proxy for the history of firm-level productivity shocks. To define notation, let $X^+ = \max\{X, 0\}$ be the positive part of variable X, and $X^- = \min\{X, 0\}$ be the negative part of X. We construct four shocks that proxy for the history of firm-specific innovations: $\Delta^+_{\text{Max}} \ln K_{i,t} \equiv [\ln K_{i,t} - \ln \bar{K}_{i,t-1}]^+$ is the increase in firm size relative to its

[30] In Table III, firm size is measured by the amount of capital. The empirical evidence based on the number of firm employees is similar in terms of magnitude and significance, and therefore, is reported in the Internet Appendix.

Table VIII

Joint Distribution of Investment and CEO Compensation

This table presents the average investment-to-capital ratio across portfolios sorted on firm size and CEO-pay-to-size ratio in the data (Panel A) and in the model (Panel B). Small and large firms represent the bottom and top size-sorted tercile portfolios; low and high firms correspond to the bottom and top tercile portfolios sorted on CEO-pay/size, respectively. Size in both the data and the model is measured by gross capital. In the Data panel, t-statistics for the difference between large and small, and high and low, firms based on the Newey and West (1987) estimator with four lags are reported in parentheses. The model-implied statistics represent population numbers that are computed using a large panel of simulated data.

Panel A: Data

	CEO-pay/Firm-Size			
Firm Size	Low	2	High	High−Low
Small	0.126	0.180	0.288	0.163 (6.10)
2	0.100	0.118	0.156	0.056 (4.51)
Large	0.079	0.104	0.113	0.035 (6.20)
Large−Small	−0.047	−0.076	−0.175	
	(−3.52)	(−4.17)	(−5.18)	

Panel B: Model

	CEO-pay/Firm-Size			
Firm Size	Low	2	High	High−Low
Small	0.149	0.171	0.180	0.030
2	0.129	0.133	0.146	0.017
Large	0.126	0.127	0.128	0.002
Large−Small	−0.023	−0.044	−0.052	

previous running maximum; $\Delta^{-}_{\text{Max}} \ln K_{i,t} \equiv [\ln K_{i,t} - \ln \check{K}_{i,t-1}]^{-}$ is the decrease in firms size relative to its running maximum; $\Delta^{+}_{\text{Min}} \ln K_{i,t} \equiv [\ln K_{i,t} - \ln \widehat{K}_{i,t-1}]^{+}$ is the increase in firm size relative to its previous running minimum; and $\Delta^{-}_{\text{Min}} \ln K_{i,t} \equiv [\ln K_{i,t} - \ln \widehat{K}_{i,t-1}]^{-}$ is the decline in firm size relative to its running minimum.[31] To estimate the joint response of firm investment and CEO compensation to productivity shocks, we run the following panel regression:

$$\begin{pmatrix} \ln i_{i,t} \\ \ln C_{i,t} \end{pmatrix} = \begin{bmatrix} \beta_{i,1} & \beta_{i,2} & \beta_{i,3} & \beta_{i,4} \\ \beta_{c,1} & \beta_{c,2} & \beta_{c,3} & \beta_{c,4} \end{bmatrix} \begin{pmatrix} \Delta^{-}_{\text{Min}} \ln K_{i,t-1} \\ \Delta^{+}_{\text{Min}} \ln K_{i,t-1} \\ \Delta^{-}_{\text{Max}} \ln K_{i,t-1} \\ \Delta^{+}_{\text{Max}} \ln K_{i,t-1} \end{pmatrix} + \alpha' X_{i,t-1} + \begin{pmatrix} \epsilon^{ik}_{i,t} \\ \epsilon^{c}_{i,t} \end{pmatrix}, \quad (30)$$

where $\ln i_{i,t} \equiv \ln \frac{I_{i,t}}{K_{i,t-1}}$ is the log investment-to-capital ratio, $\ln C_{i,t}$ is the time-t log-level of CEO compensation at firm i, and $X_{i,t-1}$ is a set of controls.

[31] Note that the signs of shocks are preserved, that is, $\Delta^{+}_{\text{Max}} \geq 0$ and $\Delta^{+}_{\text{Min}} \geq 0$, and $\Delta^{-}_{\text{Max}} \leq 0$ and $\Delta^{-}_{\text{Min}} \leq 0$.

