# ESSAYS ON ASSET PRICING AND FIRM DYNAMICS

**A THESIS**
**SUBMITTED TO THE FACULTY OF THE GRADUATE SCHOOL**
**OF THE UNIVERSITY OF MINNESOTA**
**BY**

JINCHENG TONG

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

HENGJIE AI

MARCH, 2020

Acknowledgements

I am deeply indebted to my advisor, Hengjie Ai, for his continuous support and invaluable guidance. Special thanks go to Anmol Bhandari for continuing help and advice on my thesis.

This thesis has benefited a lot from numerous helpful coommnets of mmebers of my thesis committee, Frederico Belo, Robert Goldstein, for their helpful suggestions and support.

I also appreciate the comments and suggestions from Philip Bond, Jaroslav Borovicka, Georgy Chabakauri, Riccardo Colacito, Andrea Eisfeldt, Murray Frank, Paul Ehling, Urban Jermann, Paymon Khorrami, Leonid Kogan, Jun Li, Xiaoji Lin, Erik Loualiche, Ellen Mcgrattan, Chi-Yang Tsou, Colin Ward and seminar participants at the Midwest Finance Association Annual Meeting, University of Minnesota, AFA poster session, FMA PhD Consortium, Chicago Booth Asset Pricing Conference, University of Iowa, Indiana University Bloomington, University of North Carolina Chapel Hill, University of Toronto, Goethe University of Frankfurt Asset Pricing Conference and City University of London Corporate Policy and Asset Pricing conference for their helpful comments.

I am very grateful to my parents for the unconditional love and support of my education.

Last but not the least, this work would not have been possible without constant encouragement and support from my wife Jing Wang.

## Abstract


My dissertation studies the implications of agency frictions for asset prices and firm dynamics. The first chapter embed an optimal contracting framework into an otherwise standard asset pricing paradigm to resolve some of the challenges in investment and asset pricing literature. The second chapter investigates the effect of agency frictions on CEO compensation, firm size and investment dynamics.

In Chapter one "A Dynamic Agency Based Asset Pricing Model with Production", coauthored with Chao Ying, I develop a general equilibrium model based on dynamic agency theory to study investment and asset prices. In our environment, neither firms nor workers can commit to compensation contracts that provide continuation values below their outside options. At the aggregate level, the presence of agency frictions amplifies the market price of risks and allows our model to generate a sizable equity premium with a low level of risk aversion. History dependent labor contracts generate a form of operating leverage and allow our model to match the key features of the aggregate and cross-section of investment and equity returns in the data. A variance decomposition of investment into discount rate news and cash flow news supports the mechanism of our model.

In Chapter two, "Firm Dynamics under Limited Commitment", coauthored with Hengjie Ai, Dana Kiku and Rui Li, we present a general equilibrium model with two-sided limited commitment that helps account for the observed heterogeneity in firms' investment, payout and CEO-compensation policies. In the model, shareholders cannot commit to holding negative net present value projects, and managers cannot commit to compensation plans that yield life-time utility lower than their outside options. Firms operate identical constant return to scale technologies with i.i.d. productivity growth. Our model endogenously generates power laws in firm size and CEO compensation and explains the differences in their empirical distributions. We also show that the model is able to quantitatively account for the salient features of firms' growth dynamics, the observed negative relationship between firms' investment rate and size, and the positive relationship between firms' size and their dividend and CEO payout.

# Contents

# List of Tables

# List of Figures

# Contents

# Chapter 1

# A Dynamic Agency Based Asset Pricing Model with Production

Several aspects of investment and asset returns in the data pose challenges to equilibrium asset pricing models with production. At the aggregate level, general equilibrium models typically have difficulty in simultaneously explaining the high equity premium and the considerable volatility of investment in the data. In the cross-section, it is still challenging to generate significant differences in firms' expected stock returns as well as substantial dispersion in their investments at the same time.

This paper is an attempt to resolve these challenges. We present a general equilibrium model with heterogeneous firms, making optimal investment and labor compensation decisions subject to agency frictions. The model provides a framework to reconcile several key facts about asset prices and to account for many features in firms' economic activities and macroeconomic aggregates. It generates a large equity premium as well as a low and smooth return on the risk-free asset. In the aggregate time series, the model produces substantial time variation in the market price of risk, together with a volatile aggregate investment and smooth aggregate consumption. The cross-section of firms exhibits a sizable value premium and more importantly, a substantial dispersion in corporate investment. The model is also consistent with the fact that investment negatively predicts stock returns at both the aggregate and firm level.

We embed an optimal contracting problem into a heterogeneous agent model with two types of agents, firm owners and workers. An owner meets a worker in a matching market and they form a firm to produce output. Firms' output is subject to both aggregate and firm-specific productivity shocks. Firm owners are well diversified and provide risk sharing compensation contracts to insure workers against both types of shocks. The key agency friction is limited commitment on the labor compensation contract. Both parties, firms and workers, have the option to renege on contracts and pursue their outside options if their outside values are higher than the continuation values provided in the contracts. We solve for the optimal contract under agency frictions within a general equilibrium production environment and examine its implications for quantities and asset prices. We obtain several key results that contribute to the quantitative success of the model.

Endogenous uninsurable risks in labor compensation amplify the equilibrium market price of risk. Adverse shocks to workers' compensations are not fully insured due to owner-side lack of commitment. With recursive utility and persistence in aggregate state, temporarily unconstrained workers are concerned about future uninsurable idiosyncratic states. Optimal risk sharing requires firm owners to deliver a larger fraction of total output as aggregate labor compensations in recessions than in booms. A countercyclical aggregate labor compensation share translates into a procyclical consumption share of firm owners, raising the unconditional volatility of the pricing kernel as well as generating endogenously countercyclical variation in the equilibrium market price of risk.

Countercyclical risk premia are important in jointly matching high stock market excess returns and large variation in aggregate investment. We assume that capital is produced in the capital goods sector and the supply elasticity of capital is determined by the convex adjustment cost in this sector. Low supply elasticity of capital associated with high adjustment cost parameter value leads to large variation in capital prices but smooth aggregate investment in standard models. In our model, however, countercyclical risk premia amplify the shifts in aggregate demand for capital following aggregate shocks. Suppose the economy is hit by an adverse aggregate shock. Firms' demand for capital is reduced because persistent adverse shock lowers the expected marginal benefit of investment. In addition, the demand is further decreased because of the increase in firms' cost of capital that results from higher risk premia. This feature allows us to simultaneously account for a high return on the stock market and significant time series variation in aggregate capital investment.

The interaction between risk premia and aggregate investment also implies that aggregate investment negatively predicts future stock market excess returns. Persistent variations in the equilibrium risk prices that make returns predictable come from the persistence of firm owners' consumption share. In the data, we confirm the findings in previous literature that aggregate investment rate is negatively associated with stock market excess returns in the future and this predictive power increases over forecasting horizons. This pattern is also exhibited in our model and both the slope coefficients and $R^2$ line up with the data relatively well at all horizons.

The optimal contract gives rise to a form of labor-induced operating leverage at the firm level. Firms insure workers against aggregate shocks, thus rendering the residual dividend claim more exposed to aggregate risks. Consider firms with recent histories of adverse shocks. The operating leverage effect is stronger for them because they have higher committed value to workers relative to their output due to the risk sharing nature of the contract. Hence, these firms are characterized by lower equity valuations, higher Book to Market (BM) ratios , and higher expected stock returns. Hence our model generates a sizable value premium quantitatively[1] . In addition, labor compensations are less sensitive to productivity shocks than output, negative productivity shocks are followed by decreases in investment and increases in labor operating leverage: low investments are followed by higher expected returns going forward. Our model quantitatively captures the inverse relations between firm-level investment and expected return in the data.

Finally, to illustrate our model mechanism and to contrast it with the existing literature, we develop an empirical procedure to understand what drives variations of investment at both the aggregate time series and in the cross-section of firms. In dynamic models of investment and asset pricing, investment is typically forward-looking and should respond to news on both future cash flow and discount rate. We follow

[1] This pattern of firms' BM ratios, idiosyncratic productivities, labor operating leverages, and expected returns is consistent with empirical findings. Imrohoroglu and Tuzel (2014) document the relationship between productivity and book to market ratio prior to the construction of value strategy portfolios. Donangelo et al. (2018) finds that the labor operating leverage channel is an important determinant for the value premium and labor leverage explains approximately 50% of the value premium.

Campbell and Shiller (1988) as well as Vuolteenaho (2002) and apply the variance decomposition technique to attribute time series variation of aggregate investment and the cross-sectional variation of firms' investment to the contributions of cash flow news and discount rate news. We find that at the aggregate level, discount rate news explains all the movement in investment, while the contribution of cash flow component is negligible. On the contrary, the majority of the cross-sectional difference in firms' investment comes from unexpected news to cash flow component, while the heterogeneity in firm-level discount rate news adds little. We implement the same variance decomposition methods on the simulated panel data generated by our model and we find that our model is consistent with both findings.

Our empirical procedure is also useful for distinguishing investment models based on different mechanisms, because different mechanisms have completely different implications for the major source of variations in investment. As an example, we apply our decomposition method to analyze the type of news that drives the cross-sectional variation of firm investments in a benchmark asset pricing model with capital adjustment cost and one aggregate shock, following Zhang (2005). We find that discount rate news explains most of the cross-sectional variation in firm investments, which is inconsistent with our empirical findings. Our finding adds another caveat to the class of one factor models: the calibrated high investment adjustment cost renders firms' cash flows not dispersed enough, so that most of the variation in firm investment is explained by dispersion of news related to firms stock returns.

## Related literature

Our analysis contributes to several strands of literature.

Our paper is first related to the literature on asset pricing with exogenously incomplete markets. Mankiw (1986) and Constantinides and Duffie (1996) show that countercyclical volatility in labor income raises aggregate risk prices in general equilibrium. Constantinides and Ghosh (2015) and Schmidt (2015) demonstrate that tail-risks in labor income amplify the volatility of the pricing kernel. For reasons of tractability, Constantinides and Duffie (1996) and follow-up papers typically assume that individuals face permanent income shocks, which eliminate the motives to smooth such shocks; thus individuals choose not to trade. Krueger and Lustig (2010) analyzes theoretical underpinnings for the relevance of incomplete markets to asset prices.

Second, our paper builds on the literature that studies asset pricing implications of endogenously incomplete market models. The idea that endogenously incomplete risk sharing due to agency frictions amplifies the market price of risk dates back to Alvarez and Jermann (2001) and Chien and Lustig (2010). Their asset pricing models build on the Kehoe and Levine (1993) and Alvarez and Jermann (2000) framework, which develops a theory of an endogenous incomplete market. We emphasize the role of double sided limited commitment to understand price and quantity dynamics.

The interaction between the two-sided limited commitment[2] and asset prices is the main emphasis of a recent work by Ai and Bhandari (2018). While all the models mentioned so far are set up in endowment economies, we build on the important insights of them and explore the broader implications for asset prices and macroeconomic aggregates in a production economy in which quantities and prices are jointly and endogenously determined. In addition, the heterogeneous firm setup enable us to examine not only aggregate quantities and prices, but also the cross-section of firm investments and expected returns.

Our paper is also connected to a large body of literature on asset pricing in production economies, which was recently surveyed by Kogan and Papanikolaou (2012).[3] Our work differs from this literature in two significant aspects. First, the literature typically assumes a complete market and hence the existence of a representative agent. In our model, the financial market is endogenously incomplete due to agency frictions, which amplifies the conditional volatility of the pricing kernel. Second, capital adjustment cost or other frictions in investment are the key ingredients that generate variations in the price of capital. However, strong adjustment costs lead to either implausibly smooth time series of aggregate investment, or a counterfactually high volatility of the risk-free interest rate. Our model produces a volatile time series of aggregate investment, a low

[2] The two-sided lack of commitment setup can be connected to previous works such as Lustig et al. (2011), Ai and Li (2015), Bolton et al. (2016) and Ai et al. (2018). These papers typically study similar contracting problem in a single firm or without aggregate uncertainty. We investigate the implications of two-sided limited commitment in an environment with heterogeneous firms and aggregate uncertainty.

[3] This literature aims to provide a unified framework that combines the success of the neoclassical RBC models on the quantity side with the success of asset pricing mechanisms typically derived in endowment economies (Jermann (1998), Boldrin et al. (2001), Chen (2017), Kaltenbrunner and Lochstoer (2010), Croce (2014), Kung and Schmid (2015), Corhay et al. (2017), Gourio (2012), Papanikolaou (2011))

variability of the risk-free interest rate, and a large variation of stock market excess returns.

Starting with Berk et al. (1999), Gomes et al. (2003), Carlson et al. (2005) and Zhang (2005), researchers have been investigating the capabilities of investment-based asset pricing models to explain several puzzling features of stock returns in the cross-section of firms. [4] Recent work by Clementi and Palazzo (2018) shows that models in this literature are confronted with the challenge in simultaneously explaining a large cross-section variation of investment and sizable value premium. They typically rely on capital adjustment cost or investment irreversibility to generate heterogeneity in expected returns such as the value premium. But adjustment cost lead to smooth firm-level investment. In contrast, our model does not feature firm-level adjustment cost to create dispersion in the cross-section of stock returns. The value premium results from the amplification of equilibrium risk prices and endogenous labor operating leverage, both due to agency frictions. The labor leverage process is calibrated to match the level and variation of firm-level labor share. As a result, our model is not subject to the Clementi and Palazzo (2018) critique because it can account for firm-level investment dynamics and the value premium at the same time.

To numerically solve the GE model with heterogeneous firms, we build on the solution methods developed in Krusell and Smith (1997) and Krusell and Smith (1998). From the optimal contracting perspective, we use the conceptual framework of Atkeson and Lucas (1992) and Atkeson and Lucas (1995), where promised utility[5] is a state variable in individual firm's decision problems and its distribution is an aggregate state variable. We adopt and extend the procedure in Krusell and Smith (1997) to solve for equilibrium prices with multiple market clearing conditions. Similar procedures have

[4] A large body of literature on investment based asset pricing has examined whether stock return anomalies related to firm characteristics [see Jagannathan and Wang (1996) for example] can be reconciled using neoclassical investment models with factor adjustment costs [ Kogan and Papanikolaou (2013), Kogan et al. (2018), Belo et al. (2014), Favilukis and Lin (2016), etc.] See Kogan and Papanikolaou (2012) for a comprehensive survey of the investment-based asset pricing literature. The fact that agency frictions lead to endogenous operating leverage and the amplification of the volatility of the pricing kernel draws a clear distinction between our paper and existing explanation of value premium based on sticky wages in a complete market. Sticky wages result in heterogeneity in firms' cash flow risk exposures , but it does not amplify the risk prices due to the complete market assumption.

[5] Chien and Lustig (2010) provide a multiplier method to solve for optimal contracting problems with aggregate uncertainty. Ai and Bhandari (2018) use the promised utility approach to solve a general equilibrium model with aggregate uncertainty.

also been used in Guvenen (2009) and Gomes et al. (2013).

We proceed as follows: In section 1, we describe the environment, including agents' preferences, firms' production technology, agency frictions, and shocks. In section 2, we recast the sequential optimal contract problem recursively and solve for the optimal contract using dynamic programming. Section 3 analyzes the implications of this optimal contract for asset prices and investment. In sections 4 and 5, we discuss our calibration strategies and present our main quantitative results. Section 6 formulates our empirical tests and provide a set of empirical results. Section 7 presents our conclusions.

## 1.1 Model

We start with our model environment. We develop our model in discrete time with $t = 0, 1, ....$ We fix the population of workers (mass 1). Each worker is matched with an owner. A firm starts to produce whenever an owner-worker pair is created. Both type of agents have recursive preference with Epstein-Zin type. There is no heterogeneity in preference: both types of agents have an identical risk aversion $\gamma$ and intertemporal elasticity of substitution (IES) $\psi$. The subjective discount factor $\beta$ is also common across agents.

### 1.1.1 Technology

Let $K$ and $I$ denote the capital stock and gross investment respectively. A firm's capital stock $K$ evolves according to

$$K_t = (1 - \delta)K_{t-1} + I_t \tag{1.1}$$

where $\delta \geq 0$ is the rate of depreciation.

Capital investment $I_t$ made at $t$ can be immediately used as capital input to produce. We do not assume one period time-to-build in the physical capital. As can be seen clearly in the next section, this assumption simplifies the solution to firm's recursive maximization problem and reduce the number of relevant state variables by one.

Output is produced according to a standard Cobb-Douglas production function 1.2 with capital $K_t$ and labor $L_t$ as input factors. $Z_t$ is firm's idiosyncratic productivity and $\mathbf{A}_t$ denotes aggregate productivity. Workers supply labor inelastically so we fix $L$ to be 1 at all times. $0 < \alpha < 1$ is the curvature parameter on the production function.

$$Y_t = \mathbf{A}_t \left(Z_t L_t\right)^{1-\alpha} K_t^{\alpha} \tag{1.2}$$

Aggregate productivity $\mathbf{A}_t$ depends on an exogenous shock $\theta_t$ and aggregate capital stock $\mathbf{K}_t$.

$$\mathbf{A}_t = \theta_t \mathbf{K}_t^{1-\alpha} \tag{1.3}$$

A higher aggregate capital stock renders all existing firms more productive, but firms do not take into account the effect of their capital positions on aggregate productivity. This is essentially the capital externality model in the spirit of Romer (1986): it leads to the homotheticity properties of allocations in our model and all quantity variables can be scaled by aggregate capital stock $\mathbf{K}_t$ which further simplifies our model. $\theta_t$ is a Markov process of exogenous productivity shocks and it can take two possible realizations $\{\theta_H, \theta_L\}$ with transition probability matrix $\pi$.

Firm's productivity $Z$ follows a random walk process with $\varepsilon$ being firm-specific shock:

$$\log(Z_{t+1}) = \log(Z_t) + \varepsilon_{t+1} \tag{1.4}$$

where $\varepsilon$ is i.i.d across firms. Moreover, we assume that the distribution of the idiosyncratic productivity shock depends on the aggregate state of the economy $\theta_t$, with the density function $f(\varepsilon|\theta_t)$. We denote the history of shocks up to time $t$ as $(\theta^t, \varepsilon^t) = \{\theta_s, \varepsilon_s\}_{s=0}^t$.

A firm's dividend payout is

$$D_t = Y_t - W_t - \mathbf{p}_t I_t$$

which is output less labor compensation $W_t$ and investment costs $\mathbf{p}_t I_t$. The cost of investment equals the quantity of investment made by the firm, multiplied by the unit price of capital determined in equilibrium. Physical capital investment does not incur investment adjustment cost in the model.

