Appendix: Firm Dynamics under Limited Commitment

A The First-Best Case

A.1 Optimal Solution for the First-Best Case

Proof of Proposition 1: Homogeneity implies that the value function of the optimization problem (17) is of the form $\bar{v}K$. In this case, the HJB equation that describes the value function can be simplified to:

$$(\mathbf{r}+\kappa)\,\bar{v}=\max_{i}\left\{\mathbf{A}-h(i)+(i-\delta)\,\bar{v}\right\}. \tag{A.1}$$

Because $h(i)=i+\frac{1}{2}h_0 i^2$, the first-order condition for the above maximiation problem, $h'(i)=\bar{v}$ implies

$$\bar{v}=1+h_0\hat{\imath}. \tag{A.2}$$

Together, (A.1) and (A.2) imply that $\hat{\imath}$ must satisfy

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

where we denote $\hat{r}=\mathbf{r}+\kappa+\delta$ as in Proposition 1. The relevant solution is

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

Note that under Assumption 2, $\hat{\imath}$ defined above satisfies

$$\hat{\imath}=\arg\max_{i<\hat{r}}\frac{\mathbf{A}-h(i)}{\hat{r}-i},$$

as needed.

A.2 Power Law of Firm Size in the First-Best Case

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

$$dx_t=\mu dt+\sigma dB_t.$$

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

Lemma A.1. *Let $\theta_1>0$ and $\theta_2<0$ denote the two roots of the following quadratic equation:*

$$\kappa+\mu\theta-\frac{1}{2}\sigma^2\theta^2=0. \tag{A.3}$$

1. *The stationary distribution is given by:*

$$m\left(x|\,x_0\right)=\begin{cases}\frac{1}{\sqrt{\mu^2+2\kappa\sigma^2}}e^{\theta_2(x-x_0)} & x\geq x_0\\ \frac{1}{\sqrt{\mu^2+2\kappa\sigma^2}}e^{\theta_1(x-x_0)} & x<x_0\end{cases}$$

2. *Assume* $\kappa>\mu+\frac{1}{2}\sigma^2$*, then* $\theta_2<-1$*, and*

$$\int_{-\infty}^{\infty}e^x m\left(x|\,x_0\right)dy=\frac{e^{x_0}}{\kappa-\left(\mu+\frac{1}{2}\sigma^2\right)}<\infty$$

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

$$M\left(X|\,x_0\right)=\begin{cases}\frac{1}{\sqrt{\mu^2+2\kappa\sigma^2}}e^{-\theta_2 x_0}X^{\theta_2-1} & X\geq e^{x_0}\\ \frac{1}{\sqrt{\mu^2+2\kappa\sigma^2}}e^{-\theta_1 x_0}X^{\theta_1-1} & X<e^{x_0}\end{cases}$$

   In particular, the right tail of X *obeys power law with slope* θ_2*.*

Proof. See Luttmer [4]. □

Proof of proposition 2: By proposition 1, $\log K$ is a Brownian motion:

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

with initial condition $\log\bar{K}$. Proposition 2 can be proved by applying Lemma A.1 directly.

B One-Sided Limited Commitment

B.1 Managers' Outside Options

We first derive an expression for managers' outside options upon default. In Specification 1, the manager is allowed to hire labor to produce market goods although he is excluded from intertemporal risk sharing contracts upon default. Because market goods can be used for investment, upon default, managers maximize their life-time utility,

$$\max_{\{C_{j,t},I_{j,t}\}_{t=0}^{\infty}}\left\{E\left[\int_0^{\tau}(\beta+\kappa)\,e^{-\beta t}C_{j,t}^{1-\gamma}dt\right]\right\}^{\frac{1}{1-\gamma}}\tag{B.1}$$

by choosing optimally consumption and investment decisions subject to constraints (4) and (7) with $D_{j,s}=0$ all $s>0$.

In Specification 2, upon default, managers can only use home labor to produce home consumption goods. In this case, he maximizes life-time utility (B.1) with $C_{j,t}\leq K_{j,t}^{\alpha}\left(\mathbf{z}\bar{n}\right)^{1-\alpha}$, where $K_{j,t}^{\alpha}(\mathbf{z}\bar{n})^{1-\alpha}$ is the total production of home consumption goods at time t. Because all home consumption goods must be consumed and cannot be used for investment, the law of motion of capital in this case is given by:

$$dK_{j,t} = K_{j,t}\left[-\delta dt + \sigma dB_{j,t}\right].$$

We provide the expressions for managers' outside options in Specifications 1 and 2 in the following lemma.