Table IX

Joint Dynamics of Investment and CEO Compensation

This table presents the joint dynamics of investment and CEO compensation in both the data and the model. Panel A shows the estimates from a panel regression of the log investment rate ($\ln i_{i,t} \equiv \ln \frac{I_{i,t}}{K_{i,t-1}}$) on four shocks that correspond to opposite changes in firm size relative to its running minimum and maximum. $\Delta^{-}_{Min} \ln K_{i,t-1}$ and $\Delta^{-}_{Max} \ln K_{i,t-1}$ represent declines, and $\Delta^{+}_{Min} \ln K_{i,t-1}$ and $\Delta^{+}_{Max} \ln K_{i,t-1}$ represent increases, in firm size at time $t-1$ relative to the minimum and maximum observed over the previous three years. In Panel B, we consider a similar regression for the log-level of CEO compensation ($\ln C_{i,t}$). In the data, we run a panel regression with firm and time fixed effects and a set of controls. We report the estimates and the corresponding t-statistics based on double-clustered standard errors (in parentheses). In Panel A, we control for the lagged values of gross capital, the investment rate, the change in market capitalization, Tobin's Q, and the cash-flow to assets ratio; in Panel B, we control for lagged capital and lagged CEO compensation. The model statistics represent population numbers computed using a large panel of simulated data.

	Panel A: $Y = \ln i_{i,t}$		Panel B: $Y = \ln C_{i,t}$	
Regressors	Data	Model	Data	Model
$\Delta^{-}_{Min} \ln K_{i,t-1}$	−0.458 (−6.53)	−0.125	0.078 (2.49)	0.023
$\Delta^{+}_{Min} \ln K_{i,t-1}$	−0.032 (−2.26)	−0.001	0.066 (3.07)	0.001
$\Delta^{-}_{Max} \ln K_{i,t-1}$	−0.018 (−0.61)	−0.074	−0.012 (−0.52)	0.014
$\Delta^{+}_{Max} \ln K_{i,t-1}$	−0.123 (−3.62)	−0.164	0.089 (2.24)	0.022

According to our model, when firm size exceeds its previous running maximum, the manager-side limited commitment constraint is likely to bind. It is therefore optimal to increase CEO pay and reduce investment. In contrast, when size falls below is previous running minimum, firms simultaneously reduce CEO compensation and accelerate investment to avoid the binding constraint on the firm side. Thus, our model predicts that both CEO pay and the investment rate respond to $\Delta^{+}_{\mathrm{Max}} \ln K_{i,t}$ and $\Delta^{-}_{\mathrm{Min}} \ln K_{i,t}$, but not to $\Delta^{-}_{\mathrm{Max}} \ln K_{i,t}$ and $\Delta^{+}_{\mathrm{Min}} \ln K_{i,t}$. That is, under the null of the model, $\beta_{i,j} < 0$, $\beta_{c,j} > 0$, for $j = 1, 4$, and $\beta_{i,j},\ \beta_{c,j} \approx 0$ for $j = 2, 3$. In words, the investment rate features a ∩-shaped response, and CEO compensation features a U-shaped response, to the four shocks in equation (30).[32]

Table IX presents the estimated elasticities, βs, both in the data and in the model. The regression coefficients are estimated using annual data, except we use quarterly series to obtain more accurate measures of running minimum and maximum firm sizes which we measure by the minimum and maximum of firm capital over the three prior years.[33] In the investment regression (Panel A), the set of controls includes lagged values of firm capital, the investment rate, the change in market capitalization, Tobin's Q, and the cash-flow to

[32] Note that under CRS matching technology, $\beta_{i,j} = \beta_{c,j} = 0$ for $j = 2, 3$; in our calibrated model with DRS, these elasticities may not be zero, yet investment and CEO pay continue to feature a pronounced ∩- and U-shape response, respectively.

[33] Our evidence is robust to expanding the horizon over which we measure running minimum and maximum.

assets ratio; in the CEO-pay regression (Panel B), we control for lagged capital and lagged CEO compensation. In both specifications, we also control for firm and time fixed effects and cluster standard errors by firm and time to ensure robustness of our inference to the cross-sectional dependence and serial correlation in residuals. The model-based statistics represent population numbers that are computed using a large panel of simulated data.[34]

Notice that as expected, in the model, the responses of the firm investment rate and CEO pay to productivity shocks are sizable only when they bring the firm above its running maximum or below its running minimum. Specifically, $\beta_{i,1}$ and $\beta_{i,4}$ are large and negative, $\beta_{c,1}$ and $\beta_{c,4}$ are positive, and the remaining coefficients are close to zero.[35] That is, in the model, investment and CEO compensation respond significantly to shocks that move firms toward limited commitment constraints where they respond little to productivity shocks that move firms away from the constraints.