### 1.1.2 Contracting environment

When a firm starts up, it enters a matching market and hires a worker. Firms offer workers state-contingent compensation plans as functions of histories of firm-specific and aggregate shocks. Workers accept the offers if the life-time utilities under contracts are no less than their reservation utilities. A contract also specifies a firm's investment policies that depend on histories of shocks. We denote the compensation plan under a contract for firm $j$ as $\mathcal{W}_j = \{W_{j,t}(\theta^t, \varepsilon_j^t)\}_{t=0}^{\infty}$. Firm's investment policy is denoted as $\mathcal{K}_j = \{K_{j,t}(\theta^t, \varepsilon_j^t)\}_{t=0}^{\infty}$. Firm owners are endowed with ownership claims to these firms and have no labor income. Owners are fully diversified with respect to firm-specific shocks by trading in the financial market. We assume that the matching market is competitive. Initially, any contract yields a zero present value of cash flows from the firm worker pair.

Let $\Lambda_{t,t+\tau}(\theta^{t+\tau}|\theta^t)$ denote the stochastic discount factor from aggregate state $\theta^t$ to $\theta^{t+\tau}$. Also let $\mathbf{p}_t(\theta^t)$ be the price of one unit of capital in state $\theta^t$. The value of firm $j$ under contract $(\mathcal{W}_j, \mathcal{K}_j)$ after history $(\theta^t, \varepsilon_j^t)$ is the present discounted value of a firm's dividend under this contract:

$$V_{j,t}(\mathcal{W}_j, \mathcal{K}_j|\theta^t, \varepsilon_j^t) =$$
$$\mathbf{E}\left[\left\{\sum_{\tau=0}^{\infty}\Lambda_{t,t+\tau}(\theta^{t+\tau}|\theta^t)\left(\mathbf{A}_{t+\tau}Z_{t+\tau}^{1-\alpha}K_{j,t+\tau}^{\alpha} - W_{j,t+\tau} - p(\theta^{t+\tau})(K_{j,t+\tau} - K_{j,t+\tau-1})\right)\right\}\Big|\theta^t, \varepsilon_j^t\right]$$

Workers' utility under the same contract $(\mathcal{W}_j, \mathcal{K}_j)$ solves the following Epstein-Zin preference recursion:

$$U_{j,t}(\mathcal{W}_j, \mathcal{K}_j|\theta^t, \varepsilon_j^t) = \left[(1-\beta)W_{j,t}^{1-\frac{1}{\psi}} + \beta\mathcal{R}_t U_{j,t+1}^{1-\frac{1}{\psi}}\right]^{\frac{1}{1-\frac{1}{\psi}}}$$

where the certainty equivalent operator $\mathcal{R}_t$ is defined as

$$\mathcal{R}_t U_{t+1} = \left(\mathbf{E}\left[U_{j,t+1}^{1-\gamma}(\mathcal{W}_j, \mathcal{K}_j|\theta^{t+1}, \varepsilon_j^{t+1})\Big|\theta^t, \varepsilon_j^t\right]\right)^{\frac{1}{1-\gamma}} \tag{1.5}$$

The key agency friction is that neither side of the contract, firm owners and workers, can commit to the contracts that provide lower continuation values than their outside options. In states of lower firm value, the owner has a stronger incentive to shut down the firm and terminate the contractual relationship. If firm's productivity deteriorates and its value is below 0, the owner finds it more profitable to shut down the firm and start a new firm by matching with another worker in the matching market.

On the other hand, if a worker separates with her employer, she loses a fraction of labor productivity (or human capital) because part of the labor productivity is firm-specific. She immediately matches with another firm which offers a new contract that provides initial utility depending on worker's remaining labor productivity. We use $\underline{U}(\theta^t, \varepsilon_j^t)$ to denote the outside option of a worker after history $(\theta^t, \varepsilon_j^t)$.

A worker with initial productivity $Z_0$ is offered a compensation contract that achieves a life-time utility $U_0$ and an initial amount of capital $K_0$ that worker can operate with. We suppress firm's identity $j$ and index contracts by their initial conditions $(S_0, U_0, K_0)$. The optimal contract maximizes the equity value of the firm subject to several constraints so that the contract is incentive compatible. Given a stochastic process for the SDF and capital prices $\{\mathbf{p}_t(\theta^t)\}_{t=0}^{\infty}$, the contract $\mathcal{W}(S_0, U_0, K_0) = \{W_t(\theta^t, \varepsilon^t)|(S_0, U_0, K_0)\}_{t=0}^{\infty}$ and $\mathcal{K}(S_0, U_0, K_0) = \{K_t(\theta^t, \varepsilon^t)|(S_0, U_0, K_0)\}_{t=0}^{\infty}$ solves the following sequential program:

$$\max_{\mathcal{W},\mathcal{K}} V_0[\mathcal{W}, \mathcal{K}|\theta_0, \varepsilon_0] \tag{1.6}$$

$$U_0(\mathcal{W}, \mathcal{K}|\theta^0, \varepsilon^0) \geq U_0 \tag{1.7}$$

$$U_t(\mathcal{W}|\theta^t, \varepsilon^t) \geq \underline{U}_t(\theta^t, \varepsilon^t) \quad \forall(\theta^t, \varepsilon^t) \tag{1.8}$$

$$V_t(\mathcal{W}, \mathcal{K}|\theta^t, \varepsilon^t, \eta_t) \geq 0 \quad \forall(\theta^t, \varepsilon^t) \tag{1.9}$$

Equation 1.7 is the participation constraint for the worker that says the initial utility should be no less than her reservation utility $U_0$. Equation 1.8 are the limited commitment constraints on the worker side so that we focus on a set of contract that delivers higher worker's values in the contract than her outside options in all states of the world. Equation 1.9 corresponds to owner side limited commitment constraints.

### 1.1.3 Recursive formulation of optimal contracting problem

In this section we recast firm's optimal contracting problem recursively, with the joint distribution of promised utility $U$ and firm productivity $Z$, $\Phi(U, Z)$ as a relevant state variable. At firm level, firm's promised value to workers $U$, firm's capital stock from the last period $K_{-1}$ and firm productivity are relevant state variables. Price of capital and the SDF are functions of the entire distribution of firms $\Phi$ and aggregate state $\theta$.

$$\begin{aligned} V(Z, K_{-1}, & U|\Phi,\theta) = \\ & \max_{U',K,W} \Bigg\{ \mathbf{A} Z^{1-\alpha} K^{\alpha} - W - \mathbf{p}(\Phi,\theta)(K - K_{-1}) + \\ & \qquad \mathbf{E}\Big[\Lambda'(\Phi',\theta'|\Phi,\theta) V(Z', K(1-\delta), U'|\Phi',\theta')\Big] \Bigg\} \end{aligned} \tag{1.10}$$

$$s.t. \quad U = \Big\{(1-\beta)W^{1-\frac{1}{\psi}} + \beta \mathcal{R} U'^{1-\frac{1}{\psi}}\Big\}^{\frac{1}{1-\frac{1}{\psi}}} \tag{1.11}$$

$$V(Z', K(1-\delta), U'|\Phi',\theta') \geq 0 \quad \forall (Z',\Phi',\theta') \tag{1.12}$$

$$U'(Z'|\Phi',\theta') \geq \underline{U}(Z'|\Phi',\theta') \quad \forall (Z',\Phi',\theta') \tag{1.13}$$

Firm's dividend $D$ is output less compensation and cost of new investment, $D = \mathbf{A}Z^{1-\alpha}K^{\alpha} - W - \mathbf{p}(\Phi,\theta)(K - K_{-1})$. In the expression of firm's value 1.10, firm's value $V$ depends on firm's past capital stock $K_{-1}$ only through its market value term $\mathbf{p}K_{-1}$. Past capital stock $K_{-1}$ does not interact with other constraints or control variables of this dynamic programming problem. We define $\tilde{V}$ which is the difference between firm's market value and the market value of firm's capital stock:

$$\tilde{V}(Z, U|\Phi,\theta) = V(Z, K_{-1}, U|\Phi,\theta) - \mathbf{p}(\Phi,\theta)K_{-1} \tag{1.14}$$

We replace the market value $V$ with the cash flow value $\tilde{V}$ in equation 1.14. This allows us to drop $K_{-1}$ and reduce the number of state variables by 1. The firm's value maximization problem now becomes:

$$\begin{aligned} \tilde{V}(Z, U|\Phi,\theta) = & \\ \max_{U',K,W} & \Bigg\{ \mathbf{A} Z^{1-\alpha} K^{\alpha} - W - \mathbf{p}(\Phi,\theta)K + \mathbf{E}\left[\Lambda'(\Phi',\theta'|\Phi,\theta)\tilde{V}(Z', U'|\Phi',\theta')\right]) \\ & + (1-\delta)\mathbf{E}\Big[\Lambda'(\Phi',\theta'|\Phi,\theta)\mathbf{p}'(\Phi',\theta')K\Big] \Bigg\} \end{aligned} \tag{1.15}$$

subject to constraints 1.11, 1.12 and 1.13.

### 1.1.4 Investment sector

In addition to ownerships of firms, owners are also endowed with claims to the profits generated by an investment goods producer that converts consumption numeraire into investment goods. The investment goods sector produces using a simple production technology. To produce $\mathbf{I}$ units of investment goods, investment goods producer takes $\mathbf{I}$ units of consumption numeraire and the production process incurs adjustment cost we assume to be standard convex adjustment cost form $H(\mathbf{I},\mathbf{K}) = \frac{h}{2}\left(\frac{\mathbf{I}}{\mathbf{K}}\right)^2$. The parameter $h$ governs the curvature of adjustment cost function. Production of investment goods takes one time period and investment goods producer supplies newly produced investment goods on the capital market. We formulate the recursive problem for the investment goods producer, given capital price $\mathbf{p}(\Phi,\theta)$ and stochastic discount factor $\Lambda(\Phi',\theta'|\Phi,\theta)$.

$$V_I(\mathbf{I}_{-1},\Phi,\theta) = \max_{\mathbf{I}} \mathbf{p}(\Phi,\theta)\mathbf{I}_{-1} - (\mathbf{I} + H(\mathbf{I},\mathbf{K})) + \mathbf{E}\Big[\Lambda'(\Phi',\theta'|\Phi,\theta)V_I'(\mathbf{I},\Phi',\theta')\Big]$$

The value of investment goods producer depends on its output from last period $\mathbf{I}_{-1}$ because of the sales revenue from selling $\mathbf{I}_{-1}$ units of investment at the prevailing market price $\mathbf{p}(\Phi,\theta)$. Production of new investment $\mathbf{I}$ takes $(\mathbf{I}+H(\mathbf{I},\mathbf{K}))$ units of consumption goods. The dividend payout for investment goods producer is simply $D_I = \mathbf{p}(\Phi,\theta)\mathbf{I}_{-1} - (\mathbf{I} + H(\mathbf{I},\mathbf{K})$. The maximization problem of the investment sector yields the policy function for aggregate investment $\mathbf{I}(\Phi,\theta)$.

### 1.1.5 Owners' utility maximization problem

The objective of the owners is to maximize her utility, subject to a standard budget constraint.

$$V_O(\Phi,\theta) = \max_{\mathbf{C},S'(U,Z),S_I'} \Big\{(1-\beta)\mathbf{C}^{1-\frac{1}{\psi}} + \beta\mathcal{R}(V_O')^{1-\frac{1}{\psi}}\Big\}^{\frac{1}{1-\frac{1}{\psi}}} \tag{1.16}$$

$$\mathbf{C} + \iint S'(U,Z)\Big(V(U,Z|\Phi,\theta) - D(U,Z|\Phi,\theta)\Big)\Phi(U,Z)dUdZ + S_I'(V_I(\Phi,\theta) - D_I(\Phi,\theta)) \leq \iint S(U,Z)V(U,Z|\Phi,\theta)\Phi(U,Z)dUdZ + V_I(\Phi,\theta)S_I$$

$\mathbf{C}$ denotes owners' and $V_O$ is the continuation utility. $S(U,Z)$ is the number of equity shares (in equilibrium, $S(U,Z)=1$ for all types of firms) for a particular firm type $(U,Z)$. $V(U,Z|\Phi,\theta)$ is the cum-dividend price per share; $D(U,Z|\Phi,\theta)$ is the equity payout. Similarly, $V_I(\Phi,\theta)$ is the cum-dividend price for the investment goods sector's equity; $D_I(\Phi,\theta)$ is the payout from the investment goods producer's equity. $S_I$ is the number of equity shares for investment goods producer and in equilibrium $S_I=1$.

Owners maximize utility by trading financial assets on the financial market and diversify away all idiosyncratic risks. Without loss of generality, we assume that there is one representative owner. Solving the portfolio choice problem 1.16 gives owner's consumption and continuation utility as functions of aggregate state variables, $\mathbf{C}(\phi,\theta)$ and $V_O(\phi,\theta)$ respectively. Because the owner is well diversified, her Intertemporal Marginal Rate of Substitution must coincide with the equilibrium SDF[6] . The setup implies that the SDF in our economy has the following two-factor structure: the first term captures consumption growth for the owner and the second term, involving continuation utilities, captures owner's preferences concerning uncertainty about future economic conditions.

$$\Lambda'(\Phi',\theta'|\Phi,\theta)=\beta\left(\frac{\mathbf{C}'(\phi',\theta')}{\mathbf{C}(\phi,\theta)}\right)^{-\frac{1}{\psi}}\left(\frac{V_O'(\phi',\theta')}{\mathcal{R}V_O'(\phi,\theta)}\right)^{\frac{1}{\psi}-\gamma}$$

The owner's budget constraint gives rise to the following resource constraint

$$\begin{aligned}\mathbf{C}(\Phi,\theta)+\iint W(U,Z|\Phi,\theta)\Phi(U,Z)dUdZ+&\mathbf{I}(\Phi,\theta)+H(\mathbf{I},\mathbf{K})(\Phi,\theta)\\ &\leq\iint Y(U,Z|\Phi,\theta)\Phi(U,Z)dUdZ\end{aligned}\tag{1.17}$$

Aggregate output on the right hand side of equation 1.17 equals the sum of owner's consumption $\mathbf{C}(\Phi,\theta)$, total compensation for workers $\iint W(U,Z|\Phi,\theta)\Phi(U,Z)dUdZ$ and total investment expenditure $\mathbf{I}(\Phi,\theta)+H(\mathbf{I},\mathbf{K})(\Phi,\theta)$.

The law of motion of total capital is

$$\iint K(U,Z|\Phi,\theta)\Phi(U,Z)dUdZ=\mathbf{K}(\Phi,\theta)\tag{1.18}$$

[6] In our model, unconstrained workers equalize their marginal utility with the owner and their marginal utility ratios are also valid SDFs. Therefore, the risk sharing between owner and unconstrained workers can be implemented by allowing unconstrained workers to trade state contingent assets with the owner. Hence our model also provides a theory of endogenous stock market participation based on agency frictions, which can be related with earlier works that highlight the importance of stock market participation and asset pricing such as Mankiw and Zeldes (1991), Basak and Cuoco (1998), Malloy et al. (2009), Guvenen (2009), Wachter and Yogo (2010), Elkamhi and Jo (2018)

The aggregate demand for capital equals the total capital available which consists of two components: firms' undepreciated capital and new investment made by the investment sector.

### 1.1.6 Recursive competitive equilibrium

In this section, we define a recursive competitive equilibrium for our economy. The homogeneity property resulting from assumptions on preference and technology enables us to construct a competitive equilibrium that has two aggregate state variables $(\phi, \theta)$. $\theta$ is the Markov state for aggregate productivity shock. $\phi$ is a one-dimensional density that summarizes firms' type. At the end of this section, we show that the homogeneity feature of the model allows us to reduce firms' distribution $\Phi$ to a one dimensional distribution $\phi$.

The aggregate productivity specification generates endogenous economic growth into our model. In order to solve for the competitive equilibrium, we introduce stationary aggregate state variables and scale aggregate variables by aggregate capital stock $\mathbf{K}$: de-trended firm owner's consumption $\mathbf{c} = \frac{\mathbf{C}}{\mathbf{K}}$; Normalized aggregate investment $\mathbf{i} = \frac{\mathbf{I}}{\mathbf{K}}$.

For firm level variables, we show in the appendix that the value function $\tilde{V}(U, Z|\phi, \theta)$ in 1.15 has the following representation:

$$\tilde{V}(U, Z|\Phi, \theta) = \tilde{v}(\frac{U}{\mathbf{K}Z}|\phi, \theta)\mathbf{K}Z$$

for some function $\tilde{v}(\cdot|\phi, \theta)$ that represents normalized firm value. This is true because the model features a balanced growth path and firm-level productivity follows random walk process. We scale other firm variables (promised utility $u$, compensation $w$ and capital stock $k$)

$$u = \frac{U}{\mathbf{K}Z} \quad w = \frac{W}{\mathbf{K}Z} \quad k = \frac{K}{\mathbf{K}Z}$$

We recast the optimal contracting problem 1.15, 1.11, 1.12 and 1.13 recursively using normalized variables

$$\begin{aligned}\tilde{v}(u|\phi, \theta) = \max_{u', k, w} \Big\{ &\theta k^{\alpha} - w - \mathbf{p}(\phi, \theta)k + \mathbf{E}\Big[\Lambda'(\phi', \theta'|\phi, \theta)(1 - \delta + \mathbf{i}(\phi, \theta))e^{\varepsilon'}\tilde{v}(u'|\phi', \theta')\Big] \\ &+ k(1-\delta)\mathbf{E}\Big[\Lambda'(\phi', \theta'|\phi, \theta)\mathbf{p}'(\phi', \theta')\Big]\Big\}\end{aligned} \tag{1.19}$$

$$u = \left\{ (1-\beta) w^{1-\frac{1}{\psi}} + \beta \mathcal{R}(u')^{1-\frac{1}{\psi}} \right\}^{\frac{1}{1-\frac{1}{\psi}}} \tag{1.20}$$

$$(1-\delta+\mathbf{i}(\phi,\theta))e^{\varepsilon'}\tilde{v}(u'|\phi',\theta') \geq -(1-\delta)\mathbf{p}'(\phi',\theta')k \quad \forall(\phi',\theta',\varepsilon') \tag{1.21}$$

$$u'(u|\mathbf{c}',\theta') \geq \underline{u}(\phi',\theta') \tag{1.22}$$

Solving the normalized problem 1.19 gives optimal investment decision $k(u,\phi,\theta)$, compensation policy $w(u,\phi,\theta)$ and continuation utility contingent on different values of $\eta$:$u'_{\eta'=0}(\varepsilon',u,\theta'|\phi,\theta)$ and $u'_{\eta'=1}(\varepsilon',u,\theta'|\phi,\theta)$.

We now describe the construction of our aggregate state variable $\phi$, which we will refer to as the "normalized measure". Let $w(u), k(u)$ [7] denote the compensation and investment policy for the optimal contracting problem 1.19 and recall that $\Phi(U,Z)$ denote the joint distribution of firms promise to workers $U$ and idiosyncratic productivity $Z$. In general, $\Phi(U,Z)$ is needed as a state variable in the construction of an equilibrium because we need to count the total resource produced and total compensation paid by firms in resource constraint 1.17. To illustrate the construction of the normalized measure $\phi$, we use the total labor compensation term in the aggregate resource constraint 1.17 as an example. We can equivalently express the integral on labor compensation term as

$$\begin{aligned}\iint w(\frac{U}{\mathbf{K}Z})\mathbf{K}Z\Phi(u\mathbf{K}Z,Z)dudZ &= \mathbf{K}\iint w(u)Z\Phi(u\mathbf{K}Z,Z)dudZ \\ &= \mathbf{K}\int w(u)\Big[\int Z\Phi(uZ,Z)dZ\Big]du\end{aligned}$$

We define the normalized measure $\phi(u) = \int S\Phi(uZ,Z)dZ$. Note that for a given $Z$, the expression $\Phi(uZ,Z)$ is the joint density of $(u,Z)$. As a result, the density $\phi(u)$ basically records the average productivity of firms whose normalized utility equals $u$.