Lemma A.2. *Manager's outside option upon default is given by ϖK^{ν}, where in Specification 1, $\varpi = \varpi_1$ and $\nu = 1$, and in Specification 2, $\varpi = \varpi_2$, and $\nu = \alpha$. The constant ϖ_1 is given by:*

$$\varpi_1 = \left\{\left[\mathbf{A} - \widetilde{i} - \frac{1}{2}h_0\widetilde{i}^2)\right]^{-\gamma}\left(1 + h_0\widetilde{i}\right)\right\}^{\frac{1}{1-\gamma}}, \tag{B.2}$$

where the constant h_0 is the curvature parameter of the adjustment cost function ($h(i) = i + \frac{1}{2}h_0 i^2$), and $\widetilde{i}$ is the unique solution to the following equation:

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

with $\hat{\beta} = \beta + \kappa + \delta(1-\gamma) + \frac{1}{2}\gamma(1-\gamma)\sigma^2$. The constant ϖ_2 is:

$$\varpi_2 = \left[\frac{\beta + \kappa}{(\beta+\kappa) + \delta\alpha(1-\gamma) - \frac{1}{2}\alpha(1-\gamma)(\alpha(1-\gamma) - 1)\sigma^2}\right]^{\frac{1}{1-\gamma}}(\mathbf{z}\bar{n})^{(1-\alpha)}. \tag{B.3}$$

Proof. See Technical Appendix. □

B.2 Optimality of Compensation and Investment Policies

We provide the details of the proof the optimality of proposed compensation and investment policies in part 1 of Proposition 3 in Section A of the technical appendix, where we show that under the optimal contract, managerial compensation must satisfy $\frac{C_t}{K_t^{\nu}} \geq \hat{c}$ for all t, where

$$\hat{c} \equiv \varpi\left(\frac{\zeta_1 - (1-\gamma)}{1-\gamma}\right)^{\frac{1}{1-\gamma}}, \tag{B.4}$$

where ζ_1 is given by equation (A.13) in the technical appendix. In addition, and the limited-commitment constraint binds if and only if $\frac{C_t}{K_t^{\nu}} \geq \hat{c}$ holds with equality. Here, we derive the power law for CEO pay taking the optimal compensation policy (26) and the optimal investment policy, $\frac{I_t}{K_t} = \hat{i}$ as given. Note that under this optimal investment policy, firm size follows a geometric Brown motion with a constant drift. As a result, in the right tail of the firm size distribution, optimal CEO compensation behaves like the running maximum of a geometric Brown motion. The following two lemmas characterize the right tail of the running maximum of geometric Brownian motions and prepare us for the proof of the power law of CEO compensation.

B.3 Distribution of Running Maximums of Geometric Brown Motions

The first lemma computes the integral of discounted normal density. The proof of the lemma can be found in the technical appendix.

Lemma A.3. *Assume $\kappa > 0$, $\nu > 0$ and $y \geq x_0$. Let θ_2 be the negative root of the quadratic equation (A.3) defined in Lemma A.1. Then*

$$\int_0^\infty e^{-\kappa t}\Phi_0\left(\frac{y-\nu\mu t-x_0}{\nu\sigma\sqrt{t}}\right)dt = \frac{1}{\kappa}+\frac{\nu}{\theta_2\sqrt{\mu^2+2\kappa\sigma^2}}e^{\frac{\theta_2}{\nu}(y-x_0)},$$

where $\Phi_0(\cdot)$ is the cumulative distribution function of the standard normal distribution.

The second lemma characterizes the distribution of the right tail of the running maximum of Brownian motions. Continue to consider the setup of Lemma A.1. For all j, let $\{x_{j,s}\}_{s=0}^{\infty}$ be a Brownian motion starts at x_0 as in Lemma A.1. Define the running maximum of the Brownian motions as

$$\hat{x}_{j,t} = \sup_{0<s<t} x_{j,s},$$

and let $Y_{j,t}$ be

$$Y_{j,t} = \max\{y_0, \nu_0 + \nu\hat{x}_{j,t}\},$$

where y_0 is a constant. We continue and use m to denote the stationary distribution of the particles $\{x_j\}$ which evaporate at Poisson rate κ. The following lemma characterizes the distribution of the right tail of Y.