Consistent with the model, investment and CEO compensation in the data are history dependent—firms that have recently experienced negative productivity shocks tend to reduce managerial compensation and increase investment, while firms that have recently experienced high-growth innovations tend to increase managerial compensation and reduce investment. In particular, on average, CEO pay increases by a significant 8.9% and declines by a significant 7.8% among firms that overperform their previous best and underperform their previous worst, respectively. Simultaneously, the latter increase their investment rates by almost a half and the former reduce their investment rates by about 12%, on average. Empirically, the shock elasticities of the investment rate and CEO compensation have pronounced ∩- and U-shapes, respectively, as predicted by our model.[36]

V. 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.

[34] In the model-based regressions, we control for the lagged value of firm size, and in the CEO-pay regression, we also control for the lagged value of CEO compensation as required by the model's dynamics.

[35] Note that the sample estimate of $\beta_{c,2}$ is significant, which suggests that in the data, CEO pay also responds to positive shocks that push firms above their running minimums. Incorporating other frictions in our model framework (such as moral hazard) would likely help account for this evidence.

[36] The data estimates are somewhat larger than their model counterparts because, in the data, both investment and CEO compensation are likely to respond to productivity shocks due to agency frictions above and beyond limited commitment.

Our model provides a conceptual framework that can be used to study various policy questions and several extensions may provide promising directions for future research. For example, although regulations are typically unnecessary in optimal contracting models because allocations are constrained efficient, a fully optimal policy analysis within our model requires a mechanism design approach and a deeper microfoundation for some of the incentive compatibility constraints.[37] Further, if the degree of contract enforcement is a choice variable, then improving contract enforcement may both lead to a more efficient allocation and reduce income inequality among managers. More generally, our model provides a framework to study the implication of the tradeoff between contract enforcement and income inequality for firm dynamics and economic growth. Also, in our model, outside options of firms and managers are microfounded by equilibria in a directed matching market. The interaction between the matching market equilibrium and optimal contracting represents a promising direction of the optimal policy analysis.

From a quantitative perspective, relaxing several assumptions may allow our model to account for a richer set of characteristics of CEO pay and firm dynamics. For example, we assume that both CEO compensation and investment in organization capital are publicly observable and can be contracted upon. In practice, however, investment in organization capital may not be fully observable or contractible. Relaxing these assumptions allows one to study the quantitative relevance of a broad set of agency frictions. Further, for tractability, the specification of human and organization capital dynamics in our model is quite stylized—they follow independent processes when separated and perfectly correlated processes once matched to a corporation. A more general and perhaps more plausible specification would allow managers' human capital and firms' organization capital to be positively but not perfectly correlated and may provide further insights into managerial compensation and firm dynamics.

Finally, our model predicts that the impact of agency frictions depends on the returns to scale of the matching technology. Our preliminary industry analysis discussed in Section VIII of the Internet Appendix suggests that consistent with the model, the strength of the inverse relationship between firm size and investment varies significantly with the returns to scale of the matching technology. Further research in this area may help us better understand how and why CEO compensation and investment policies vary across different industries.

Initial submission: July 30, 2016; Accepted: November 21, 2019
Editors: Stefan Nagel, Philip Bond, Amit Seru, and Wei Xiong

[37] For example, firm side limited commitment in equation (11) involves equilibrium prices and gives room for the existence of pecuniary externalities. An earlier version of our paper shows that this constraint can be replaced by a coalition rationality constraint. Under that interpretation, the equilibrium is constrained efficient.

REFERENCES

Ai, Hengjie, and Anmol Bhandari, 2015, Asset pricing with endogenously uninsurable tail risk, *Econometrica*, forthcoming.

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, 775–797.

Arellano, Cristina, Yan Bai, and Jing Zhang, 2012, Firm dynamics and financial development, *Journal of Monetary Economics* 59, 533–549.

Atkeson, Andrew, and Patrick Kehoe, 2005, Modeling and measuring organization capital, *Journal of Political Economy* 113, 1026–1053.