With this construction we now bring back all state variables and the aggregate resource constraint 1.17 now reads

$$\mathbf{c}(\phi,\theta) + \int w(u,\phi,\theta)\phi(u|\theta)du + \mathbf{i}(\phi,\theta) + \frac{h}{2}\mathbf{i}(\phi,\theta)^2 = \int k(u,\phi,\theta)^\alpha \phi(u|\theta)du \tag{1.23}$$

[7] We ignore the dependence of policy function on the distribution itself and aggregate state. This slight abuse of notation makes explain the construction of normalized measure more easily by focusing on relevant state variable $u$.

The capital market clearing condition 1.18 can be written as

$$\int k(u,\phi,\theta)\phi(u|\theta)du = 1 \tag{1.24}$$

We can also characterize the law of motion of the normalized measure $\phi$. Given the optimal policy for continuation utilities $u'(u,\theta',\varepsilon'|\phi,\theta)$, the law of motion of $\phi$, $\phi' \equiv \Gamma(\theta'|\theta,\phi)$ is given by

$$\forall \tilde{u} \quad \phi'(\tilde{u}|\theta') = (1-\kappa)\int \phi(u|\theta)\int e^{\varepsilon'} f(\varepsilon'|\theta')\mathcal{I}_{u'(u,\theta',\varepsilon'|\phi,\theta)=\tilde{u}} + \kappa \int e^{\varepsilon'} f(\varepsilon'|\theta')\mathcal{I}_{u'(\overline{u},\theta',\varepsilon'|\phi,\theta)=\tilde{u}} \tag{1.25}$$

where $\mathcal{I}$ is an indicator that takes the value one if the continuation utility for firm-specific shock $\varepsilon'$ equals $\tilde{u}$.

### Entry, exit and workers' outside options

We allow entry and exit of firms to maintain a stationary distribution of firm-specific productivities in the model. Firms and workers exit the economy if they receive an exogenous death shock at the rate of $\kappa$ per period. For simplicity, we assume that within a firm-worker pair, the death shock of the firm and the death shock of the worker are perfectly correlated. Once hit by the death shock, the capital of the firm and firm-specific productivity evaporate.

A measure of $\frac{1}{\kappa}$ of new firm start every period with average productivity $Z_0 = 1$ and $\overline{u}(\phi,\theta)$. $\overline{u}(\phi,\theta)$ is determined by the zero profit condition which says competition in the matching market drives the initial value of a firm-worker pair to be zero.

If a worker defaults on the contract, she keeps a $\lambda$ fraction of productivity and immediately matches with a firm that starts with initial utility $\overline{u}(\phi,\theta)$. By homogeneity, we can show that worker's outside value $\underline{u}(\phi,\theta)$ in 1.22 is given by $\underline{u}(\phi,\theta) = \lambda\overline{u}(\phi,\theta)$.

### Equilibrium

We now define the recursive competitive equilibrium.

A recursive competitive equilibrium consists of a law of motion for the normalized measure $\phi$, $\Gamma(\theta'|\phi,\theta)$; a set of prices: the SDF $\{\Lambda(\theta'|\phi,\theta)\}$[8] , prices of capital

[8] In principle, value, policy and pricing functions should all depend on the normalized measure in the next period $\phi'$. However, given that all agents are able to forecast the normalized measure using $\Gamma$, putting $\phi'$ as an argument of equilibrium outcome functions is redundant.

$\{\mathbf{p}(\phi, \theta)\}$; firm's franchise value $\tilde{v}(u|\phi, \theta)$ and policy function for wages $w(u|\phi, \theta)$, investment $k(u|\phi, \theta)$ and continuation utilities $u'(u, \theta', \varepsilon|\phi, \theta)$; policy functions for owner's consumption share $\mathbf{c}(\phi, \theta)$ and aggregate investment $\mathbf{i}(\phi, \theta)$ such that

- under the price of capital $\mathbf{p}(\phi, \theta)$, capital market clearing condition 1.24 holds.
- the owner solves her consumption and portfolio choice problem 1.16 and the equilibrium SDF is consistent with owner's IMRS derived from the solution to problem 1.16.
- Given the SDF, capital price and the law of motion of the normalized measure, the value function and the policy functions solve the optimal contracting problem 1.19.
- given the policy functions for continuation utilities, the law of motion of the normalized measure $\phi$ satisfies 1.25.
- the policy functions and the normalized measure $\phi$ satisfy the goods market clearing condition 1.17.

We apply Krusell and Smith (1998) type of technique and forecast the law of motion of the normalized measure $\phi$ and price of capital using the owner's consumption share $\mathbf{c}$. All details are included in the appendix.

## 1.2 Implications of optimal contract

We begin by analyzing properties of optimal allocations without agency frictions. We then characterize the optimal contract with agency concerns.

### 1.2.1 Uninsurable tail risk and firm investment

With no agency frictions (the case without limited commitment constraint 1.21 and 1.22), our model specializes to the neoclassical benchmark. First, firms can fully insure workers against idiosyncratic shocks and workers' continuation utility do not respond to idiosyncratic shocks. Second, given the productivity specification and the homogeneity of production technology, capital allocation in the first best case follows a simple linear rule: a firm's capital is linearly proportional to its productivity, $K = Z\mathbf{K}$.

With perfect risk sharing, continuation utility does not respond to idiosyncratic shocks and thus normalized continuation value $u'(\varepsilon', u, \theta'|\phi, \theta)$ must be inversely proportional to the idiosyncratic shock $e^{\varepsilon'}$. In particular, an extremely negative shock in $\varepsilon'$ pushes $u'(\varepsilon', u, \theta'|\phi, \theta)$ toward infinity. However, this cannot be true in the presence of firm side incentive constraint 1.21. Keeping an unproductive worker adds negative value to the firm because the cash flow produced by the worker is not enough to pay for her promised compensations.

In the first best case, firm's share in aggregate capital $\mathbf{K}$ only depends on its productivity $Z$. With agency frictions, risk-sharing considerations also contribute to the determination of firm's optimal investment policy. When a firm's value declines, the firm-side constraint is likely to bind, leading to poor risk-sharing conditions for workers. To avoid further losses and to improve risk sharing, these firms accelerate investment to grow out of the constraint. The effect of agency frictions on utility provision and investment policy under the optimal contract is summarized by the following propositions.

**Proposition 1.** *There exists $\hat{\varepsilon}(u, \theta'|\phi, \theta)$ (let $\hat{u}(u, \theta'|\phi, \theta)$ denote the continuation utility associated with $\hat{\varepsilon}(u, \theta'|\phi, \theta)$) and $\underline{\varepsilon}(u, \theta'|\phi, \theta)$ such that*

$$u'(\varepsilon', u, \theta'|\phi, \theta) = \begin{cases} \tilde{v}^{-1}(e^{\hat{\varepsilon}(u,\theta'|\phi,\theta)-\varepsilon'}\tilde{v}(\hat{u}(u, \theta'|\phi, \theta)|\phi', \theta') & \forall \varepsilon' \leq \hat{\varepsilon}(u, \theta'|\phi, \theta) \\ \underline{u}(\phi, \theta) & \forall \varepsilon' \geq \underline{\varepsilon}(u, \theta'|\phi, \theta) \end{cases}$$

*And the continuation utility $\hat{u}(u, \theta'|\phi, \theta)$ at the point $\hat{\varepsilon}(u, \theta'|\phi, \theta)$ is determined by the following expression*

$$(1 - \delta + \boldsymbol{i}(\phi, \theta))e^{\hat{\varepsilon}(u,\theta'|\phi,\theta)}\tilde{v}(\hat{u}(u, \theta'|\phi, \theta)|\phi', \theta') = -(1 - \delta)\boldsymbol{p}'(\phi', \theta')k(\hat{u}, \phi, \theta)$$

*where $k(\hat{u}, \phi, \theta)$ represents the investment policy function.*

*For all $\varepsilon' \in (\hat{\varepsilon}(u, \theta'|\phi, \theta), \underline{\varepsilon}(u, \theta'|\phi, \theta))$, the continuation utility is determined by the risk sharing condition which equalizes owner's SDF with worker's:*

$$\Lambda(\theta'|\phi, \theta) = \beta e^{-\gamma\varepsilon'}(1-\delta+\boldsymbol{i}(\phi, \theta))^{-\gamma}\Big(\frac{w(u'(\varepsilon', u, \theta'|\phi, \theta), \phi', \theta')}{w}\Big)^{-\frac{1}{\psi}}\Big(\frac{u'(\varepsilon', u, \theta'|\phi, \theta)}{\mathcal{R}u'}\Big)^{\frac{1}{\psi}-\gamma}$$

Figure 1.1 illustrates the main message of the above proposition. Figure 1.1 plots continuation utility $u'(\varepsilon', u, \theta'|\phi, \theta)$ (on the y-axis) as a function of realizations of idiosyncratic shocks $\varepsilon'$ (on the x-axis), holding all other state variables fixed. For shocks lower than the cutoff shock $\hat{\varepsilon}(u, \theta'|\phi, \theta)$, continuation value is still inversely related with the shock but the curvature is not as steep as in the region where worker is perfectly insured, leading to compensation reduction if firm receives a large adverse idiosyncratic shock and firm side constraint 1.21 binds.

Proposition 1 also highlights the interaction between risk sharing and firm investment. Figure 2 presents an experiment in which we slightly deviate investment from its optimal level $k(u, \phi, \theta)$ and increase it to $k_2$. We observe that the cutoff shock decreases from $\hat{\varepsilon}(u, \theta'|\phi, \theta)$ to $\hat{\varepsilon}_2$. The continuation utility at the cutoff point raises from $\hat{u}(u, \theta'|\phi, \theta)$ to $\hat{u}_2$. Since firm side constraint binds for shocks lower than the cutoff value, the decrease in cutoff value is associated with lower probability of binding firm side constraint and compensation cut for the worker which improves worker's risk sharing condition. Firm takes into account the benefit of investment on improving risk sharing when making investment decision. Proposition 2 characterizes the trade offs of investment

**Proposition 2.** *Investment policy function $k(u, \phi, \theta)$ satisfies the investment equation below*

$$\begin{aligned} \boldsymbol{p}(\phi, \theta) = \alpha\theta k(u, \phi, \theta)^{\alpha-1} &+ (1-\delta)\mathbf{E}\Big[\Lambda'(\phi', \theta'|\phi, \theta)\boldsymbol{p}'(\phi', \theta')\Big] \\ &+ (1-\delta)\mathbf{E}\Big[\lambda'(\varepsilon', u, \theta'|\phi, \theta)\Lambda'(\phi', \theta'|\phi, \theta)\boldsymbol{p}(\phi', \theta')\Big] \end{aligned} \tag{1.26}$$

*where $\lambda'(\varepsilon', u, \theta'|\phi, \theta)$ corresponds to the Lagrangian Multiplier of the firm-side incentive constraint 1.21.*

Proposition 2 says that the cost of an additional unit of capital $\mathbf{p}(\phi, \theta])$ must also equal the present discounted value of marginal profits of having an extra unit of capital. The direct benefit of an extra unit of capital is to have $\alpha\theta k^{\alpha-1}$ extra units of revenue contemporaneously. Also, having more capital directly raises the market value of capital which is captured by the term $(1-\delta)\mathbf{E}\Big[\Lambda'(\phi', \theta'|\phi, \theta)\mathbf{p}'(\phi', \theta')\Big]$. More importantly, an extra unit of capital reduces the likelihood of experiencing binding firm side limited commitment constraint which improves worker's risk sharing condition. The term $(1-\delta)\mathbf{E}\Big[\lambda'(\varepsilon', u, \theta'|\phi, \theta)\Lambda'(\phi', \theta'|\phi, \theta)\mathbf{p}(\phi', \theta')\Big]$ reflects the risk-sharing benefit of an extra unit of capital.

### 1.2.2 Uninsurable tail risks and the amplification of market price of risk

In our model uninsurable tail risks amplify the volatility of the equilibrium pricing kernel. The amplification of risk prices can connect to the mechanism in early literature such as Alvarez and Jermann (2001), Chien and Lustig (2010) and especially Ai and Bhandari (2018). In our calibration, firm-level tail shocks are more likely to happen in recessions than booms, which reflects the finding that firm-level tail rare events are more likely to take place in economic downturns. Therefore, workers experience wages cuts more frequently in economic downturns. With recursive utility and persistent countercyclical idiosyncratic risks, the prospect of future lack of risk sharing raises workers' current marginal utilities in downturns. As a result, the optimal risk sharing scheme compensates workers by allocating a higher share of aggregate output to the workers as labor compensations in downturns. Labor share moves negatively with the aggregate output. The countercyclicality of labor share translates into a procyclical consumption share of the owner and amplifies risk prices.

The main difference with the existing literature is that our model is setup in a production economy. This introduces two new forces to the literature on endogenous incomplete market and asset pricing, which typically assume the environment of an endowment economy: first of all, as demonstrated in proposition 2, to alleviate agency frictions and improve risk sharing introduces a new benefit of investment. Secondly, with the access to an aggregate investment sector, risk-averse owner can always smooth out the cyclical variations in her consumption, by adjusting investment and production accordingly. Quantitatively, we show that the amplification of risk prices is crucial for our model to jointly explain asset prices and investment dynamics at both the aggregate and firm level.

### 1.2.3 Endogenous labor operating leverage

In figure 1.4 and 1.3, we plot the normalized firm value ($\tilde{v}$) and compensation $w$ as functions of normalized continuation utility $u$. With firm-side limited commitment, firm value is strictly concave with respect to $u$ while the compensation policy is strictly convex. Firm values are lower because the dividends are discounted more with incomplete markets. The steeper curvature reflects the higher marginal cost of providing insurance as the limited commitment constraint is likely to bind for high $u$ firms. Since the principal cannot deliver higher $u'$s, in the future, wages per unit of output increases more than proportionately with $u$ in the limited commitment case.

An alternative interpretation is that $\tilde{v}(u|\phi, \theta)$ is the valuation of the firm's claim to its equity holders and $u$, the promised value is the valuation of liabilities to workers. Thus $\frac{u}{\tilde{v}+u}$ ratio is a measure of operating leverage. Under the optimal contract, firms with a sequence of adverse productivity shocks remain solvent by delaying compensation payments but accumulating debt. Higher debt makes the firm more risky.

Another observation in 1.3 and 1.4 is the variation of labor leverage and firm value across aggregate states. A firm with identical committed value to its worker have a higher labor leverage in $\theta_L$ than in $\theta_H$. The steeper curvature of the compensation policy in $\theta_L$ reflects that marginal cost of providing insurance is higher for firms with equal committed value in $\theta_L$ than in $\theta_H$.

## 1.3 Numerical algorithm

We solve for recursive competitive equilibrium numerically. Because the financial market is endogenously incomplete, one cannot solve for allocations first and then obtain prices as in a standard complete market model. Allocations and prices must be solved at the same time. Second, normalized measure $\phi$ is a state variable with infinite dimension. We use a procedure adopted from Krusell and Smith (1998) and replace $\phi$ with sufficient statistics that can accurately characterize the dynamics of $\phi$ over time and states. We further build on the numerical procedure in Krusell and Smith (1997) and Guvenen (2009) to jointly look for the market clearing prices for both markets.

The distribution $\phi$ enters the problem through its effect on the stochastic discount factor and market clearings. To approximate the law of motion of $\Gamma$ and capital price $\mathbf{p}$, we conjecture that agents forecast the future SDF and capital prices using the current owner's consumption share $\mathbf{c}$ and aggregate states of the economy. In particular, we assume that $\mathbf{c}' = \Gamma_{\mathbf{c}}(\theta'|\theta, \mathbf{c})$ and price of capital depends on owner's consumption share and aggregate state, $\mathbf{p}(\mathbf{c}, \theta)$.

Using $\mathbf{c}$ as the state variable is both numerically efficient and computationally convenient. Note that the stochastic discount factor depends on $\phi$ for two reasons. First, $\phi$ affects capital owners consumption through the aggregate resource constraint 1.23. Second, distribution of $\phi$ is a slow moving variable that induces persistent variation in owner's continuation utility. Hence, persistence in the distribution $\phi$ translates into predictable variable in owner's consumption share $\mathbf{c}$.

The equilibrium price of capital is determined such that the aggregate demand for capital equals the aggregate capital. Different need for relaxing binding firm side constraint introduces significant heterogeneity in firms' capital demand and hence we do not obtain any aggregation results which could simplify the search for capital price. We adopt the numerical algorithm in the heterogeneous agent literature such as Krusell and Smith (1997) and Guvenen (2009), that look for market clearing price of bonds. The basic idea is that we first solve for firm policy functions for compensation and investment. Then we move to the simulation step as in the original Krusell and Smith (1998) method. At each simulation time point, we introduce another step to solve for firms' demand function for capital as a function of an arbitrary chosen capital price. We solve an associated dynamic programming problem at each time point in the simulation path assuming that firms perceive future SDF and capital prices as given by the existing forecasting rules. Price of capital becomes another state variable of this programming problem. We apply a root-finding routine and search for the capital price that guarantees the market clearing condition 1.24 is satisfied with enough numerical accuracy. We then use firms' compensation and investment policies under market clearing price of capital to compute the owner consumption share at this time point. We describe these steps in details in the appendix.

## 1.4 Quantitative results

In this section, we assess the model's ability to replicate key moments of quantities and asset returns at both the aggregate and firm-level, using the numerical algorithm discussed in the previous section. We focus on a long sample of U.S. annual data including the pre Second World War period as our calibration targets for macroeconomic quantities and aggregate financial moments. Macro variables are de-trended and deflated using their corresponding price index available from the Bureau of Economic Analysis (BEA). Firm-level statistics in the data are computed using a shorter sample that ranges from 1978 to 2017 because two important data variables, capital expenditure and labor compensation, are sparse in earlier years.

### 1.4.1 Parameter values

The parameters of our benchmark model can be divided into five groups. The first group of parameters governs the stochastic process for aggregate productivity. We assume that aggregate productivity shock $\theta$ follow a two-state Markov Chain and $\theta$ can take two values $\{\theta_H, \theta_L\}$. The transition probability between different states is denoted as $\pi$. We label $\theta_L$ as recession states and $\theta_H$ as boom states. We choose the transition probability from boom to boom to be 0.9802 at a quarterly frequency to target the duration of U.S. economic booms which is about 12 years on average. The transition probability from recession to recession is set to match the 4 years' length of U.S. recessions on average. We set the unconditional volatility of the shock to be 2.2% annually that is consistent the volatility of TFP shock typically assumed in the RBC literature such as King and Rebelo (1999).