Lemma A.4. *Assume $\mu > 0$. For y large enough,*

$$m(Y_j > y) \sim -\frac{\nu}{\theta_2\sqrt{\nu^2\mu^2+2\kappa\nu^2\sigma^2}}e^{\frac{\theta_2}{\nu}[y-(\nu_0+\nu x_0)]},$$

where θ_2 is the negative root of the quadratic equation (A.3) defined in Lemma A.1.

Proof. Note that for $y > y_0$, $Y_{j,t} = \max\{y_0, \nu_0 + \nu\hat{x}_{j,t}\} > y$ is equivalent to $\nu_0 + \nu\hat{x}_{j,t} > y$, or $\hat{x}_{j,t} > \frac{1}{\nu}(y-\nu_0)$. Using equation (9.4) on page 15 of Harrison [3], for any x,

$$P(\hat{x}_{j,t} < x) = \Phi_0\left(\frac{x-x_0-\mu t}{\sigma\sqrt{t}}\right) - e^{-\frac{2\mu(x-x_0)}{\sigma^2}}\Phi_0\left(\frac{-x+x_0-\mu t}{\sigma\sqrt{t}}\right),$$

where Φ_0 is the cumulative distribution function of the standard normal distribution. Therefore,

$$P\left(\hat{x}_{j,t} > \frac{1}{\nu}(y-\nu_0)\right) = 1-\Phi_0\left(\frac{\frac{1}{\nu}(y-\nu_0)-x_0-\mu t}{\sigma\sqrt{t}}\right)+e^{-\frac{2\mu\left(\frac{1}{\nu}(y-\nu_0)-x_0\right)}{\sigma^2}}\Phi_0\left(\frac{-\frac{1}{\nu}(y-\nu_0)+x_0-\mu t}{\sigma\sqrt{t}}\right), \tag{B.5}$$

Note that the second normal cdf can be written as

$$\begin{aligned}\Phi_0\left(\frac{-\frac{1}{\nu}(y-\nu_0)+x_0-\mu t}{\sigma\sqrt{t}}\right) &= \Phi_0\left(\frac{-y+\nu_0+\nu x_0-\nu\mu t}{\nu\sigma\sqrt{t}}\right) \\ &= 1-\Phi_0\left(\frac{y+\nu\mu t-(\nu_0+\nu x_0)}{\nu\sigma\sqrt{t}}\right).\end{aligned}$$

As a result, (B.5) can be written as:

$$\begin{aligned}P\left(\hat{x}_{j,t} > \frac{1}{\nu}(y-\nu_0)\right) &= 1-\Phi_0\left(\frac{y-\nu\mu t-(\nu_0+\nu x_0)}{\nu\sigma\sqrt{t}}\right) \\ &\quad +e^{-\frac{2\mu(y-\nu_0-\nu x_0)}{\nu\sigma^2}}\left[1-\Phi_0\left(\frac{y+\nu\mu t-(\nu_0+\nu x_0)}{\nu\sigma\sqrt{t}}\right)\right]. \end{aligned} \tag{B.6}$$

Because particles evaporate at rate κ, the law of large numbers implies that the total measure of particles that satisfies $\hat{x}_{j,t} > \frac{1}{\nu}(y-\nu_0)$ is given by:

$$m(Y_j > y) = \int_0^\infty e^{-\kappa t} P\left(\hat{x}_{j,t} > \frac{1}{\nu}(y-\nu_0)\right) dt.$$

Using (B.6),

$$\begin{aligned}m(Y_j > y) &= \frac{1}{\kappa} - \int_0^\infty e^{-\kappa t}\Phi_0\left(\frac{y-\nu\mu t-(\nu_0+\nu x_0)}{\nu\sigma\sqrt{t}}\right) dt \\ &\quad + e^{-\frac{2\mu(y-\nu_0-\nu x_0)}{\nu\sigma^2}}\left[\frac{1}{\kappa} - \int_0^\infty e^{-\kappa t}\Phi_0\left(\frac{y+\nu\mu t-(\nu_0+\nu x_0)}{\nu\sigma\sqrt{t}}\right) dt\right]. \end{aligned} \tag{B.7}$$