Atkeson, Andrew, and Patrick Kehoe, 2007, Modeling the transition to a new economy: Lessons from two technological revolutions, *American Economic Review* 97, 64–88.

Atkeson, Andrew, and Robert Lucas, 1992, On efficient distribution with private information, *Review of Economic Studies* 59, 427–453.

Axelson, Ulf, and Philip Bond, 2015, Wall street occupations, *Journal of Finance* 70, 1949–1996.

Axtell, Robert L., 2001, Zipf distribution of U.S. firm sizes, *Science* 293, 1818–1820.

Baker, George, and Brian Hall, 2004, Ceo incentives and firm size, *Journal of Labor Economics* 22, 767–798.

Ballister, Marta, Garcia-Ayuso Mannuel, and Joshua Livnat, 2003, The economic value of the R&D intangible asset, *European Accounting Review* 21, 605–633.

Berk, Jonathan, Richard Stanton, and Josef Zechner, 2010, Human capital, bankruptcy, and capital structure, *Journal of Finance* 65, 891–926.

Bernstein, Jeffery, and Theofanis Mamuneas, 2006, R&D depreciation, stocks, user costs and productivity growth for US R&D intensive industries, *Structural Change and Economic Dynamics* 17, 70–98.

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, 2019, Optimal contracting, corporate finance, and valuation with inalienable human capital, *Journal of Finance* 74, 1363–1429.

Clementi, Gian, and Hugo Hopenhayn, 2006, A theory of financing constraints and firm dynamics, *Quarterly Journal of Economics* 121, 229–265.

Clementi, Gian Luca, and Thomas Cooley, 2010, Executive compensation: Facts, Working Paper, NYU.

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* 54, 1232–1287.

Edmans, Alex, Xavier Gabaix, Tomas Sadzik, and Yuliy Sannikov, 2012, Dynamic ceo compensation, *Journal of Finance* 67, 1603–1647.

Eisfeldt, Andrea, and Dimitris Papanikolaou, 2013, Organization capital and the cross-section of expected returns, *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, 1–8.

Edward Fee, C., 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.

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

Field, Laura C., and Jonathan Karpoff, 2002, Takeover defenses of ipo firms, *Journal of Finance* 57, 1857–1889.

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.

Gollin, Douglas, 2008, Nobody's business but my own: Self-employment and small enterprise in economic development? *Journal of Monetary Economics* 55, 219–233.

Gomez, Matthieu, 2016, Asset prices and wealth inequality, Working paper, Columbia University.

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.

Haltiwanger, Jophn, Ron Jarmin, and Javier Miranda, 2013, Who creates jobs? Small versus large versus young, *Review of Economics and Statistics* 95, 347–361.

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.

Hsieh, Chang-Tai, and Peter J. Klenow, 2009, Misallocation and manufacturing tfp in China and India, *Quarterly Journal of Economics* 124, 1403–1448.

Jovanovic, Boyan, and Peter Rousseau, 2001, Vintage organization capital, NBER working paper, No. 8166.

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, and Edward Prescott, 1982, Time to build and aggregate fluctuations, *Econometrica* 50, 1345–1370.

Lev, Baruch, and Theodore Sougiannis, 1996, The capitalization, amortization, and value-relevance of R&D, *Journal of Accounting and Economics* 21, 107–138.

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.

Loughran, Tim, and Jay R. Ritter, 2004, Why has ipo underpricing changed over time? *Financial Management* 33, 5–37.

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, 2015, A duality approach to continuous-time contracting problems with limited commitment, *Journal of Economic Theory* 159, 929–988.

Nadiri, Ishaq, and Ingmar Prucha, 1996, Estimation of the depreciation rate of physical and R&D capital in the U.S. total manufacturing sector, *Economic Inquiry* 34, 43–56.

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.

Parrino, Robert, 1997, Ceo turnover and outside succession a cross-sectional analysis, *Journal of Financial Economics* 46, 165–197.

Prescott, Edward, and Michael Visscher, 1980, Organization capital, *Journal of Political Economy* 88, 446–461.

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.

Schankerman, Mark, and Ariel Pakes, 1986, Estimates of the value of patent rights in european countries during the post-1950 period, *Economic Journal* 96, 1052–1076.

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, *Review of Economic Studies* 55, 541–554.

Supporting Information

Additional Supporting Information may be found in the online version of this article at the publisher's website:

Internet Appendix.
Replication Code.