On the production technology side, the elasticity of firm output to capital parameter $\alpha$ is set to be 0.22. We calibrate this parameter so that the average capital share in our model is close to 0.33 as in the U.S. data. In the model, parameter $\alpha$ cannot be interpreted as capital share as in the standard RBC because owner's income not only comes from firms' profit but also from the profit of investment sector. We set the annual depreciation rate of physical capital $\delta$ to be 10% as in Kydland and Prescott (1982) and King and Rebelo (1999).

Preference parameters including risk aversion $\gamma$, IES $\psi$ and the discount factor $\beta$. IES is set to be 2 which is consistent with the long run risk literature following Bansal and Yaron (2004). We assume that all agents have a preference for early resolution of uncertainty which is a common preference assumption in the long run risk literature. We require a significantly lower risk aversion 4.5 than that in a typical long run risk model. With a low level of risk aversion, we achieve sufficient amplification in the equilibrium risk prices. The quarterly discount factor is 0.99 to guarantee that the average risk-free rate is plausibly low as in the data.

Firm-specific productivity follows a random walk process. We parameterize our idiosyncratic shock process using firm level sales data as in Bloom (2009) and Salgado et al. (2017). In particular, we assume that $f(\varepsilon|\theta)$ is Gaussian in booms and follows a mixture of a Gaussian and a fat-tailed distribution with negative exponential form in recessions with one extra parameter. This assumption leaves us with two parameters $\{\mu_\theta, \sigma_\theta\}_{\theta \in \theta_H, \theta_L}$ for the Gaussian distribution per aggregate state, two parameters $\{\iota, \varepsilon^{max}\}$ for the negative exponential, and $p \in (0, 1)$ as the mixing probability that represents the probability from a draw of the negative exponential distribution. We normalize the mean of exponential of shock to be 1. We further assume a conditional mean of unity for each of the individual distribution in the mixture too. These restrictions imply $\mu_\theta = -\frac{\sigma_\theta^2}{2}$ and $\varepsilon^{max} = \log \frac{1+\iota}{\iota}$. Overall, these assumptions reduce the number of parameters that govern idiosyncratic shock process to four, including $[\sigma_H, \sigma_L, \iota, p]$.

We search for values of these four parameters to match the cross-sectional distribution of annual sales growth moments documented in Salgado et al. (2017): the $50^{\text{th}}$-to-$10^{\text{th}}$ log percentiles differential in recession, the Kelley Skewness in recession and the unconditional cross-sectional standard deviation. In our parameterization, the standard deviation of sales growth is set to be identical across business cycles because we focus on the asset pricing implications of individual firm-level tail events that are captured by the negative exponential distribution. Note that Salgado et al. (2017) also reports that in expansions, firms' annual sales growth distribution exhibit a right tail. The functional form we take on shocks is flexible enough to model right tails in good aggregate state which does not bring additional asset pricing insights. Therefore, we choose to carefully calibrate idiosyncratic shock process in recessions.

The parameter $\lambda$ that is workers' human capital loss upon separation, determines workers' incentives to renege on the labor contract. For the calibration of $\lambda$, we use information from Davis and von Wachter (2011) that estimates the present value of earning losses due to separation. We target the consumption equivalent of the value of separation to be 70% of pre-separation wages

### 1.4.2 Aggregate quantity dynamics

In this section, we show that our benchmark calibration is largely consistent with macroeconomic aggregates observed in the data. In addition, the volatility of owner's consumption is crucial to the success of our model to explain a large equity premium. To put further discipline on the dynamics of the pricing kernel, we calibrate the owner's consumption volatility to that of wealthy investors in the data.

The quantity dynamics produced by our main calibrations are shown in panel A and B of table 1.2. The model retains the basic successes of the RBC framework: aggregate consumption growth, which is the sum of the owner and all workers' consumption, is about half as volatile as output growth. The growth rate of aggregate investment is more than three times as volatile as output growth. Moreover, aggregate consumption growth is moderately autocorrelated as in the data. The persistence in aggregate consumption is inherited from the persistent fluctuations of our endogenous state variable $\mathbf{c}$ which keeps close track of the law of motion of slow-moving distribution of firms. Our model overestimates the contemporaneous correlation between aggregate consumption and investment which is also a well-known drawback of the RBC model.

Capital owner's consumption is the relevant pricing factor to price risky securities. We calibrate our model to be consistent with the large volatility of wealthy investors' consumption, as a further discipline on the equilibrium risk prices. Malloy et al. (2009) and Wachter and Yogo (2010) report the volatility of consumption growth of wealthy stock holders. Specifically, using the method proposed in Malloy et al. (2009), Goldstein and Yang (2015) estimates the consumption volatility for the very top stock holders in the CEX data to be 13.5% per year. In our model, the consumption volatility of capital owner is 9.7% that is close to Goldstein and Yang (2015) estimates.

Figure 1.1: continuation utility

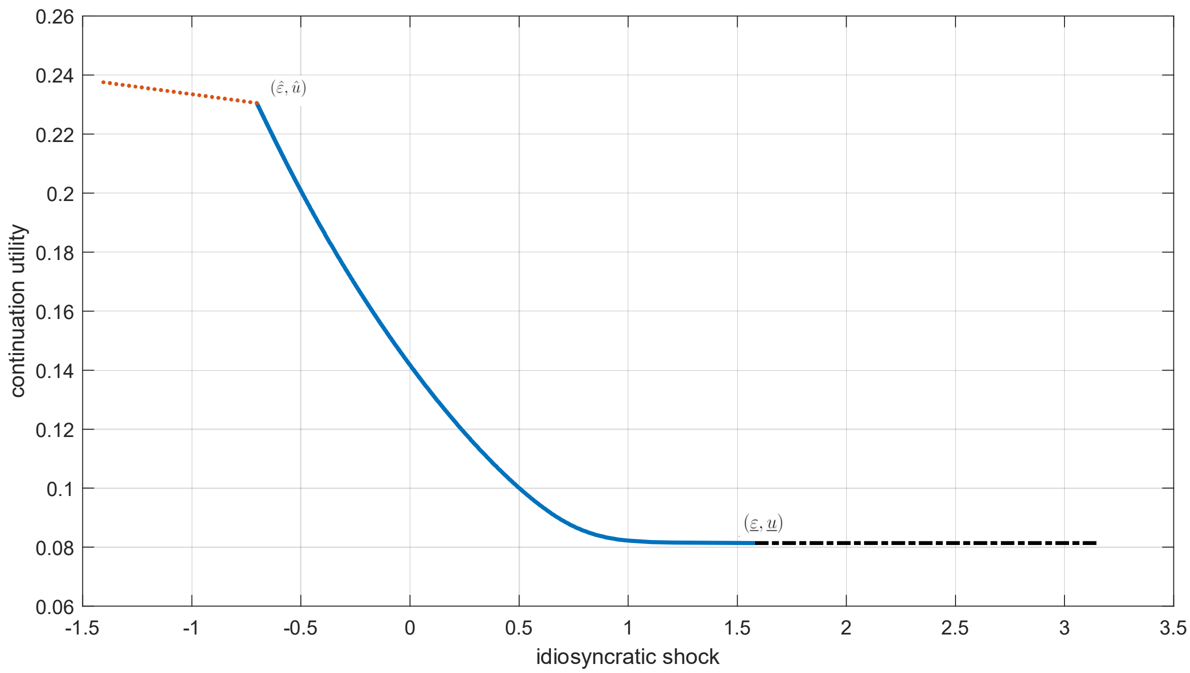


Figure 1.2: continuation utility.

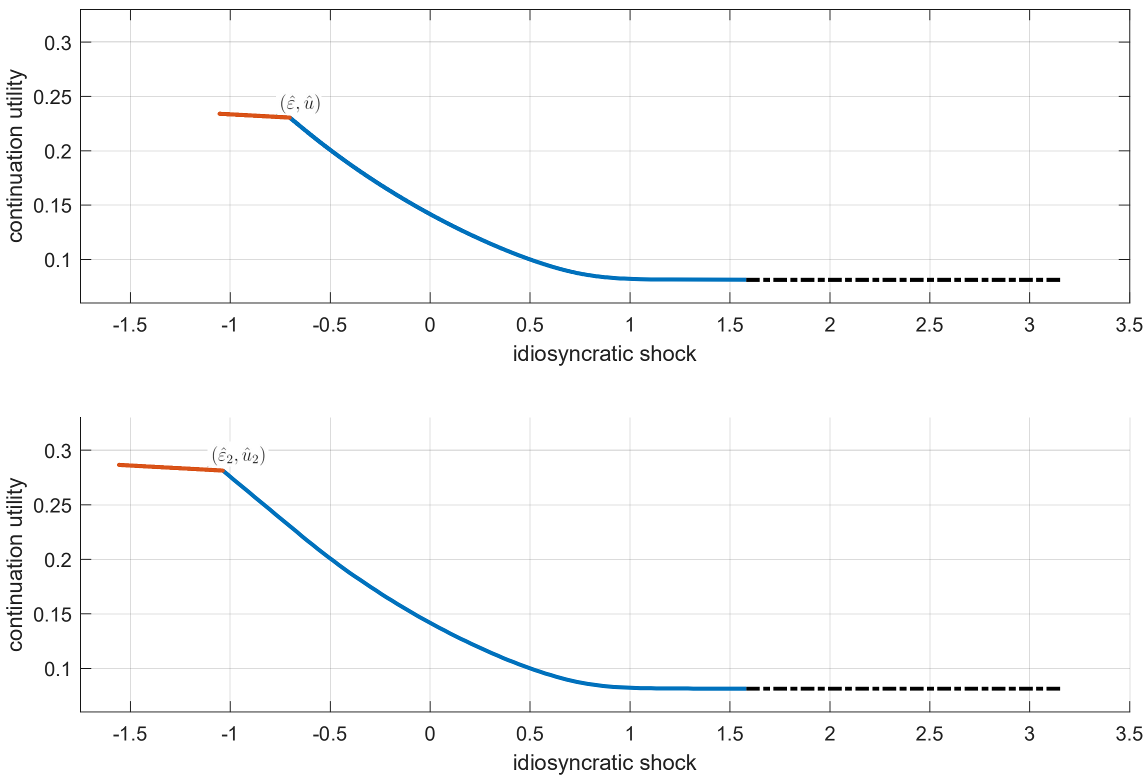


Table 1.1: Key parameters

We calibrate the model at quarterly frequency. To facilitate the comparison of key parameter values with the literature, we convert quarterly parameter values into annual ones and report the annual parameter values in this table with a few exceptions. The volatility of idiosyncratic shock parameter is simply its quarterly value multiply by $\sqrt{12}$. The probability of receiving tail shock in recession is 1 minus the probability of not getting a tail shock for four quarters of consecutive recessions. The conversion of probability of selling capital at quarterly level to annual level is quite similar with the conversion for the probability of receiving tail shocks.

| Parameter | symbol | value |
|---|---|---|
| *aggregate productivity* | | |
| unconditional volatility of productivity shock | $\sigma_\theta$ | 2.20% |
| transition probability from boom to boom (Quarterly) | $\pi(\theta_H\|\theta_H)$ | 98.02% |
| transition probability from recession to recession (Quarterly) | $\pi(\theta_L\|\theta_L)$ | 93.71% |
| *production technology* | | |
| capital elasticity | $\alpha$ | 0.22 |
| aggregate adjustment cost | $h$ | 20 |
| depreciation of physical capital | $\delta$ | 10% |
| *preference* | | |
| risk aversion | $\gamma$ | 4.5 |
| Elasticity of intertemporal substitution | $\psi$ | 2 |
| subjective discount factor | $\beta$ | 0.99 |
| *firm-specific shocks* | | |
| volatility of idiosynratic shock | $\sigma_F$ | 30.41% |
| probability of receiving tail shock: recession | $\rho$ | 7.76% |
| *agency frictions* | | |
| present value of earning loss upon separation | $\mu$ | 30% |

Table 1.2: Aggregate Quantities and Prices

This table presents annualized macroeconomic and asset pricing moments from the data and the benchmark model. Panel A reports the average growth rate and volatilities of aggregate output growth $\Delta\log(Y)$, aggregate consumption growth $\Delta\log(\mathbf{C}+W)$ that is the sum of owner's consumption and workers' total consumption, physical investment growth $\Delta\log(\mathbf{i})$ and wealthy owner's consumption $\Delta\log(\mathbf{C})$. Panel B reports the mean and volatility of risk-free rate $r^f$ and market excess return $R^e$. We report un-levered excess return. The model statistics are computed from 100 parallel samples and each sample contain 10000 quarters of simulated data using the policy functions obtained from the model solutions. Data moments are computed using US annual data from 1930 to 2016.

| Moments | Data | Model |
|---|---|---|
| Panel A: *Macro Quantities* | | |
| $\mathbf{E}(\Delta\log(Y))$ | 1.70% | 1.90% |
| $\mathbf{E}(\frac{\mathbf{C}+W}{\mathbf{C}+W+\mathbf{I}})$ | 78.67% | 84.70% |
| $\mathbf{E}(\frac{\mathbf{I}}{\mathbf{C}+W+\mathbf{I}})$ | 21.33% | 15.31% |
| Panel B: *Volatilities and Autocorrelations* | | |
| $\sigma(\Delta\log(\mathbf{C}+W))$ | 2.53% | 1.83% |
| $\sigma(\Delta\log(\mathbf{i}))$ | 16.40% | 14.56% |
| $\sigma(\Delta\log(\mathbf{i}))/\sigma(\Delta\log(Y))$ | 3.32 | 3.83 |
| $\sigma(\Delta\log(\mathbf{C}+W))/\sigma(\Delta\log(Y))$ | 0.52 | 0.64 |
| $\sigma(\Delta\log(\mathbf{C}))$ | 13.05% | 9.10% |
| $\mathrm{AC}_1(\Delta\log(\mathbf{C}+W))$ | 0.49 | 0.39 |
| $\rho(\Delta\log(\mathbf{C}+W),\Delta\log(\mathbf{I}))$ | 0.39 | 0.77 |
| Panel C: *Asset Prices* | | |
| $\mathbf{E}(R^e)$ | 7.68% | 5.98% |
| $\sigma(R^e)$ | 20.26% | 13.61% |
| $\mathbf{E}(R^f)$ | 0.41% | 0.42% |
| $\sigma(R^f)$ | 0.93% | 0.22% |

The observation that our model produces large volatility in aggregate investment despite fairly high adjustment cost parameter is noteworthy. In standard models, the calibration of adjustment cost parameter that can account for the equity premium typically leads to smooth aggregate investment, which is not the case in our calibration. This quantitative success hinges on a novel mechanism that links the risk premium component of cost of capital to aggregate investment [9] . Suppose the economy is in downturn, the demand for aggregate investment is low because recession state is persistent so that the future marginal benefit of investment is low. Investment sector should reduce investment expenditure significantly but it is not economically beneficial to do so because high adjustment cost imposes a penalty on large reduction of investment. Now suppose the risk premium is higher in recession and capital owner discount future marginal benefit more in recessions. A higher cost of capital will further lower the demand for aggregate investment, thus leading to the decline of aggregate investment despite capital adjustment cost.

### 1.4.3 Asset price dynamics

In this section, we examine the aggregate asset pricing implications of the model. The amplification of risk prices due to agency frictions enable our model to deliver a sizable risk premium, a low and smooth risk-free interest rate, under a low risk aversion coefficient 4.5 and in the absence of financial leverage. We present our aggregate asset pricing results in panel C of table 1.2.

The well known asset pricing puzzles, such as the equity premium puzzle and the stock market volatility puzzle, are even harder to address in production economies. To deliver a high equity premium requires more than sufficiently volatile risk prices, but also significant firm cash flow risk exposures that are endogenously determined in production economies. It is well understood that capital adjustment cost is a useful modeling ingredient that generates firm cash flow exposures in the production asset pricing literature and our model is no exception. Without adjustment cost or other frictions in the investment process, the owner is able to smooth out any risks in consumption once

[9] Winberry (2018) highlight how time-varying cost of capital affects the relationship between plant-level investment lumpiness and aggregate investment dynamics. Winberry (2018) focuses on the risk-free interest component of cost of capital while we study the effect of risk premium component on aggregate investment.

they have access to a production technology and produce their own consumption. A second channel that contributes to the high equity premium is the labor induced operating leverage channel. In our model, risk sharing contract insures workers against fluctuations in labor earnings and that leaves the residual dividend more exposed to aggregate shocks.

### 1.4.4 The predictability of aggregate investment on future stock market return

In this section, we show that in the model, future market excess returns can be predicted by aggregate investment rate. In the left panel of table 1.3, we regress cumulative future realized market excess returns on contemporaneous aggregate investment rate across different forecasting horizons. Our empirical findings are consistent with results in Cochrane (1991) and Belo and Yu (2013): the $R^2$ rises with forecasting horizons, from 9% at one year horizon to about 45% at the five year horizon. The model implied predictability of future excess return is consistent with the empirical evidence. The slope coefficients in the multi-horizon regressions using model simulated data have the right sign and magnitudes are broadly comparable to those in the real data.

Table 1.3: the predictability of aggregate investment on future excess return

This table presents the predictability of aggregate investment on future market excess returns. We run regression of the type $\sum_{J=1}^{H} r^e_{t\to t+J} = \alpha + \beta \log(\frac{I}{K})_t + e_t$. To construct the investment to capital ratio at the aggregate level, we closely follow Belo and Yu (2013) and a detailed description of data construction is included in the appendix.

| | Data | | | Model | |
|---|---|---|---|---|---|
| Horizon | coefficient | t-stat | $R^2$ | coefficient | $R^2$ |
| 4 | -0.38 | -3.03 | 0.09 | -0.27 | 0.1 |
| 8 | -0.69 | -3.91 | 0.17 | -0.45 | 0.14 |
| 12 | -1.04 | -6.17 | 0.29 | -0.56 | 0.21 |
| 16 | -1.35 | -7.42 | 0.41 | -0.63 | 0.24 |
| 20 | -1.60 | -7.92 | 0.45 | -0.63 | 0.26 |

It is useful to compare the predictability of excess return generated in our model with a standard Q-theory model with convex adjustment cost. Under the Hayashi (1982) conditions, the predictability of market return with aggregate investment rate comes from the well-known identity that equalizes investment rate to Tobin's Q. In our economy, firm-side limited commitment constraint drives a wedge between the marginal cost of investment and the present discounted value of marginal product of capital for each single firm as in equation 1.26. As a result, neo-classical Q theory, which states that marginal cost of investment equals to Tobin's Q, does not hold at both the aggregate level [10] and firm level. The logic behind the predictability in standard Q-theory model does not go through here.

In fact, the predictability comes from the time variation of risk premium. In recessions, investment falls because of lower future marginal benefit and higher cost of capital. Stock market valuation is lower also due to higher cost of capital. As a result, low investment is associated with low stock market valuation and hence higher expected return going forward.