Using Lemma A.3, for $y > (\nu_0+\nu x_0)$,

$$\int_0^\infty e^{-\kappa t}\Phi_0\left(\frac{y-\nu\mu t-(\nu_0+\nu x_0)}{\nu\sigma\sqrt{t}}\right) dt = \frac{1}{\kappa} + \frac{\nu}{\theta_2\sqrt{\nu^2\mu^2+2\kappa\nu^2\sigma^2}} e^{\frac{\theta_2}{\nu}[y-(\nu_0+\nu x_0)]},$$

where θ_2 is the negative root of the quadratic equation (A.3). Similarly,

$$\int_0^\infty e^{-\kappa t}\Phi_0\left(\frac{y+\nu\mu t-(\nu_0+\nu x_0)}{\nu\sigma\sqrt{t}}\right) dt = \frac{1}{\kappa} + \frac{\nu}{\tilde{\theta}_2\sqrt{\nu^2\mu^2+2\kappa\nu^2\sigma^2}} e^{\frac{\theta_2}{\nu}[y-(\nu_0+\nu x_0)]}, \tag{B.8}$$

where $\tilde{\theta}_2$ is the negative root of the quadratic equation:

$$\kappa - \mu\tilde{\theta} - \frac{1}{2}\sigma^2\tilde{\theta} = 0.$$

Note that $\tilde{\theta}_2 = -\theta_1$, where θ_1 is the positive root of (A.3). Therefore, equation (B.8) can be written as:

$$\int_0^\infty e^{-\kappa t}\Phi_0\left(\frac{y+\nu\mu t-(\nu_0+\nu x_0)}{\nu\sigma\sqrt{t}}\right)dt = \frac{1}{\kappa} - \frac{\nu}{\theta_1\sqrt{\nu^2\mu^2+2\kappa\nu^2\sigma^2}}e^{-\frac{\theta_1}{\nu}[y-(\nu_0+\nu x_0)]}. \tag{B.9}$$

Summarize (B.7), (B.8) and (B.9), we have:

$$m(Y_j > y) = -\frac{\nu}{\theta_2\sqrt{\nu^2\mu^2+2\kappa\nu^2\sigma^2}}e^{\frac{\theta_2}{\nu}[y-(\nu_0+\nu x_0)]} + \frac{\nu}{\theta_1\sqrt{\nu^2\mu^2+2\kappa\nu^2\sigma^2}}e^{-\frac{2\mu+\theta_1\sigma^2}{\nu\sigma^2}[y-(\nu_0+\nu x_0)]}$$

It is straight forward to show that under the assumption $\mu > 0$, $\frac{\theta_2}{\nu} > -\frac{2\mu+\theta_1\sigma^2}{\nu\sigma^2}$. Therefore the first term dominates and determines the tail behavior of $m(Y_j > y)$ for large y. This proves the lemma. □

B.4 Proof of Proposition 3

The optimal compensation policy in equation (26) implies

$$\ln C_{j,.t} = \max\{\ln c(u_{MIN}) + \nu\ln K_{j,t},\ \ln C_0\}.$$

Under the optimal investment policy, the law of motion of $K_{j,t}$ is given by:

$$d\ln K_{j,t} = (\hat{\imath}-\delta)\,dt + \sigma dB_{j,t},$$

where the Brownian motion $B_{j,t}$ is independent across j. Under Assumption 2, $\hat{\imath}-\delta > 0$. Using the result of Lemma A.4 above, for y large,

$$m(\ln C_{j,.t} > y) \sim -\frac{\nu}{\xi\sqrt{\nu^2(\hat{\imath}-\delta)^2+2\kappa\nu^2\sigma^2}}e^{\frac{\xi}{\nu}[y-(\ln c(u_{MIN})+\nu\ln C_0)]}.$$

Clearly, the right tail of C obeys a power law with tail slope $\frac{\xi}{\nu}$.

C Two-Sided Limited Commitment

C.1 Law of Motion of Continuation Utility

In the main text, we defined continuation utility as in (6). Here we show that (6) is equivalent to the recursive formulation (12) and an integrability condition. Because it is more convenient to work with the additively separable (expected utility) representation of the same preference, we define

$$\hat{U}_t = \frac{1}{1-\gamma}U_t^{1-\gamma} = E_t\left[\int_t^\tau e^{-\beta(s-t)}(\beta+\kappa)\frac{1}{1-\gamma}C_s^{1-\gamma}ds\right].$$