### 1.4.5 the Value Premium

After documenting our model's implications on aggregate asset prices, we now turn to the cross-section of firms and ask if the model can replicate puzzling features of stock return at firm-level. We show that our model generates a sizable value premium. In table 1.4, we report the performance of value investing strategy constructed using simulated data and compare the model implied value premium with the data . We simulate parallel samples with large number of firms. For each sample, we sort firms into five BM ratio groups and construct value-weighted portfolios every year. We then compute one year holding period returns for each portfolio. We find that in our model, a strategy that long high BM ratio firms and short a low BM firms earn an average return which is about 0.26% per month (about 3.12% per year). As a comparison, we compute the five BM sorted portfolio returns using the data available on Ken French's website. We find that in the data, the value strategy generates a monthly average return which is about 0.40% and it is consistent with our model's prediction.

In the model, the difference in firm-level productivity drives the heterogeneity in

[10] By integrating equation 1.26 across all firms, we obtain an equation that equates the price of capital to the sum of average marginal product of capital and the average Lagrangian multiplier associated with firm-side limited commitment constraint.

Table 1.4: the Value Premium

This table presents the model's implications on the cross-section of stock returns (the Value Premium). Panel A reports the Value Weighted 5 Book to Market (BM) ratio sorted portfolio returns. We present monthly returns computed using data available on Ken French's website. The sample period is from 1978 to 2017.

| | Low BM | 2 | 3 | 4 | High BM | 5-1 |
|---|---|---|---|---|---|---|
| | | | Panel A: Data | | | |
| | 0.72% | 0.82% | 0.87% | 0.93% | 1.13% | 0.41% |
| t-stat | 4.53 | 5.29 | 5.87 | 5.94 | 6.11 | 2.89 |
| | | | Panel B: Model | | | |
| return | 0.47% | 0.51% | 0.54% | 0.63% | 0.73% | 0.26% |

firm's BM ratio, investment and expected stock returns. To illustrate the role of firm-level productivity, we rank firms by recent growth rate of productivity and divide firms into 5 groups. We further examine the average BM ratios, investment rate within each group and see if our model's implications on the relations among these characteristics are consistent with existing empirical evidence. Table 1.5 presents our main findings on these relevant characteristics.

Table 1.5: Descriptive Statistics for Productivity Sorted Portfolios in the Model

This table presents descriptive annual statistics of productivity sorted portfolios in the model. We sort firms according to firms' productivity growth in the past year and group them into 10 portfolios. For each variable, averages are first taken over all firms in that portfolio then over years. BM is the ratio of firm's capital stock $K$ dividend by market value of firm $V$. BM is demeaned every year. IK is the sum of quarterly investment $I$ in the past year dividend by the capital stock $K_{-1}$ at the beginning quarter of last year.

| | Low Prod. | 2 | 3 | 4 | High Prod |
|---|---|---|---|---|---|
| BM | 1.54 | 1.02 | 0.89 | 0.81 | 0.75 |
| IK | -0.26 | -0.09 | 0.04 | 0.19 | 0.48 |

Prior ro portfolio formations, low productivity firms tend to be value firms with high BM ratio. They downsize their capital stock to be in accordance with recently fallen productivity levels. Conversely, high productivity group is typically associated with low BM ratio. These firms expand their physical capital stock so as to catch up with recently rising productivities. These results are in line with existing empirical evidence (see Imrohoroglu and Tuzel (2014)) on firms' characteristics prior to BM-sorted portfolio formation. The model's prediction that value firms downsize their capital stock draws a clear distinction between our model and models that rely on costly irreversibility as the main mechanism to generate the value premium.

### 1.4.6 The return spread between investment rate sorted portfolios

Next, we show that our model is able to generate another features of the data that investment negatively predicts stock return at the firm-level. In particular, a large empirical literature (Titman et al. (2004) and Kogan and Papanikolaou (2013)) document that a trading strategy that long low investment rate firms and short high investment rate firms earn a significant positive returns. It turns our that although our model is not calibrated to match this pattern at firm-level, it can deliver similar results as in the data.

Table 1.6: the Investment Rate Sorted Portfolio Returns

This table presents the model's implications on the cross-section of stock returns ( the Investment Rate Sorted Portfolio Returns ). Panel A reports the Value Weighted 5 Investment rate sorted portfolio returns. The definition of Investment rate is capital expenditure (CAPX) dividend by property plant and equipment(PPENT). We exclude finance utility industries and use NYSE breaking points. We sort and form portfolios based on firms' investment rates in the past fiscal year. The sample period is from 1978 to 2017. The sample period is consistent with the period of data we use to compute other firm characteristics.

| | Low INV | 2 | 3 | 4 | High INV | 5-1 |
|---|---|---|---|---|---|---|
| | | | All firms Data | | | |
| return | 0.80% | 0.66% | 0.59% | 0.54% | 0.49% | -0.31% |
| t-stat | 4.51 | 4.43 | 3.81 | 3.41 | 2.42 | -1.76 |
| | | | Model | | | |
| return | 0.62% | 0.58% | 0.55% | 0.51% | 0.46% | -0.16% |

In table 1.6, we first present the empirical evidence on the investment rate sorted portfolio spread. Our definition of investment rate is capital expenditure (CAPX) divided by property plant and equipment (PPENT) following Kogan and Papanikolaou (2013). This measure of investment rate is different from the investment sorted portfolio returns available on Ken French's website because the latter measures firm's asset growth which does not quite fit with our model.

We find that in our model, a strategy that long high investment rate firms and short a low investment firms earns a negative return that is about -0.16% per month (about 1.92% per year). As a comparison, we compute the five IK sorted portfolio returns empirically. We find that in the data, this investment rate strategy generates a monthly average return which is about -0.31% and it is consistent with our model's prediction.

The model is able to generate a significant IK sorted portfolio spread. In the model, the evolution of normalized promised utility $u$ drives both firm's investment decisions and expected return going forward. Consider a firm that is not binding by either side of limited commitment constraints. Following an unfavorable shock, firm cuts investment expenditure dis-proportionally to the size of the shock due to agency frictions. Meanwhile, optimal risk-sharing contract stipulates that the firm should increase the normalized worker's promised utility as insurance. Upon the impact of adverse shock, firm cuts investment expenditure and its value immediately drops because of lower output and higher labor commitment. Therefore, firm investment negatively predicts stock return as shown in table 1.6.

### 1.4.7 Firm-level quantities

Finally, we assess our model's predictions on firm-level quantities. We find that the model's success in fitting the cross-sectional asset pricing moments does not come at the cost of empirically implausible implications for firm-level quantities.

Examining the first two rows of table 1.7, we find that our model can explain the significant cross-sectional variation of firm-level investment. This is noteworthy in itself because existing investment-based models that can quantitatively account for the cross-section of asset returns fail to match the cross-section of firm investment. Investment models featuring rich investment frictions are silent on the cross-section of firm stock returns. Another notable feature of the data is that a non-trivial fraction of firms choose to downsize capital stock every year. This feature of the data can hardly be reconciled in an asset pricing model that relies on costly irreversibility as its main mechanism to match the cross-section of asset returns.

In the third and fourth row of table 1.7, we show that firm-level output moments are similar to the second and third moments of sales growth in the data. This is not

Table 1.7: Implications on firm-level quantities

This table presents the model's implications on firm-level quantities dynamics. We report the cross-sectional standard of firm-level investment, the fraction of firms that undertake negative net investment, the cross-sectional standard deviation of firm-level sales growth, the cross-sectional mean and standard deviation and the Kelley skewness of sales growth in economic downturns. The construction of these firm-level moments is standard and a detailed data explanation is included in the appendix. Since data on capital expenditure in earlier years is very sparse, we set our sample period to start from 1978 until 2017.

| moment | data | model |
|---|---|---|
| $\sigma(\frac{I}{K})$ | 28.50% | 32.06% |
| fraction of negative investment firms | 19% | 29.82% |
| $\sigma(\Delta \log(\text{sales}))$ | 33.40% | 28.01% |
| skewness$(\Delta \log(\text{sales}))$\|recession | -2.12% | -1.15% |
| mean $\frac{W}{Y}$ | 53.13% | 65.11% |
| $\sigma(\frac{W}{Y})$ | 65.50% | 71.66% |

surprising because the higher moments of sales growth are targeted moments so as to discipline our parameterizations of idiosyncratic productivity shock process.

In the last two rows of table 1.7, we report our model's results on firm-level labor share, defined as total labor compensation to firm's output ratio. For the data moment, we compute labor share as the Compustat item Total Staff Expense for labor compensation, normalized firm value added. Compustat item XLR is sparsely populated and we observe missing values under this item for many firm quarter observations. Donangelo (2018) provides reassuring evidence that the statistical property of XLR is similar to other measures of total staff expense of a firm such as industry average wage per employee times the number of employees of a firm. We find that the mean firm-level labor share in the data is about 53.13% and our model overshoots its level a little bit. In addition, firm-level labor share also exhibits significant variation across firms in the data and the cross-sectional variation of labor share in our model is similar to that in the data.

## 1.5 Empirical Analysis

In previous sections, we have demonstrated that our model is consistent with a wide range of empirical facts on asset prices and investment both at the aggregate and firm-level. In this section, we present a set of tests to further illustrate the main mechanisms of our model and to contrast our model with existing literature. We design an empirical procedure that builds on Campbell and Shiller (1988) and Vuolteenaho (2002) and attribute time series variation of aggregate investment as well as cross-sectional variation of firm-level investment to variations in cash flow news and discount-rate news in the data. This procedure sheds light on the key driving forces behind aggregate and firm-level investment both in our model and in the data. This exercise is also informative about dynamic structural models of investment and asset price, which typically predict that variations of investment should be either due to future cash flow news or future discount rate news. Different class of models will have completely different implications on the importance of each type of news to account for movements in investment and therefore our empirical findings provide further guidance and discipline on structural models of investment and asset prices.

### 1.5.1 Investment decomposition framework

The main idea of decomposing the variance of investment is that forward looking investment can either respond to news to cash flow or news to discount rate. Therefore, variations in investment should be driven by these news component. This logic is analytically clear in a neoclassical Q model with convex adjustment cost. We should emphasize that our empirical exercise does not depend on the assumption that marginal Q equals average Q under the Hayashi (1982) conditions. The presentation of the following Q model is only for conveying the main logic behind our empirical tests. Equation 1.27 provides a link between investment and cash flow and discount rate innovations when firm's marginal cost of investment equals its average Q [11] .

$$\eta \ln(\frac{I_t}{K_t}) \approx \text{constant} + \mathbf{E}_t \left[ \ln \frac{D_{t+1}}{K_{t+1}} + \sum_{j=1}^{\infty} (\rho^j \Delta \ln D_{t+j+1} - \rho^{j-1} \ln R_{t+j}) \right] \tag{1.27}$$

[11] see Kogan and Papanikolaou (2012) for detailed derivation of equation 1.27

Equation 1.27 highlights a relation between three endogenous firm variables: firm's investment rate $\frac{I_t}{K_t}$ is positively affected by its future profitability which is expressed as expected dividend to capital stock ratio $\frac{D_{t+1}}{K_{t+1}}$; A firm's investment is also positively related to future cash flow growth $\Delta \ln D_{t+j+1}$ and negatively related with future discount rates $\ln R_{t+j}$. $\rho$ is a parameter that captures average level of stock market prices. $\eta$ is an adjustment cost curvature parameter.

Equation 1.27 holds ex-ante as well as ex-post in this theoretical framework. Therefore, to decompose the variance of investment, it is essential to express changes in expectation of investment rate as a linear combination of revisions in expected future cash flows and returns.

$$\begin{aligned} N_{ik,t} &= \eta \ln \frac{I_t}{K_t} - \mathbf{E}_{t-1} \eta \ln \frac{I_t}{K_t} \\ N_{DR,t} &= (\mathbf{E}_t - \mathbf{E}_{t-1}) \sum_{j=1}^{\infty} \rho^{j-1} \ln R_{t+j} \\ N_{CF,t} &= (\mathbf{E}_t - \mathbf{E}_{t-1}) \left[ \ln \frac{D_{t+1}}{K_{t+1}} + \sum_{j=1}^{\infty} \rho^j \Delta \ln D_{t+j+1} \right] \end{aligned} \tag{1.28}$$

Re-writing equation 1.27 with expressions for news component 1.28, we obtain the following key condition that is crucial to our empirical analysis:

$$N_{ik,t} = N_{CF,t} - N_{DR,t} \tag{1.29}$$

Condition 1.28 which we can consistency condition, says that an increase in expected future cash flow is associated with more investment today, while an increase in expected discount rate leads to less investment today.

### 1.5.2 The VAR method

The empirical implementation of the investment decomposition requires a model that forecasts future cash flows and discount rates. A vector autoregressive (VAR) system provides such a mechanism to forecast these variables and allows us to compute the revisions in expectations of these variables. We decompose the time series variance of aggregate investment as well as cross-sectional variation of firm investment. Therefore, these two exercises require us to estimate two separate VAR models including different set of variables as relevant forecasting variables. We first introduce the general procedure that extracts news components given an estimated VAR system. We then comment on our choices of forecasting variables at both the aggregate VAR and firm-level VAR.

The behavior of aggregate economy or a firm is described by a vector $z_t$ of some relevant state variables. In particular, $z_t = [ik_t, CF_t, r_t, f_t]$. And a vector $z_t$ in total include m variables. The first component of $z_t$ is the log of investment to capital ratio $ik_t$ which could either be aggregate investment to capital ratio for an aggregate VAR model or investment rate for a particular firm for a firm-level VAR model. $CF_t$ corresponds to a measure of cash flow and $r_t$ is the log of stock return. $f_t$ is a vector that include all other variables that help predict investment, cash flow and return.

The vector of $z_t$ evolves according to a first-order VAR:

$$z_{t+1} = c + \Gamma z_t + e_{t+1} \tag{1.30}$$

where $z_{t+1}$ is a m-by-1 state vector. $c$ and $\Gamma$ are m-by-1 constant vector and m-by-m matrix of parameters. The assumption that VAR is first-order is standard as in Vuolteenaho (2002) and Campbell et al. (2013). The qualitative findings of this paper do not change when we estimate VAR models with higher orders.

From the estimated VAR system 1.30, innovations to investment is simply the time series of residuals for the first equation in the VAR system that regresses current period of investment on lagged state variable vector. That is,

$$N_{ik,t+1} = \mathbf{e}_1' u_{t+1} \tag{1.31}$$

where $\mathbf{e}_i$ is a vector whose $i^{\text{th}}$ entry is one with the rest entries being 0.

Campbell (1991) shows that the discount rate news can be computed as following:

$$N_{DR,t+1} = \mathbf{e}_3' \rho \Gamma (I - \rho \Gamma)^{-1} e_{t+1} \tag{1.32}$$

where $\Gamma$ is the VAR coefficient matrix, $e_{t+1}$ is the error term of VAR system and $\mathbf{e}_3$ is a vector where its third element is one and zero otherwise since we put return variable as the third variable in the state vector $z$. And $\rho$ is a parameter that depends on average pd ratio at the aggregate level or firm-level. Previous literature find that setting $\rho$ to be an arbitrary number between 0.95 to 0.99 does not matter for the variance decomposition of stock returns. We experiment with different values of $\rho$ and find similar patterns that different values of $\rho$ does not drive our empirical results at all. The above formula demonstrates that any unexpected shocks to current state variables will be transmitted into returns for all future periods because discount rates are predictable using the VAR model 1.30.

The cash flow news can be indirectly computed by consistency condition 1.29, as the residual of unexpected investment rate $N_{ik,t+1}$ and discount rate news $N_{DR,t+1}$. An alternative approach is to compute cash flow news directly using the VAR system. However, we choose not to do so because earnings or cash flows could come from other business activities in addition to firms' investment. To attribute variation of investment to movement in cash flows driven by other factors than investment itself could potentially contaminate our variance decomposition results.

### 1.5.3 VAR data

The choice of state variables is crucial in empirical implementation of cross-sectional variance decomposition of firm investment. We include variables that are standard to the literature of firm-level return decomposition. In particular, we include firm investment, firm cash flow measures, firm valuation ratio which is the market value of equity to its book value ratio and firm realized stock return.

Our aggregate VAR estimation method involves identifying a set of state variables for the aggregate VAR. Our choice of state variables follow Campbell and Vuolteenaho (2004) and Campbell et al. (2013) closely. Our aggregate VAR includes standard variables to better forecast the dynamics of aggregate investment, cash flow and discount rate. A list of state variables, except aggregate investment, corporate dividend and market excess return, include 1) Price to Earning (PE) ratio. 2) the term yield as the difference between the log yield on the 10-year U.S. bond and short-term treasury yield. 3) default spread. 4) small stock value spread. See appendix for detailed explanation on data construction.

### 1.5.4 Firm-level evidence

We report the firm-level variance decomposition results in Panel A of table 1.8. We first estimate time-series VAR models for each individual firm using all available observations over time. The variance decomposition results demonstrate the contribution of each type of news on the time-series variation of each individual firm. We average VAR coefficients and variance decomposition results across firms. The predictive coefficient for investment on the lagged stock return is negative while the coefficient on two measures of cash flows are both positive. This supports our theoretical prediction that firm's investment is positively related with future cash flow news while and negatively related with discount rate news. In the columns under "decomposition coefficients", we report the coefficients of regressing discount rate news and cash flow news on unexpected investment. We find that cash flow news dominate discount rate news in explaining the time-series variation of firm investment: for example, discount rate news explains about 20% of investment variance and news to ROE explains roughly 80% of investment variance.

Table 1.8: Firm-level investment variance decomposition

This table first reports firm-level investment variance decomposition results in the data. The forecasting variables include firm-level investment, ROE and stock returns. We follow Chen and Zhao (2009) to select firms that are included in our sample and our sample starts from 1978 to 2017. All variables are at the annual frequency. Firm by firm time series refers to the estimation of VAR conducted for each individual firm in our sample. We then report the average decomposition coefficients across firms. This is essentially decomposing the time series variation of a firm's investment which is analogue to the decomposition of time series variation of aggregate investment. We then estimate panel VAR models with firm fixed effect and with the addition of year fixed effect.

| | Decomposition Coefficients | |
|---|---|---|
| | cash flow | discount rate |
| Panel A: Data | | |
| firm by firm time series | 79.28% | 20.72% |
| panel var with firm fixed effect | 94.34% | 5.66% |
| panel var with firm and time fixed effect | 93.68% | 6.32% |
| Panel B: Model | | |
| Benchmark | 88.94% | 11.06% |
| Zhang(2005) | 21.34% | 78.66% |

We then estimate panel VAR models. With the presence of lagged dependent variables as regressors, OLS estimates would be biased even when the number of firms is large (See Nickell (1981)). Therefore, we use a GMM estimators based on Arellano and Bover (1995) for our panel VAR analysis with fixed effects. We first report the panel VAR controlling for the firm fixed effects, which is the same as de-meaning all firm-level variables time-series wise. Again, we find that for both measures of cash flow, firm investment increases when cash flow increases or when firm stock return drops. In addition, when examining the cross-sectional variance decomposition, we find that almost all cross-sectional variation of investment is driven by dispersion in cash flow news while only a small fraction of investment variance is due to discount rate news. We repeat our VAR analysis by controlling for both the firm and time fixed effects and we find very similar results.

We further investigate implications of our model, as well as a stylized investment-based asset pricing model with high adjustment cost such as Zhang (2005), on the variance decomposition of firm-level investment. Panel B of table 1.8 presents these results and draws a clear distinction between our model and adjustment cost models.

In our model, we generate a time series of aggregate productivity shocks with a length of 3000 quarters. Then we sample firm idiosyncratic productivities from the aggregate state-dependent distribution of firm-specific shocks for each time period. The number of firms is set to be 10000. We then simulate firms' optimal investment, compensation and dividend payout policies using relevant policy functions. Similarly, we generate a large sample of firms in Zhang (2005) using the program and simulation procedure available in Lin and Zhang (2013).

We report the model implied investment variance decomposition results. There is no ambiguity on the definition of cash flow in both models and the measure of cash flow is simply dividend payout. In both models, investment is positively related with dividend but negative correlated with realized returns next period. In our framework, as we have shown in previous section, promised value to workers simultaneously affect firms' current investment policy as well as firms' future cash flows because promised utility is forward looking and summarizes all future state-contingent firms' compensation policies. Therefore, an unexpected favorable firm-specific shock raises firm's investment because a favorable shock reduces firm's labor commitment going forward and hence increases firm's dividend payout for multiple period in the future. As a result, in our model, variation in firm investment is almost entirely driven future cash flow news.

On the contrary, applying our method to understand the cross-sectional variation of investment in adjustment cost models such as Zhang (2005), we uncover another empirical irregularity of that class of models with high adjustment cost. Clementi and Palazzo (2018) already documents that the cross-sectional standard deviation of firm investment in that class of model is counter-factually small. In Panel B of table 1.8, we find that show that about 80% of cross-sectional variation of firm investment is driven by news to discount rate. With high adjustment cost, investment is smooth so that volatility of firm investment is low. To make matters worse, high adjustment cost renders firms' cash flows not dispersed enough so that most of the variation in firm investment is explained by dispersion of news to firms stock returns which is not consistent with what data suggests. It is unclear to us whether the failure to account for the composition of investment variance decomposition is a special case that only shows up under a particular parameterization of firms' investment technology as in Zhang (2005). Or such a failure could be common to a large class of models that emphasize the importance of investment adjustment costs to understand stock return anomalies in the cross-section of firms.

### 1.5.5 Aggregate evidence

In the last section, we conclude both in the data and in our model that firm-level investment is largely driven by unexpected news to cash flows. In this section, we examine the aggregate VAR models and identify the key factor that determines the time series variation of aggregate investment.

In table 1.9, we report the variance decomposition results. The first row in the data panel presents the variance decomposition of aggregate investment rate. As a robustness test, we follow Hall (2001), Andrei et al. (2018) to construct an alternative measure of aggregate investment rate. For both of measures, we find that more than 100% of time series variance of aggregate investment is due to discount rate news. This finding is an echo of similar findings in understanding times series variation of stock market returns as summarized in Cochrane (2011). The value of the decomposition coefficients do not have to be between 0 and 100 percent. For example, $-18.21\%$ and $118.21\%$ happens

Table 1.9: Aggregate investment variance decomposition

| | Decomposition Coefficients | |
|---|---|---|
| | cash flow | discount rate |
| **Data** | | |
| investment rate defined as in Belo and Yu (2013) | -18.21% | 118.21% |
| investment rate defined as in Hall (2001) | -7.07% | 107.07% |
| **Model** | | |
| Benchmark | 13.90% | 86.07% |

This table first reports aggregate investment variance decomposition results in the data. The forecasting variables include standard variables as in Campbell et al. (2013). Our sample starts from 1978 to 2017. All variables are at the annual frequency. We construct investment to capital ratio in two different ways, following Belo and Yu (2013) and Hall (2001).

because high investment rate seem to forecast lower real cash flow growth. They must do forecast really low returns and returns must account for more than 100% of aggregate investment variation. This logic is confirmed by the VAR coefficients. The loadings of cash flow growth on lagged aggregate investment rate is negative while the loadings of return on lagged investment is also negative.

In our model however, at the aggregate level, investment and future cash flow growth are positively related. Therefore, the contribution of cash flow news will not be more than 100% as in the data. However, but discount rates news dominates the cash flow news since return news explain more than 80% of variation in aggregate investment. the highly variable firm-level cash flow-news component is largely diversified away in aggregate portfolios. While cash flow information is largely firm specific, expected-return information has a common component that is predominantly driven by aggregate shock.

## 1.6 Conclusion

We present a dynamic agency based asset pricing model with production. We show that risk premia are amplified by agency frictions because agency frictions give rise to endogenously incomplete market. We further explore the the implications of agency frictions on quantities and asset prices both at the aggregate level and in the cross-section of firms. Specially, our model is able to generate a sizable value premium, a large return spread between high-low investment firms without sacrificing its performance on matching the stylized facts on firm-level investment.

Figure 1.3: compensation.

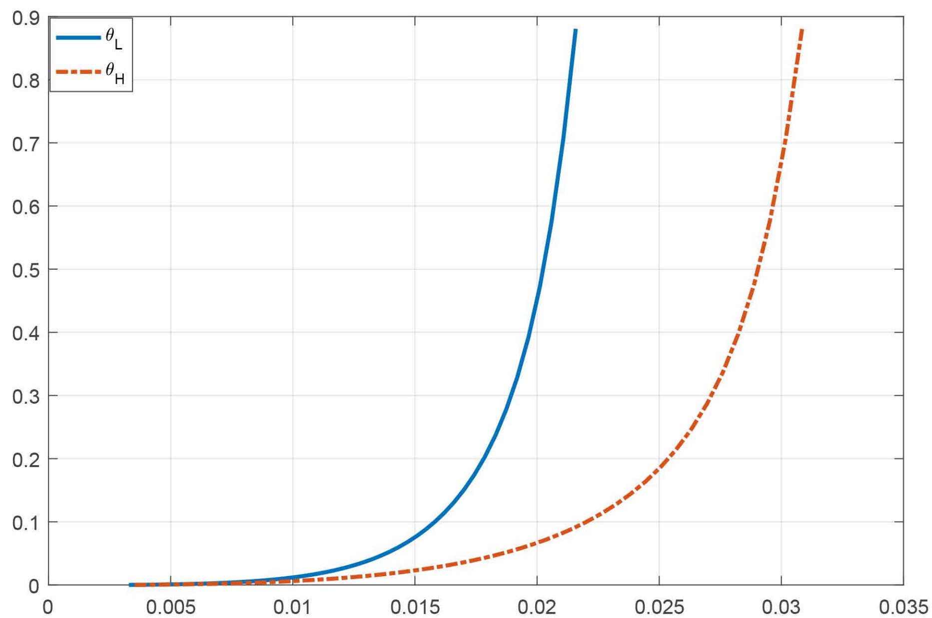
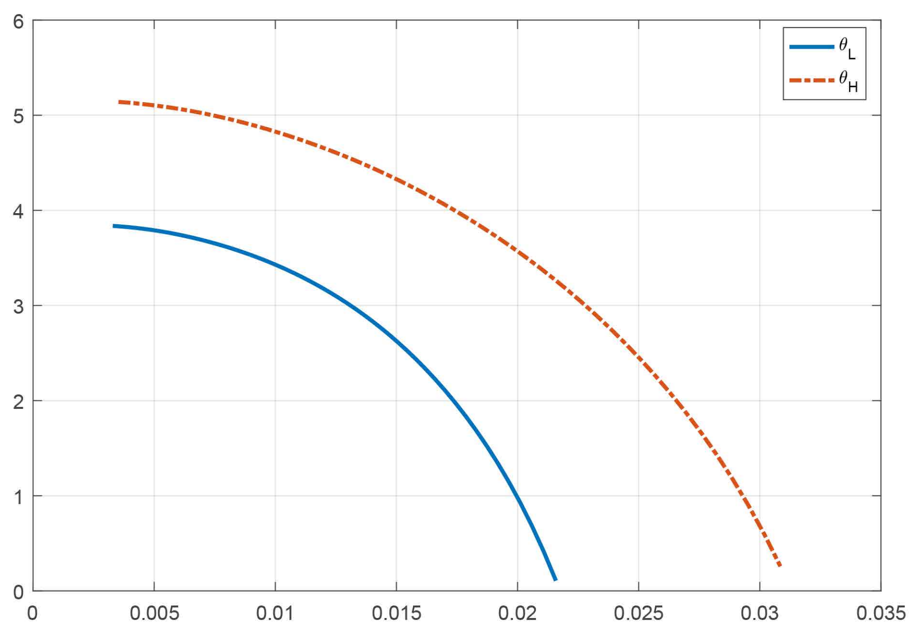


Figure 1.4: value function



# Chapter 2

# Firm Dynamics under Limited Commitment

## 2.1 Introduction

The purpose of this paper is to incorporate assortative matching into the dynamic agency theory of limited commitment to provide a unified theory of the dynamics of firm size, CEO compensation and investment.

Contracting models of limited commitment provides a powerful theoretical framework to study the dynamic features of firm growth (Albuquerque and Hopenhayn (2004)) and CEO compensation (Harris and Holmstrom (1982)). To evaluate the quantitative relevance of the the agency friction of limited commitment, it is important to discipline the choice of the outside options of firms and managers, which are key determinant of the implications of the model. In most models of limited commitment, outside options are taken as exogenous parameters.

The assortative matching model of Gabaix and Landier (2008) provides a convincing account for the power law of CEO pay taking the power law of firm size as given. No attempt so far has been made to integrate the static model of assortative matching into a dynamic equilibrium framework to study time-series dynamics and the cross-sectional distribution of firms jointly.

In this paper, we propose a unified theory of limited commitment and assortative matching. On one hand, assortative matching provides a micro-foundation for the outside options of firms and managers. On the other hand, the power law distribution of firm size, which is taken as exogenous in the Gabaix and Landier (2008) model is jointly determined as an equilibrium outcome in our model. Merging the theory of optimal contracting with limited commitment and assortative matching allows us to evaluate the quantitative implications of theory through a disciplined calibration exercise not only because it ties the exogenous assumptions made in both literature to deeper micro-foundations, but also because in our dynamic model, the cross-section distribution of firms size, investment and CEO compensation will be endogenous outcomes of the time-series dynamics implied by the theory. We demonstrate that our calibrated model is able to account for many stylized facts on both the time-series dynamics and the cross-sectional distribution of firm investment, dividend payout and CEO compensation.

We start by presenting a continuous-time version of the organizational capital model of AK. Because the technology shocks are i.i.d., without agency frictions, Gibrat (1931)'s law holds and the distribution of firm size follows a power law. However, contrary to the data, this model rules out any dependence of investment and growth on firm size. Further, if shareholders are well-diversified and managers are risk averse, then the optimal compensation contract prescribes constant managerial pay, which is inconsistent with the large inequality in CEO compensation in the data.

Our main model features a market where managers and firms meet to form new relationships. In our model, the organization capital of a firm is jointly determined by managerial human capital and firm productivity. As in Gabaix and Landier (2008), the efficient outcome requires assortative matching and implies a power law of CEO compensation. Once a production relationship is formed, firms' investment, CEO compensation and dividend payout policies are determined by optimal contracting subject to limited commitment. Due to the presence of the matching market, firms and managers and voluntarily choose to terminate their relationship and search for new matches. Due to assortative matching, the outside options of firms and managers are increasing functions of their productivity and ability, respectively.

As in Harris and Holmstrom (1982), risk sharing implies that CEO compensation must be constant whenever the limited commitment constraint is not binding. However, a sequence of positive shocks to human capital raises CEO's outside options and a constant compensation contract will eventually lead the manager to voluntarily separate. Because separation is associated with loss of human capital and reduces efficiency, we show that the optimal compensation contract implements a minimum increase in CEO compensation to prevent separation whenever the manager's limited commitment constraint binds. Similarly, because compensation contracts insure managers against productivity shocks, a squence of negative productivity shocks lowers the equity value of the firm. Whenever firm value falls below its outside options, managerial compensation needs to be cut to lower firms' incentive to default.

We solve an equilibrium model with two-sided limited commitment and establish several important results. First, we demonstrate that limited commitment on the manager side translates a power law in firm size into a power law in CEO compensation. In particular, we prove that managerial compensation follows a power law with an exponent that depends on the power law in firm size and a return to scale parameter on the matching market similar to the Gabaix and Landier (2008) model.

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

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

The tradeoff between risk sharing and limited commitment in our model builds on the earlier work of Kehoe and Levine (1993), Kocherlakota (1996), and Alvarez and Jermann (2000). Our specification of managers' outside options is similar to the one in Kiyotaki and Moore (1997), and Albuquerque and Hopenhayn (2004). Several more recent papers also study optimal contracting problems related to ours. Biais et al. (2010a) solve the optimal labor contract in a model with limited commitment and capital structure decisions. Rampini and Viswanathan (2010, 2012) study the implications of limited commitment for risk management and capital structure.[1]

Our paper is more closely related to two recent papers on limited commitment and firm dynamics. Cooley et al. (2013) consider a model with two-sided limited commitment where shareholders cannot commit to compensation plans that provide utility higher than managers' outside options. As a result, managers in their model always receive their outside options. In contrast, we assume that shareholders cannot commit to negative NPV projects. Therefore, the optimal contract in our model allows for risk sharing. Lustig et al. (2011) consider a model with limited commitment on the manager side and study the link between the inequality of CEO compensation and productivity growth. Their model also generates a power law in firm size, but unlike us, they do not characterize the power law in CEO compensation and its relationship with the power law in firm size. In addition, different from Lustig et al. (2011), in our model, investment decisions are endogenously determined by the optimal contract, which allows us to explore the implications of limited commitment for the cross-sectional distribution and life-cycle dynamics of firms' investment and growth.

This paper is related to the large literature on agency frictions and managerial

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

compensation. Edmans and Gabaix (2016) provide a comprehensive review of the earlier literature; more recent papers include Biais et al. (2010b), and Bond and Axelson (2015). Our paper is also related to the literature on firm dynamics and power law in economics and finance (see Gabaix (2009) and Luttmer (2010) for a survey). The neoclassical model without frictions considered in our paper is essentially an interpretation of the model in Luttmer (2007). Terviö (2008), and Gabaix and Landier (2008) are assortative matching models that link CEO compensation to firm size taking size distribution as given. Our model provides an alternative, mechanism-design based explanation of the level of CEO pay and its dependence on firm size. In our model, both the distribution of firm size and CEO compensation are endogenous outcomes of the optimal contract.

The continuous-time methodology of our paper builds on the fast growing literature on continuous-time dynamic contracting, for example, Sannikov (2008), DeMarzo and Sannikov (2006), DeMarzo and Sannikov (2012), He (2009), Biais et al. (2010a).[2] The optimal contracting design in our paper is related to the one in Ai and Li (2015), and Bolton et al. (2016).

The rest of the paper is organized as follows. In Section 2.2, we set up a general equilibrium model with limited commitment. In Section 2.4, we consider a frictionless Arrow-Debreu economy and discuss its limitations. In Section 2.5, we introduce and discuss the implications of the manager-side and the shareholder-side limited commitment separately. Our full model with two-sided limited commitment is presented and analyzed in Section 2.6. We calibrate the model and evaluate its quantitative implications in Section 2.7. Concluding remarks are provided in Section 2.8.

## 2.2 Model setup

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

### 2.2.1 Production technology

Time is continuous and infinite. There is a continuum of firms indexed by $j$. As in Atkeson and Kehoe (2005), the output of firm $j$, denoted $y_j$, is produced from organization

[2] For an excellent survey of this literature see Biais et al. (2004).

capital ($Z_j$), physical capital ($K_j$), and labor ($N_j$) using a standard Cobb-Douglas production technology: $y_j = Z_j^{1-\nu}\left(K_j^{\alpha}N_j^{1-\alpha}\right)^{\nu}$, where $\nu$ is the span-of-control parameter.[3]

The operating profit of firm $j$ is defined as

$$\pi\left(Z_j\right) = \max\left\{Z_j^{1-\nu}\left(K_j^{\alpha}N_j^{1-\alpha}\right)^{\nu} - MPK\cdot K_j - MPL\cdot N_j\right\}, \tag{2.1}$$

where $MPK$ is the rental rate of physical capital and $MPL$ is the equilibrium wage for unskilled workers. We assume that the unskilled labor, $N$ and physical capital, $K$ are not firm-specific and can be hired on perfectly competitive markets. As a result, the constant return to scale (CRS) of the production function implies that the profit function is linear in organization capital, i.e., $\pi\left(Z_j\right) = AZ_j$, where $A$ is the equilibrium marginal product of organization capital. Due to the CRS technology, the optimal choice of $K_j$ and $N_j$ is proportional to $Z_j$. Hence, in our model, all three are equivalent measures of firm size.

The organization capital $Z_j$ is firm-specific, and we assume that the accumulation of organization capital depend on manager's investment:

$$dZ_{j,t} = Z_{j,t}\left[\left(i_{j,t}-\delta\right)dt + \sigma^T dB_{j,t}\right], \tag{2.2}$$

where $i_{j,t} = \frac{I_{j,t}}{Z_{j,t}}$ is the investment-to-organization capital ratio, $\delta$ is the depreciation rate of organization capital, $dB_{j,t}$ is a firm-specific Brownian motion shock, and $\sigma$ is the sensitivity of $Z$ with respect to Brownian motion shocks. As in Atkeson and Kehoe (2005), and Luttmer (2011), firm dynamics are driven by the accumulation of organization capital. The cost of investment in organization capital is specified by a standard quadratic adjustment cost: $h\left(\frac{I}{K}\right)K$, where $h\left(i\right) = i + \frac{1}{2}h_0\cdot i^2$ with $h_0 > 0$.

### 2.2.2 Entry, exit, and matching

Operating a firm requires a manager, who are the only agents in the economy that can efficiently build up organization capital for firms. A measure one of firms and a measure one of managers arrive at the economy per unit of time. The initial level of human capital of managers and the initial level of firms are both normalized to one.

[3] Organization capital is firm-specific knowledge that makes physical capital and labor more productive. Examples of organization capital include corporate culture, team work, firm-specific human capital, etc. (see Kydland and Prescott (1982)).

Newly arrived firms are idle firms and need to be matched with a manager to produce output.

At Poisson rate $\kappa_D$, an operating firm and its manager are hit by a death shock and exit the economy. Once hit by the death shock, the organization capital of the firm and the human capital of the manager evaporates. In addition, at Poisson rate $\kappa_S$, an operating firm-manager pair exogenously separate. Upon separation, the firm retain a fraction $\lambda \in (0,1)$ of the organization capital as firm-specific organization capital and become an idle firm. The manager retain a fraction $\lambda$ of the organization capital as manager-specific human capital. Formally, let $\tau_S$ denote the stopping time of separation,

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

where $Y$ denote the firm-specific organization capital of an idle firm and $X$ denote the human capital of a unmatched manager.