Note that for $t < \tau$

$$\int_0^t e^{-\beta s} (\beta + \kappa) \frac{C_s^{1-\gamma}}{1-\gamma} ds + e^{-\beta t} \hat{U}_t = E_t \left[\int_0^\tau e^{-\beta s} (\beta + \kappa) \frac{C_s^{1-\gamma}}{1-\gamma} ds \right] \tag{C.1}$$

is a martingale. Therefore, by the *Martingale Representation Theorem*, (C.1) has the representation

$$\int_0^t e^{-\beta s} (\beta + \kappa) \frac{C_s^{1-\gamma}}{1-\gamma} ds + e^{-\beta t} \hat{U}_t = \int_0^t e^{-\beta s} G\left(K_s, \hat{U}_s\right) dB_s + \int_0^t e^{-\beta s} G_D\left(K_s, \hat{U}_s\right) \left(dN_s^D - \kappa ds\right) \tag{C.2}$$

for some adapted processes $\left\{G_B\left(K_t, \hat{U}_t\right)\right\}$ and $\left\{G_D\left(K_t, \hat{U}_t\right)\right\}$ such that

$$E_0 \left[\int_0^\tau \left(e^{-\beta t} G_B\left(K_t, \hat{U}_t\right)\right)^2 dt \right] \text{ and } E_0 \left[\int_0^\tau \left(e^{-\beta t} G_D\left(K_t, \hat{U}_t\right)\right)^2 dt \right] < \infty.$$

In equation (C.2), $\left\{N_t^D\right\}_{t \geq 0}$ represent the poisson process (with intensity κ) of managers' health shocks. Because manager receives a terminal utility of zero and exit the economy whenever the health shock hits, $G_D(K_t, \hat{U}_t) = \hat{U}_t$ for all t. Therefore, conditioning on not being hit by a health shock, (C.2) can be written as:

$$d\hat{U}_t = (\beta + \kappa) \left[\hat{U}_t - \frac{C_t^{1-\gamma}}{1-\gamma} \right] dt + G_B\left(K_t, \hat{U}_t\right) dB_t.$$

Since $U_t = \left((1-\gamma)\hat{U}_t\right)^{\frac{1}{1-\gamma}}$, if we define $G(K, U) = U^\gamma G_B\left(K, \hat{U}\right)$ Itô's lemma implies (12).

C.2 Proof of Proposition 4

For any pair of state variables (K, U) such that $u_{MIN}(K) < \frac{U}{K} < u_{MAX}(K)$, the limited commitment constraints do not bind and the value function $V(U, K)$ must satisfy the following HJB equation in the interior.

$$rV(K, U) = \max_{C,i,g} \left\{ \begin{array}{c} \mathbf{A}K - h(i)K - C + V_K(U,K)K[i - \delta] + \frac{1}{2}V_{KK}(U,K)K^2\sigma^2 \\ +V_U(U,K)U\left[\frac{\mathbf{r}}{1-\gamma}\left(1 - \left(\frac{C}{U}\right)^{1-\gamma}\right) + \frac{1}{2}\gamma g^2\sigma^2\right] \\ +\frac{1}{2}V_{UU}(U,K)U^2 g^2\sigma^2 + V_{K,U}(U,K)KUg\sigma^2 \end{array} \right\}. \tag{C.3}$$

Taking first-order condition with respect to C, we have

$$rU^\gamma V_U(K, U) = \frac{1}{C^{-\gamma}}.$$

To prove that compensation must be constant, it is enough to establish

$$d\left[U^\gamma V_U(K, U)\right] = 0,$$

which follows from the proof of Lemma 5 in Ai and Li [2]. This proves part 2) of Proposition 4.

Next, to derive the appropriate boundary conditions, it is more convenient to use $\ln\left(\frac{U}{K}\right)$ and $\ln K$ as state variables. To simplify notation, we denote

$$\mathbf{k} = \ln K \text{ and } \mathbf{u} = \ln U - \ln K,$$

and work the normalized value function: $v(\mathbf{k}, \mathbf{u}) = e^{-\mathbf{k}} V\left(e^{\mathbf{k}}, e^{\mathbf{u}+\mathbf{k}}\right)$. Then (C.3) becomes

$$0 = \max_{c,i,g} \left\{ \begin{array}{c} [\mathbf{A} - c - h(i)] - (\mathbf{r} + \kappa + \delta - i) v(\mathbf{k}, \mathbf{u}) \\ + v_{\mathbf{k}}(\mathbf{k}, \mathbf{u}) \left[i - \delta + \frac{1}{2}\sigma^2\right] + \frac{1}{2} v_{\mathbf{kk}}(\mathbf{k}, \mathbf{u}) \sigma^2 \\ + v_{\mathbf{u}}(\mathbf{k}, \mathbf{u}) \left[\frac{\beta+\kappa}{1-\gamma}\left(1 - \left(\frac{c}{\exp(\mathbf{u})}\right)^{1-\gamma}\right) - (i - \delta) + \frac{1}{2}(\gamma - 1) g^2\sigma^2 + g\sigma^2 - \frac{1}{2}\sigma^2\right] \\ + \frac{1}{2} v_{\mathbf{uu}}(\mathbf{k}, \mathbf{u}) (g-1)^2 \sigma^2 + v_{\mathbf{ku}}(\mathbf{k}, \mathbf{u}) (g - 1) \sigma^2 \end{array} \right\}, \tag{C.4}$$

The first-order conditions implies the following optimal choices of c, i and g (recall that $h(i) = i + \frac{1}{2}h_0 i^2$):

$$\begin{aligned} c^* &= \mathbf{u}\left(-(\beta+\kappa) v_{\mathbf{u}}\right)^{\frac{1}{\gamma}}, \\ i^* &= \frac{1}{h_0}\left(v + v_{\mathbf{k}} - \mathbf{u} v_{\mathbf{u}} - 1\right), \\ g^* &= -\frac{v_{\mathbf{u}} + v_{\mathbf{ku}} - v_{\mathbf{uu}}}{(\gamma - 1) v_{\mathbf{u}} + v_{\mathbf{uu}}}. \end{aligned} \tag{C.5}$$

Therefore, the HJB equation (C.6) becomes the following PDE.

$$0 = \left\{ \begin{array}{c} \left[\mathbf{A} - \mathbf{u}\left(-(\beta+\kappa)v_{\mathbf{u}}\right)^{\frac{1}{\gamma}} - h\left(\frac{1}{h_0}\left(v + v_{\mathbf{k}} - \mathbf{u}v_{\mathbf{u}} - 1\right)\right)\right] \\ -\left(\mathbf{r} + \kappa + \delta - \frac{1}{h_0}\left(v + v_{\mathbf{k}} - \mathbf{u}v_{\mathbf{u}} - 1\right)\right) v + v_{\mathbf{k}}\left[\frac{1}{h_0}\left(v + v_{\mathbf{k}} - \mathbf{u}v_{\mathbf{u}} - 1\right) - \delta + \frac{1}{2}\sigma^2\right] \\ + v_{\mathbf{u}}\left[\frac{\beta+\kappa}{1-\gamma}\left(1 - \frac{\mathbf{u}(-(\beta+\kappa)v_{\mathbf{u}})^{\frac{1-\gamma}{\gamma}}}{\exp((1-\gamma)\mathbf{u})}\right) - \left(\frac{1}{h_0}\left(v + v_{\mathbf{k}} - \mathbf{u}v_{\mathbf{u}} - 1\right) - \delta\right)\right. \\ \left. + \frac{1}{2}\left(\gamma\left(-\frac{v_{\mathbf{u}} + v_{\mathbf{ku}} - v_{\mathbf{uu}}}{(\gamma-1)v_{\mathbf{u}} + v_{\mathbf{uu}}}\right)^2 - \left(-\frac{\gamma v_{\mathbf{u}} + v_{\mathbf{ku}}}{(\gamma-1)v_{\mathbf{u}} + v_{\mathbf{uu}}}\right)^2\right)\sigma^2\right] \\ + \frac{1}{2}v_{\mathbf{kk}}\sigma^2 + \frac{1}{2}v_{\mathbf{uu}}\left(-\frac{\gamma v_{\mathbf{u}} + v_{\mathbf{ku}}}{(\gamma-1)v_{\mathbf{u}} + v_{\mathbf{uu}}}\right)^2\sigma^2 + v_{\mathbf{ku}}\left(-\frac{\gamma v_{\mathbf{u}} + v_{\mathbf{ku}}}{(\gamma-1)v_{\mathbf{u}} + v_{\mathbf{uu}}}\right)\sigma^2 \end{array} \right\}, \tag{C.6}$$