Idle firms, including the new arrivals and the separated firms, meet intantaneously with managers on a centralized market, where idle firms offer competititive contracts to managers to form productive firms. The inital level of the organization capital of a newly formed firm-manager pair depends on the firm-specific organziation capital, $Y$ and the manager-specific human capital $X$ through a Cobb-Douglas production function:

$$Z_\tau = Y_\tau^{\psi_Y} X_\tau^{\psi_X}, \quad \text{with } \psi_Y, \psi_X \in (0,1) \text{ and } \psi_Y + \psi_X \leq 1, \tag{2.3}$$

where $\tau$ denotes the stopping time at which a match is formed and the new firm starts operation.

### 2.2.3 Preferences

Let $\{C_s\}_{s=t}^{\tau_D \wedge \tau_S}$ denote the compensation contract offered by a firm to a manager from time $t$ until separation, where $\tau_D \wedge \tau_S$ stands for $\min\{\tau_D, \tau_S\}$. Given the contract, the manager evaluate his utility under the contract using a standard CRRA utility. We denote the continuation utility of the manager at time $t$ as:

$$U_t = \left\{ E_t \left[ \int_0^{\tau_D \wedge \tau_S} (r+\kappa)\, e^{-r(s-t)} C_s^{1-\gamma} ds + e^{-\tau_S} U^M \left(\lambda Z_{\tau_S}\right)^{1-\gamma} \right] \right\}^{\frac{1}{1-\gamma}}. \tag{2.4}$$

In the above expression, we use $U(X)$ to denote the maximum utility a manager can obtain upon separation by going to the matching market to search for an idle firm. The term $U(\lambda Z_{\tau_S})$ reflects the fact that the manager and keep $\lambda$ fraction of the organization capital as human capital upon separation. Our formulation (2.4) normalizes managers' utility to be homogenous of degree on in consumption for tractability.

The firm evaluate the present value of the cash flow at discount rate $r$. As is common in the dynamic contracting literature, it is convenient to use promised utility as the state variable. We denote $V(Z_t, U_t)$ as the present value of the cash flow to a firm with current level of organization capital $Z_t$ and with promised utility to the manager, $U_t$:

$$V(Z_t, U_t) = E_t \left[ \int_s^{\tau_D \wedge \tau_S} e^{-r(s-t)} \left[ AZ_s - C_s - h\left(\frac{I_s}{Z_s}\right) Z_s \right] ds + e^{-\tau_S} V^M(\lambda Z_{\tau_S}) \right], \tag{2.5}$$

where the term $V^M(\lambda Z_{\tau_S})$ reflects the fact that upon separation, the firm retains $\lambda$ fraction of organization capital and therefore obtain a maximum value on the matching market of $V^M(\lambda Z_{\tau_S})$.

### 2.2.4 Limited Commitment

The key agency friction in our model is two-sided limited commitment. We assume that both firms and managers can chose to separate at any time and go to the matching market to search for a new match. This requires that before separation, $t < \tau_S$, the continuation utilility of managers must be higher than what they can obtain from separation:

$$U_t \geq U^M(\lambda Z_t), \quad \forall t \leq \tau_S. \tag{2.6}$$

Similarly, the continuation value of a firm must be higher than their outside option before $\tau_S$:

$$V(Z_t, U_t) \geq V^M(\lambda Z_{\tau_S}), \quad \forall t \leq \tau_S. \tag{2.7}$$

Here, we use the fact that perfect competition on the matching market implies zero profit for firms.

Because both firm value and manager utility is strictly increasing, if one of the inequalities in (2.6) and (2.7) holds is strict, the firm can always reallocate the cash flow to make both parties better off to prevent separation. Therefore, under the optimal contract, both (2.6) and (2.7) must hold with equality at $t = \tau_S$.

### 2.2.5 Matching

Idle firms offer competitive contracts to workers to form productive firms. Competition on the matching market determines the outside options of firms and workers. Consider an idle firm with organization capital $Y$. Because the initial organization capital of firm is given by (2.3), if matched with a manager with human capital $X$, the value of the firm is $V\left(Y^{\psi_Y}X^{\psi_X}, U\right)$, where $U$ is the inital promised utility to the manager. Because $U^M(X)$ is what 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:

$$Y(X) \in \arg\max_{Y} V\left(Y^{\psi_Y}X^{\psi_X}, U^M(X)\right). \tag{2.8}$$

In equilibrium, firms outside options are determined by the value they can obtain on the matching market:

$$V^M(Y(X)) = V\left(Y(X)^{\psi_Y}X^{\psi_X}, U^M(X)\right). \tag{2.9}$$

## 2.3 Equilibrium

### 2.3.1 Equilibrium concept

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

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 (2.2) and the law of motion of $U_t$ by specifying its sensitivity with respect to the Brownian motion shocks, $G(Z,U)$.[4] To save notation, we use

[4] Given $G(Z,U)$, equation (2.4) implies that the law of motion of $U$ can be constructed as: $dU = \left[-\frac{\beta+\kappa}{1-\gamma}\left(C^{1-\gamma}U^{\gamma} - U\right) + \frac{1}{2}\frac{G(K,U)^2}{U}\right]dt + G(K,U)\,dB$. This formulation is similiar to the representation in Sannikov (2008), except that we use a monotonic transformation so that utility is measured in consumption units.

$E^*$ to denote the expectation operator under the law of motion of the state variables implied by the optimal contract.

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, $Y(X)$ that satisfy the following conditions:

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

2. Given the equilibrium outside option, $U^M(X)$, the optimal contract maximizes firm value (2.5) subject to the limited commitment constraint (2.6) and (2.7), where the initial condition is determined by $Z_\tau = Y(X_\tau)^{\psi_Y} X_\tau^{\psi_X}$ and $U_\tau = U^M(X_\tau)$.

3. Given the equilibrium outside option of managers, $U^M(X)$, the equilibrium matching rule satisfies (2.8).

4. The outside options of firms are determined by the optimal matching rule: (2.9).

5. The matching market clears. That is, firms and managers match instantaneously upon birth or separation.

### 2.3.2 Assortative matching

We conjecture and later verify that the equilibrium features assortative matching. That is, more productive firms are matched with managers with high human capital. Because the law of motion of firm organization capital and manager human capital are identical and there initial conditions are both normalized to one, assortative matching implies a simple matching rule: $Y(X) = X$.

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

$$V_Z\left(X^{\psi_Y+\psi_X}, U^M(X)\right)\psi_X X^{\psi_Y+\psi_X-1} + V_U\left(X^{\psi_Y+\psi_X}, U^M(X)\right)\frac{d}{dX}U^M(X) = 0. \tag{2.10}$$

Given the functional form of $V(Z,U)$, the above equation determines the slope of managers' outside options, $U^M(X)$.

Because the production technology of initial organization capital is strictly concave, we conjecture that there exists a minimum level of $Z$, $Z_{MIN}$ such that it is optimal to separate once $Z_t \leq Z_{MIN}$. We assume that at $Z_{MIN}$, separated firm earns zero profit:

$$V_Z\left((\lambda Z_{MIN})^{\psi_Y+\psi_X}, U^M(\lambda Z_{MIN})\right) = 0, \tag{2.11}$$

which together with equation (2.10), determines the equilibrium outside option of managers.

### 2.3.3 Special case: $\psi_Y + \psi_X = 1$

In this case, the optmal matching rule implies

$$V_Z\left(X, U^M(X)\right)\psi_X + V_U\left(X, U^M(X)\right)\frac{d}{dX}U^M(X) = 0.$$

We conjecture that $U^M(X) = u_{MAX}X$ is linear. Note that $V(Z,U) = v\left(\frac{U}{Z}\right)Z$. Therefore,

$$V_Z = v(u) + v'(u) - \frac{U}{Z^2}Z = v(u) - uv'(u); \quad V_U = v'(u).$$

The above is written as:

$$\psi_X\left[v(u_{MAX}) - u_{MAX}v'(u_{MAX})\right] + v'(u_{MAX})u_{MAX} = 0$$

That is,

$$\psi_X v(u_{MAX}) + \psi_Y v'(u_{MAX})u_{MAX} = 0.$$

The workers' outside option is therefore $\lambda u_{MAX}Z$ and the firms' outside option is $\lambda v(u_{MAX})Z$.

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

$$C(K,U) = c(u)K; \qquad I(K,U) = i(u)K. \tag{2.12}$$

The normalized value function can be characterized by the solution to an ordinary differential equation (ODE) with appropriate boundary conditions. Under Specification 2, the homogeneity property no longer holds because managers' outside options are no longer linear in firm size. However, we will continue using the state pair $(u, K)$ to facilitate comparison across the two specifications.

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

Second, health shocks make managers effectively less patient than shareholders because they evaluate utility using an effective discount rate of $\beta + \kappa$. However, because the cash flows completely evaporate after managers are hit by health shocks, shareholders value firm's cash flow with the same discount rate of $\mathbf{r} + \kappa$, i.e., $V(K_t, U_t) = E_t\left[\int_0^\infty e^{-(\mathbf{r}+\kappa)s} D_{t+s} ds\right]$. From the optimal contract design perspective, our model should therefore be interpreted as one where the principal and the agent have the same discount rate because $\beta = \mathbf{r}$.[5]

## 2.4 The First-Best Case

In the first-best case, shareholders maximize the present value of firm's cash flow subject to the manager's participation constraint in equation (2.9). Using the budget constraint in equation (2.7), the present value of cash flow can be written as:

$$E_0\left[\int_0^\infty e^{-(\mathbf{r}+\kappa)t}\left(\mathbf{A}K_t - h\Big(\frac{I_t}{K_t}\Big)K_t\right)dt\right] - E_0\left[\int_0^\infty e^{-(\mathbf{r}+\kappa)t} C_t dt\right]. \tag{2.13}$$

Note that the participation constraint affects only the choice of managerial compensation in the second term. As a result, the profit maximization problem is separable and can be solved in two steps. The first step is to maximize the total value of the firm in the first term of equation (2.13) by choosing the optimal investment policy. The second

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

step is to select the optimal managerial compensation to minimize the cost subject to the manager's participation constraint.[6]

The firm value maximization problem in the first step is standard as in Hayashi (1982). The solution to the cost minimization problem is also straightforward: risk aversion of the manager and the fact that the principal and the agent have identical discount rates imply a constant consumption of the manager: $C_t = \bar{U}$ for all $t$.[7] We make the following assumptions to guarantee that firm value is finite and the maximization problem is well defined.

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

$$\mathbf{A} > \mathbf{r} + \delta + \kappa > \frac{-1 + \sqrt{1 + 2h_0\mathbf{A}}}{h_0} \tag{2.14}$$

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

**Proposition 3.** *The first-best Case*

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

$$V(K, U) = \bar{v}K - \frac{1}{\mathbf{r} + \kappa}U, \tag{2.15}$$

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

$$\hat{\imath} = \arg\max_{i<\hat{r}} \frac{\mathbf{A} - h(i)}{\hat{r} - i} = \hat{r} - \sqrt{\hat{r}^2 - \frac{2}{h_0}(\mathbf{A} - \hat{r})}, \tag{2.16}$$

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

*Proof.* See the online appendix of Ai et al. (2018). □

The first term $\bar{v}K = h'(\hat{\imath})K$ in equation (2.15) is the firm value in the neoclassical model with capital adjustment costs. The second term is the present value of the cost of managerial compensation. In the absence of aggregate uncertainty, perfect risk sharing implies a constant managerial compensation, the resent value of which is simply given by the Gordon (1959)'s formula: $\frac{1}{\mathbf{r}+\kappa}U$.[8]

---

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

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

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

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

$$\mathbf{A} = \alpha \mathbf{z} \left(\frac{1}{\mathbf{K}}\right)^{1-\alpha} = \alpha \mathbf{z} \left(\frac{\kappa + \delta - \hat{\imath}}{\bar{K}}\right)^{1-\alpha}. \tag{2.17}$$

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

$$H'\left(\bar{K}\right) = \bar{v} = 1 + h_0 \hat{\imath}. \tag{2.18}$$

The three equations (2.16), (2.17), and (2.18) jointly determine three equilibrium quantities: $\mathbf{A}$, $\hat{\imath}$, and $\bar{K}$. Given them, we can construct the equilibrium wage as the marginal product of labor, $\mathbf{W} = (1-\alpha)\, \mathbf{z} \mathbf{K}^{\alpha}$, and the aggregate consumption of the representative shareholder is the total output less the cost of investment.[9]

Equation (2.16) in Proposition 3 shows that the investment-to-capital ratio in this economy is constant across firms. As a result, Gibrat's law holds, growth rates are i.i.d. across firms and the distribution of firm size follows a power law as in Luttmer (2010), which is summarized in the following proposition.

**Proposition 4.** *Power Law of Firm Size*

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

$$\phi\left(K\right) = \begin{cases} \frac{1}{\sqrt{\left(\hat{\imath}-\delta-\frac{1}{2}\sigma^2\right)^2 + 2\kappa\sigma^2}} \bar{K}^{-\xi} K^{\xi-1} & K \geq \bar{K} \\ \frac{1}{\sqrt{\left(\hat{\imath}-\delta-\frac{1}{2}\sigma^2\right)^2 + 2\kappa\sigma^2}} \bar{K}^{-\eta} K^{\eta-1} & K < \bar{K}, \end{cases} \tag{2.19}$$

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

[9] We need to make assumptions on the primitive parameters of the model so that the assumption in Proposition 3 is satisfied and $V\left(\bar{K}, \bar{U}\right) \geq H\left(\bar{K}\right)$ to ensure a positive entry in equilibrium. Although these assumptions are not here listed to conserve space, we make sure that they are satisfied in our calibration.

*Proof.* See the online appendix of Ai et al. (2018). □

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

## 2.5 One-Sided Limited Commitment

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

### 2.5.1 Limited Commitment on the Manager Side

In the case of the lack of commitment on the manager side, shareholders maximize firm value subject to managers' time-zero participation constraint (equation (2.9)), and the requirement that managers' continuation utility must be higher than their outside options at all times. Our main objective in this section is to show that limited commitment on the manager side generates a power law in managerial compensation and to relate the power law in CEO pay to the power law in firm size.

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

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

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

**Assumption 2.**

$$\mathbf{A} > \mathbf{r} + \kappa + \delta > \hat{\imath}, \tag{2.20}$$

*and*

$$\frac{\mathbf{A} - \hat{\imath}}{\mathbf{r} + \kappa + \delta - \hat{\imath}} - \frac{\nu\gamma}{(\mathbf{r} + \kappa)(\zeta_1 - 1)} \zeta_1^{\frac{\nu\gamma - 1}{\nu(\gamma - 1)}} (\zeta_1 - (1 - \gamma))^{\frac{-\nu\gamma}{\nu(\gamma - 1)}} \varpi^{\frac{1}{\nu}} \geq 1, \tag{2.21}$$

*where*

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

Inequality (2.20) imposes a lower bound on the marginal product of capital, $\mathbf{A}$. Inequality (2.21) is a restriction on the magnitude of managers' outside options, $\varpi$, which is a measure of the severity of agency frictions. Together, they guarantee that the technology is productive enough and that the agency frictions are not too large so that it is always optimal for firms to invest at the maximum rate of $\hat{\imath}$. The above assumptions allow us to simplify firms' investment decisions and focus on the implications of the model for power laws in firm size and CEO compensation. We summarize our main results in the following proposition.

**Proposition 5.** *Power Law in CEO Compensation*

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

$$C_t = \max\left\{\hat{c} \max_{0 \leq s \leq t} K_s^{\nu},\ C_0\right\}, \tag{2.22}$$

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

2. *The right tail of CEO compensation obeys a power law with a slope coefficient of $\frac{\xi}{\nu}$ with $\xi$ being defined in Proposition 4.*

*Proof.* See the online appendix of Ai et al. (2018). $\square$

In the model with limited commitment on the manager side, the compensation contract is downward rigid, as in Harris and Holmstrom (1982). Compensation has to increase to match the manager's outside option whenever the limited commitment constraint binds. Otherwise, due to risk sharing, it must remain constant. Because the manager's outside option is an increasing function of firm size, the above dynamics imply that managerial compensation must be an increasing function of the running maximum of firm size.[10]

Due to the special form of the adjustment cost function, firm investment rate is constant and Gibrat's law holds. As a result, Proposition 4 applies and firm size follows a power law with slope $\xi$. It is straightforward to show that if the distribution of $K$ follows a power law with slope coefficient $\xi$, the distribution of $K^{\nu}$ obeys a power law with slope coefficient $\frac{\xi}{\nu}$. By part 1 of Proposition 5, managerial compensation is a linear function of the running maximum of $K_t^{\nu}$. Intuitively, the running maximum of a power law process obeys a power law with the same slope coefficient. Therefore, managerial compensation in our model follows a power law with slope $\frac{\xi}{\nu}$. Proposition 5 thus links the power law in CEO pay to the power law in firm size and the elasticity of CEOs' outside options with respect to firm size. In our calibration exercise, we show that this relationship generalizes to the case with smooth adjustment costs, where Gibrat's law does not hold.

Under Specification 1, because managers are allowed to participate in the labor market after default, the constant return to scale of the production function implies that the profit function is linear in size. Therefore, the manager's outside option is also linear in $K$. In this case, Proposition 5 implies firm size and CEO compensation have the same power law, $\xi$. Under Specification 2, after default, the manager can only use his home labor to produce home consumption goods but is not allowed to participate in the labor market. In this case, total output after default is $\mathbf{z}K^{\alpha}\bar{n}^{1-\alpha}$ and the manager's outside option is proportional to $K^{\alpha}$. Because managers are excluded from the labor market after default, capital is less scalable and the distribution of CEO pay has a thinner tail, with a slope of $\frac{\xi}{\alpha}$.[11] Hence, qualitatively, Specification 2 is able to account for the observed differences in the right tails of the distributions of managerial

[10] See also Lustig et al. (2011), Grochulski and Zhang (2011), and Miao and Zhang (forthcoming).

[11] Recall that under Specification 2, $\nu = \alpha$.

compensation and firm size.

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

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

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

Finally, we note two main differences between our framework with one-sided commitment and Albuquerque and Hopenhayn (2004). First, the Albuquerque and Hopenhayn (2004) model features decreasing return to scale technologies and stationary productivity. As a result, firms eventually reach their optimal size, where the limited-commitment constraint does not bind. Therefore, neither the distribution of firm size nor that of the managerial compensation in their model have fat right tails. Second, managers are risk-neutral in Albuquerque and Hopenhayn (2004). Therefore, they receive no payment as long as the limited commitment constraint binds, and payment policy is undetermined once the firm grows out of the constraint.