With a little abuse of notations, we denote the policy functions characterizing the optimal contract by $c(\mathbf{k}, \mathbf{u})$, $i(\mathbf{k}, \mathbf{u})$ and $g(\mathbf{k}, \mathbf{u})$. Then under the optimal contract

$$\begin{aligned} d\mathbf{k} &= \left(i(\mathbf{k}, \mathbf{u}) - \delta - \frac{1}{2}\sigma^2\right) dt + \sigma dB \\ d\mathbf{u} &= \left[\frac{\beta+\kappa}{1-\gamma}\left(1 - \left(\frac{c(\mathbf{k}, \mathbf{u})}{\exp(\mathbf{u})}\right)^{1-\gamma}\right) - (i(\mathbf{k}, \mathbf{u}) - \delta) + \frac{1}{2}\left(\gamma g(\mathbf{k}, \mathbf{u})^2 - (g(\mathbf{k}, \mathbf{u}) - 1)^2\right)\sigma^2\right] dt \\ &\quad + (g(\mathbf{k}, \mathbf{u}) - 1)\sigma dB. \end{aligned}$$

To derive the relevant boundary conditions, note that the limited commitment constraint on the manager side implies $U \geq \varpi K^{\nu}$, which is equivalent to $\mathbf{u} + (1-\nu)\,\mathbf{k} \geq \ln \varpi$. Therefore, on the boundary where $\mathbf{u} + (1-\nu)\,\mathbf{k} = \ln \varpi$, we must have $d\left[\mathbf{u} + (1-\nu)\,\mathbf{k}\right] \geq 0$ with probability one. A necessary condition for this is that the diffusion component of $d\left[\mathbf{u} + (1-\nu)\,\mathbf{k}\right]$ vanishes, which requires $(g(\mathbf{k},\mathbf{u}) - 1)\,\sigma + (\nu - 1)\,\sigma = 0$, or $g(\mathbf{k},\mathbf{u}) = -\nu$ for all $(\mathbf{k},\mathbf{u})$ such that $\mathbf{u} + (1-\nu)\,\mathbf{k} = \ln \varpi$. Using (C.5), we obtain a boundary condition for $v(\mathbf{k},\mathbf{u})$ on $\mathbf{u} + (1-\nu)\,\mathbf{k} = \ln \varpi$ (or $\frac{U}{K} = u_{MIN}(K)$):

$$\frac{v_{\mathbf{u}} + v_{\mathbf{ku}} - v_{\mathbf{uu}}}{(\gamma - 1)\,v_{\mathbf{u}} + v_{\mathbf{uu}}} = \nu.$$

Similarly, the limited-commitment constraint on shareholder side implies that $v(\mathbf{k},\mathbf{u}) \geq 0$. Therefore, on the boundary $v(\mathbf{k},\mathbf{u}) = 0$, we must have $dv(\mathbf{k},\mathbf{u}) \geq 0$ with probability one. Again, a necessary condition for this to happen is that the diffusion component on $dv(\mathbf{k},\mathbf{u})$ vanishes on the boundary. That is, $v_{\mathbf{k}}(\mathbf{k},\mathbf{u}) + v_{\mathbf{u}}(\mathbf{k},\mathbf{u})\,(g(\mathbf{k},\mathbf{u}) - 1) = 0$, or $g(\mathbf{k},\mathbf{u}) = 1 - \frac{v_{\mathbf{k}}}{v_{\mathbf{u}}}$. Using (C.5), we obtain the following boundary condition on $v(\mathbf{k},\mathbf{u}) = 0$ (or $\mathbf{u} = \ln u_{MAX}(K)$):

$$-\frac{v_{\mathbf{u}} + v_{\mathbf{ku}} - v_{\mathbf{uu}}}{(\gamma - 1)\,v_{\mathbf{u}} + v_{\mathbf{uu}}} = 1 - \frac{v_{\mathbf{k}}}{v_{\mathbf{u}}}.$$

Finally, we refer the reader to Proposition 1 and Proposition 3 in Ai and Li [2], which can be adapted to prove part 3) of the proposition.

D Aggregation and the Summary Measure m

In this section, we describe a procedure to solve the cross-sectional distribution of firm characteristics. We refer the reader to Ai [1] for the technical details and proofs. Because firm types can be summarized by (K,u), the cross-section distribution of firm characteristic can be equivalently characterized by a distribution on the (K,u) space. We denote this measure by $\tilde{\Phi}(K,u)$. Using $\tilde{\Phi}(K,u)$, the resource constraints (14) and (15) can be written as:

$$C + \int \left[c(u)\,K + h(i(u))\,K\right] d\tilde{\Phi}(K,u) + H\left(\bar{K}\right) = \int K^{\alpha}\,(zn(u)\,K)^{1-\alpha}\,\tilde{\Phi}(K,u)\,, \tag{D.1}$$

and

$$\int n(u)\,K d\tilde{\Phi}(K,u) = 1. \tag{D.2}$$

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

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

Using the summary measure, $m(\cdot)$, the resource constraints (D.1) and (D.2) can be written as:

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

and

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

respectively. Given Proposition 4, the equilibrium can be completely characterized by one-dimensional policy functions, $\{c(u), i(u), n(u)\}$ and the one dimension measure, $m(u)$.

In addition, as shown in Ai [1], $m(u)$ must satisfy the following ODE:

$$[\kappa + \delta - i(u)]\, m(u) = \frac{1}{2}\sigma^2 \frac{d^2}{d(\log u)^2}\left\{m(u)[g(u)-1]^2\right\} + \frac{d}{d\log u}\{b(u)\, m(u)\}, \tag{D.3}$$

where

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

E Limited Commitment on Shareholder Side

In this section, we provide the details of the optimal contracting problem for the case of limited commitment on the shareholder side discussed in Section 4. The key properties of the policy functions are summarized in Proposition 3 below, the proof of which can be adapted from Ai and Li [2].

Proposition A.1. *One-Sided Limited Commitment*

1. *There exists a $u_{MAX} > 0$ such that under the optimal contract, $0 < u_t \le u_{MAX}$ for all t, and $v(u) \ge 0$ for all $u \in (0, u_{MAX}]$. In addition, the limited commitment in equation constraint (9) binds if and only if $u_t = u_{MAX}$.*

2. *The normalized value function, $v(u)$ solves the following ODE*

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

   with the boundary condition $v(u_{MAX}) = 0$, $v''(u_{MAX}) = \infty$, and $\lim_{u\to 0} v(u) = \bar{v}$.

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

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

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

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

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

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

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

Third, due to limited commitment on the shareholder side, firms' investment rate is decreasing in size, which is qualitatively consistent with the data. As stated in Proposition A.1, investment is increasing in u. Intuitively, as capital stock gets depleted and the normalized utility goes up, the firm gets closer to the boundary u_{MAX}, which involves imperfect risk sharing and is welfare reducing. To avoid hitting the constraint, the firm must increase investment to re-build its capital stock and move away from the distress region. Note that due to risk sharing, promised utility is less sensitive to productivity shocks than firm size. Therefore, under the optimal contract, firm size is negatively correlated with normalized utility. Hence, under limited commitment on the shareholder side, small firms invest more and feature higher growth rates compared with large firms.

F Data Description

The cross-sectional data that we use consists of US non-financial firms and come from the Center for Research in Securities Prices (CRSP) and Compustat. We measure executive compensation by the total compensation figure from ExecuComp database, which comprises of salary, bonuses, the value of restricted stock granted, the Black-Scholes-based value of options granted and long-term incentive payouts. For each firm we collect market capitalization, the number of firm employees, the book value of firm assets, the gross value of property, plant and equipment to measure capital, capital expenditure to measure investment, and the amount of common dividends. We measure firm age by the number of years since the firm's founding date. Our evidence remain virtually intact if we use the date of incorporation to define firm age. Firm exit rates are computed using Compustat deletion series that account for acquisitions and mergers, bankruptcy, liquidation, reverse acquisition and leverage buyout. All nominal quantities are converted to real using the consumer price index compiled by the Bureau of Labor Statistics. The data are sampled on the annual frequency and cover the period from 1992 till 2011.

G Power-Law Estimates

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

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

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

$$\tilde{\zeta} = N\left[\sum_{i=1}^{N} \log \frac{x_i}{\tilde{x}}\right]^{-1}, \quad x_i \geq \tilde{x},\ i = 1, ..., N. \tag{G.2}$$

Our estimates of x_{min} and ζ is the pair that yields the power-law distribution that provides the best fit to the observed data according to the KS criteria, i.e.,

$$\{\hat{x}_{min}, \hat{\zeta}\} = \min_{\tilde{x},\tilde{\zeta}} \Big\{ KS\!-\!distance \Big\} \equiv \min_{\tilde{x},\tilde{\zeta}} \Big\{ \max_{x \geq \tilde{x}} \left| F(x; \tilde{x}, \tilde{\zeta}) - \hat{F}(x) \right| \Big\}, \tag{G.3}$$

where $F(x; \tilde{x}, \tilde{\zeta})$ is the candidate power-law cumulative distribution function and $\hat{F}(x)$ is the empirical distribution.

References

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

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

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

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