### 2.5.2 Limited Commitment on the Shareholder Side

In the model with limited commitment on the shareholder side, shareholders maximize firm value subject to managers' time-zero participation constraint (equation (2.9)) and the requirement that firm value must stay non-negative at all times. Because the objective function and the constraints are linear in size, the value and policy functions are homogenous in $K$ and it is convenient to work with the normalized value and policy functions defined in equation. The normalized value function $v(u)$ is strictly decreasing in the normalized continuation utility of the manger. Intuitively, as we use more of the firm's cash flows to support a higher continuation utility of the manager, shareholder value declines. If we define $u_{MAX}$ to be the highest normalized utility that can be supported by the optimal contract without violating shareholder's limited-commitment constraint, that is,

$$u_{MAX} = \sup \{u : v(u) \geq 0\}, \tag{2.23}$$

then under the optimal contract the shareholder constraint, $E_t\left[\int_0^\infty e^{-(\mathbf{r}+\kappa)t+s} D_{t+s} ds\right] \geq 0$, is equivalent to $u_t \leq u_{MAX}$. We describe two properties of the normalized value function and the normalized investment policy but relegate the details of the optimal contract to the online appendix of Ai et al. (2018).

1. The normalized value function, $v(u)$, is strictly decreasing and concave on $(0, u_{MAX}]$ with $v(u_{MAX}) = 0$ and $\lim_{u \to 0} v(u) = \bar{v}$, where $\bar{v}$ is the firm value in the first-best case defined in Proposition 3.

2. The investment-to-capital ratio, $i(u)$, is strictly increasing in $u$ and $\lim_{u \to 0} i(u) = \hat{\imath}$, where $\hat{\imath}$ is the first-best investment level defined in equation (2.16).

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

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

## 2.6 Two-Sided Limited Commitment

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

### 2.6.1 Optimal Contract with Two-sided Limited Commitment

The optimal contract maximizes firm value subject to managers' time-0 participation constraint (equation (2.9)), the shareholder-side limited-commitment constraint, and the manager-side limited-commitment constraint. Note that the limited-commitment constraint on the manager side requires $U_t \geq \varpi K^{\nu}$, or $\frac{U_t}{K_t} \geq \varpi K_t^{\nu-1}$ for all $t$. This motivates the following definition of the lower boundary of the normalized continuation utility: $u_{MIN}(K) = \varpi K^{\nu-1}$. The manager-side limited commitment constraint can therefore be written as $u_t \geq u_{MIN}(K_t)$. Note that in Specification 1, $u_{MIN}(K) = \varpi_1$ and does not depend on $K$. Similarly, generalizing equation (2.23), we can also define an upper bound for the normalized continuation utility (that is, the highest normalized utility of the manager that can be supported by the optimal contract):

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

The properties of the optimal contract are summarized in the following proposition.[12]

**Proposition 6.** *Two-Sided Limited Commitment*

1. *Under the optimal contract, $u_{MIN}(K_t) \leq u_t \leq u_{MAX}(K_t)$ for all $t$. In the domain $\left\{(K,U) : u_{MIN}(K) \leq \frac{U}{K} \leq u_{MAX}(K)\right\}$, the value function satisfies an HJB equation together with the associated boundary conditions. Given the value function, the optimal CEO compensation and optimal investment can be constructed from the optimality conditions.* [13]

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

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

[12] A recent paper by Bolton et al. (2016) shows how similar contacts can be implemented by corporate liquidity and risk management policies.

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

*on $K$. The normalized compensation policy satisfies*

$$\ln c\left(u_t\right) = \ln C_0 - \ln K_t + l_t^+ - l_t^-, \tag{2.24}$$

*where $\left\{l_t^+, l_t^-\right\}_{t=0}^{\infty}$ are the minimum increasing processes such that $c\left(u_{MIN}\right) \leq c\left(u_t\right) \leq c\left(\varpi_1\right)$ for all $t$.*[14] *The optimal investment, $i\left(u\right)$, is a strictly increasing function of $u$.*

*Proof.* See Appendix of Ai et al. (2018). □

### 2.6.2 Domain of the State Variables

In Figure 2.1, we plot the boundaries of the normalized utility, $u_{MIN}\left(K\right)$ (dashed line) and $u_{MAX}\left(K\right)$ (solid line), for Specification 1 in the top panel and Specification 2 in the bottom panel. For convenience, both normalized utility and firm size are plotted in log units. Under Specification 1, because the value function and policy functions are homogenous of degree one in $K$, the boundary for the normalized utility do not depend on $K$, as shown in Panel (a) of Figure 2.1. Intuitively, $u_{MAX}\left(K\right)$ is the highest level of the normalized utility that can be supported by the optimal contract. Higher levels of $u$ result in negative firm value. The lower boundary, $u_{MIN}\left(K\right)$, is the lowest level of the normalized utility that is required to retain the manager and to prevent him from taking the outside options. Note that under Specification 1, it is never optimal for the shareholder to default because in that case capital cannot be productively deployed. Also, in equilibrium, the manager never takes the outside option because it destroys risk-sharing. Therefore, under the optimal contract, the normalized continuation utility always stays within the two boundaries.

The boundaries $u_{MIN}\left(K\right)$ and $u_{MAX}\left(K\right)$ for Specification 2 are plotted in Panel (b) of Figure 2.1. Recall, that in this specification, the minimum promised utility required to match managers' outside option is $\varpi_2 K^\alpha$. Therefore, $\log u_{MAX}\left(K\right) = \log\left(\frac{\varpi_2 K^\alpha}{K}\right) = \log \varpi_2 - (1-\alpha)\log K$, represented by the dashed line, has a negative slope of $1-\alpha$. As firm size gets smaller, managers' outside option shrinks at a lower rate than firm's cash

[14] *Formally, $l^+$ and $l^-$ are the unique pair of continuous and nondecreasing processes that satisfy the following conditions: i) $l_0^+ = l_0^+ = 0$, ii) $\ln c\left(u_{MIN}\right) \leq \ln C_0 - \ln K_t + l_t^+ - l_t^- \leq \ln c\left(\varpi_1\right)$ for all $t$, iii) $l^+$ increases only when $\ln C_0 - \ln K_t + l_t^+ - l_t^- = \ln c\left(u_{MIN}\right)$ and $l^+$ increases only when $\ln C_0 - \ln K_t + l_t^+ - l_t^- = \ln c\left(\varpi_1\right)$. For alternative equivalent constructions of two-sided regulators, see Harrison (1985).*

flow. Eventually, as $K \searrow K_{MIN}$, the cash flow is not enough to support managerial pay and it is optimal to shut down the firm and let the manager take his outside option. As a result, Specification 2 allows for endogenous firm exit. Intuitively, as $K$ gets smaller, the home production technology dominates the market technology and it is optimal to dissolve the manager-firm match. Hence, $K_{MIN}$ is the smallest firm size in this economy. At $K_{MIN}$, the net cash flow of the firm is actually negative and the firm value is zero (note that the firm value is still non-negative because is incorporates the option value of future growth). A further decline in capital lowers the firm value and results in the optimal shut-down of the firm. Therefore, Specification 2 implies that small firms exit markets more frequently relative to large firms.

As $K$ increases, the market production technology becomes more efficient than the home technology and $u_{MAX}(K)$ increases. In the limit, as $K \to \infty$, $u_{MAX}(K)$ converges to the dash-dotted line, which is the upper bound on the normalized utility for the case with the shareholder-side limited commitment only. Intuitively, as firm size increases, the manager's outside option becomes negligible compared with the firm's cash flow, and limited commitment on the manager side almost never binds.

### 2.6.3 Value Functions

The normalized value function, $v(u, K)$, for the two-sided limited commitment model is plotted in Figure 2.2 and is represented by a solid line.

For the purpose of comparison, in the same figure, we also plot the value function for the first-best case (dash-dotted line) and that for the case with the shareholder-side limited commitment only (dashed line).[15]

The top panel corresponds to Specification 1, under which $v(u)$ does not depend on $K$ and is defined over $u \in [u_{MIN}, u_{MAX}]$. In the first-best case, the normalized value function is linear with slope $-\frac{1}{\mathbf{r}+\kappa}$ on its domain $(0, \infty)$, as shown in equation (2.15). In the case of limited commitment on the shareholder side only, $u$ is bounded from above by the limited commitment constraint on shareholders. As $K \to \infty$, because optimal

[15] For illustrative purposes, in Figures 2.1 and 2.2, we assume that the marginal product of capital is the same across all economies. Hence, the comparison between the first-best case and cases with limited commitment is a partial equilibrium one. In general equilibrium, fixing preference and technology parameters of the model and adding limited commitment will result in an endogenous change in the steady-state level of capital and, therefore, a different marginal product of capital.

risk sharing implies that changes in $U$ are slower than changes in $K$, $u = \frac{U}{K} \to 0$ in both the first-best case and the shareholder-side limited commitment case. As $u$ approaches 0, the probability of a binding limited-commitment constraint vanishes and the normalized value function converges to that in the first-best case, $\bar{v} - \frac{1}{\mathbf{r}+\kappa}u$. As the figure illustrates, more contracting frictions reduce the efficiency of risk sharing, lower the value of the firm and shrink the region of feasible continuation utilities, $u$.

In Panel (b) of Figure 2.2, we plot the normalized value function, $v(u, K)$, for Specification 2 for for three levels of firm size: $K_1 < K_2 < K_3$. Note that as $K$ increases, the domain of the normalized utility widens and the level of firm value rises. This is because as firm size increases, the market technology, which is linear in $K$, becomes more efficient compared to concave home production technology. As a result, the limited-commitment constraint on the manager side is less likely to bind, and risk sharing becomes more efficient. In the limit, as $K \to \infty$, the value function converges to that for the case with the shareholder-side limited commitment only.

### 2.6.4 Dynamics of State Variables

In Figure 2.3, we plot the diffusion of the state variables in the $(\ln K, \ln u)$ space for both model specifications. The thick arrows indicate the direction of the state variables upon a positive realization of the Brownian motion shock and the thin arrows represent movements of the state variables upon a negative realization of $dB_t$. The arrows always point toward the southeast or the northwest direction in the interior of the domain, indicating that $u$ and $K$ are negatively correlated. This is the implication of the optimal risk sharing: an increase in $K$ is associated with a decrease in $u = \frac{U}{K}$ because continuation utility ($U$) is less sensitive to shocks than firm size ($K$). As a result, in the stationary distribution, small firms tend to be in the northwest region of the domain, and large firms tend to be in the southeast region of the domain. Also, because $(\ln K, \ln u)$ must stay in its domain with probability one (unless at $K_{MIN}$ where firms exit the economy) for the limited commitment constraint to hold, the direction of the diffusion must be tangent to $u_{MIN}(K)$ and $u_{MAX}(K)$ on the boundaries.

Figure 2.4 plots the expected change (or the drift) of the state variables in the $(\ln K, \ln u)$ space. Note that at the two boundaries, the limited-commitment constraints bind and prevent perfect risk sharing. As a result, the arrows point away from these boundaries because efficiency requires that firms stay away from the binding constraints as much as possible. In addition, arrows point strongly to the right for large values of $u$. Because the length of the arrows is proportional to their magnitude, this pattern indicates rapid growth in size for firms close to $u_{MAX}(K)$. By comparison, the magnitude in growth rates is much smaller for firms close the manager-side limited commitment constraint, $u_{MIN}(K)$. As we show below, these patterns are due to dynamically optimal risk-sharing between shareholders and managers and are consistent with the inverse relationship between size and growth rates in the data.

### 2.6.5 Compensation Policy

Proposition 6 summarizes two important properties of the compensation policy. First, part 2 of the proposition implies that compensation must stay constant whenever the limited-commitment constraints do not bind. The intuition for this result is best illustrated using an expected utility representation of the manager's continuation utility. If we define $\hat{U} = \frac{1}{1-\gamma}U^{1-\gamma}$, it can be shown that $\hat{U}$ is additively separable with respect to consumption across time and states. It is not hard to prove that $\frac{\partial V}{\partial \hat{U}} = \frac{1}{C^{-\gamma}}$ . Intuitively, this condition says that the marginal cost of utility provision for the shareholder must equal the inverse of the marginal utility of the manager: providing one addition unit of utility to the manager requires $1/C^{-\gamma}$ unit of consumption goods.[16] Note that optimality requires the marginal cost of utility provision to be equalized across time and future states if incentive compatibility constraint does not require otherwise. In our model, this means that in the interior of the domain, $\frac{\partial V}{\partial \hat{U}}$ must remain constant. Hence, $C$ must also stay constant.[17]

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

[16] This is commonly known as the "inverse Euler equation" in discrete time dynamic contracting problems (see, for example, Rogerson (1985)).

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

$c(u_t)$, equation (2.24) in part 3 of Proposition 6 implies

$$\ln C_t = \log C_0 + l_t^+ - l_t^-. \tag{2.25}$$

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

In Figure 2.5, we illustrate the sample path of a firm starting from the interior of $[u_{MIN}, u_{MAX}]$. The top panel is the trajectory of the log size of the firm, and the second panel is the path of the normalized utility. The third panel is the corresponding realizations of the value of the firm, and the bottom panel shows the log managerial compensation.[18]

At time 0, the firm starts from the interior of the normalized utility space, $u_0 < u_{MAX}$. A sequence of negative productivity shocks from time 0 to time 1 lowers capital stock of the firm (top panel). For $t < 1$, $u_t < u_{MAX}$ is in the interior (second panel). In this region, firm value is strictly positive (third panel) and managerial compensation is constant (bottom panel). At $t = 1$, $u_t$ hits the boundary $u_{MAX}$ and cannot increase further despite subsequent negative productivity shocks. For $t \in (1, 2)$, the firm continues to receive a sequence of negative productivity shocks and the total capital stock of the firm shrinks further (top panel). During this period, $u_t$ stays at $u_{MAX}$, where the shareholder-side limited-commitment constraint binds, as shown in the second panel of Figure 2.5. The firm value remains at zero and does not cross over to the negative region due to reductions in managerial compensation, which keeps decreasing until the firm starts to experience positive productivity shocks. From time $t = 2$ to $t = 3$, the firm receives a sequence of positive productivity shocks followed by a sequence of negative productivity shocks. As a result, firm value bounces back to the positive region and decreases afterwards. Because the normalized utility $u_t$ stays in the interior before $t = 3$ (second panel), managerial consumption stays constant (bottom panel), although

[18] $\log K_t$ is a Brownian motion with a drift, therefore its sample path has an unbounded variation. To illustrate the basic properties of the optimal contract we plot smooth sample paths.

at a lower level than $C_0$. At time $t = 3$ the size of the firm hits its previous running minimum, and $u_t$ reaches $u_{MAX}$ again. As before, firm value stays at zero, and managerial consumption keeps decreasing, until the firm starts to receive positive productivity shocks for the next time.

In Figure 2.6, we plot a sample path of a firm with $u_0$ close to the left boundary, $u_{MIN}$.

At time 0, the firm starts from the interior of the normalized utility space, $u_{MIN} < u_0 < u_{MAX}$. A sequence of positive productivity shocks from time 0 to 0.5 increases capital stock of the firm (top panel). For $t < 0.5$, $u_t > u_{MIN}$ is in the interior (second panel) and manager's consumption is constant (bottom panel). During this period, both the size of the firm and the normalized firm value, $v(u_t)$, increase. At time 0.5, the normalized continuation utility reaches the left boundary, $u_{MIN}$, and the manager-side limited-commitment constraint binds. Further realizations of positive productivity shocks from $t = 0.5$ to $t = 1$ translate directly into increases in managerial compensation (bottom panel), but the normalized continuation utility (second panel), and the normalized firm value (third panel) remain constant. At time $t = 1$, the firm starts to experience a sequence of negative productivity shocks. As a result, the size of the firm shrinks, and the normalized utility $u_t = \frac{U_t}{K_t}$ increases because risk sharing implies that the continuation utility $U_t$ is less sensitive to shocks than $K_t$ (part 3 of Proposition 6). During $t \in (1, 3)$, $u_t$ stays in the interior of $[u_{MIN}, u_{MAX}]$ and manager's consumption stays constant. At time $t = 2$, the firm starts to receive a sequence of positive productivity shocks. During this period, $u_t$ stays in the interior of its domain until the size of the firm, $K_t$, reaches its previous running maximum at $t = 5$, at which time, the manager-side limited-commitment constraint starts to bind again and, as a result, manager's compensation increases.

### 2.6.6 Investment Policy

The optimal investment-to-capital ratio monotonically increases with the normalized continuation utility under both model specifications. We plot the optimal investment rate as a function of the normalized utility in Figure 2.7. Because the optimal investment rate in Specification 2 depends both on the normalized utility and firm size, in Panel (b) we plot $i(u, K)$ for three levels of firm size: $K_1 < K_2 < K_3$. Consistent with the pattern in value functions, the dashed line, which represents the optimal investment-to-capital ratio for the case with shareholder side limited commitment only, converges to the first-best investment level as $u \to 0$. Also note that in Specification 2, as $K \to \infty$, the optimal investment rate converges to that in the case with the shareholder-side limited commitment only because managers' outside options become negligible compared to the size of the firm's cash flow.

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

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

## 2.7 Quantitative Results

In this section, we present the quantitative implications of our two-sided limited commitment model and discuss its ability to account for key characteristics of the empirical distribution of firm size, CEO compensation, investment and dividend policies. The cross-sectional data that we use consist of US non-financial firms and come from the Center for Research in Securities Prices (CRSP) and Compustat. They are sampled on the annual frequency and cover the period from 1992 till 2011. Our data set is standard and we refer to the online appendix of Ai et al. (2018) for a detailed description of the data.

### 2.7.1 Calibration

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

We follow the standards of the macroeconomics literature to calibrate preference and technology parameters; see, for example, Kydland and Prescott (1982), King and Rebelo (1999) and Rouwenhorst (1995). We choose a risk aversion ($\gamma$) of 2 and set the discount rate ($\beta$) that determines the risk-free interest rate at 4% per year.[19] The capital share parameter $\alpha$ is set at 0.36. We calibrate the exogenous firm death rate, $\kappa$, to be 5% per year and choose $\delta = 7\%$ that together with the exit rate imply a 12% effective annual depreciation rate of capital. The volatility parameter $\sigma$ is set at 35% to match the average volatility of firms' sales growth in the data.

We calibrate the equilibrium marginal product of capital, $\mathbf{A}$, without explicitly specifying the level of aggregate productivity, $\mathbf{z}$. Note that at steady state, aggregate investment must be related to aggregate capital stock by $\mathbf{I} = (\kappa + \delta)\,\mathbf{K}$. Because the aggregate labor supply is normalized to one, the output-to-investment ratio is given by:

$$\frac{\mathbf{Y}}{\mathbf{I}} = \frac{\mathbf{z}\mathbf{K}^{\alpha}}{(\kappa+\delta)\,\mathbf{K}} = \frac{1}{\kappa+\delta}\mathbf{z}\mathbf{K}^{\alpha-\mathbf{1}}. \tag{2.26}$$

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

$$\mathbf{A} = \alpha\mathbf{z}\mathbf{K}^{\alpha-\mathbf{1}} = \alpha\,(\kappa+\delta)\,\frac{\mathbf{Y}}{\mathbf{I}}. \tag{2.27}$$

[19] This allows our model to match the average return of risky and risk-free assets in the data, as in Kydland and Prescott (1982).

We set $\mathbf{A} = 0.231$, which implies an investment-to-output ratio of 18.7%.[20]

We choose capital adjustment cost parameter $\phi = 5$ to account for the average market-to-book ratio (Tobin's Q) in the data and set the initial promised utility, $\bar{U}$, at 0.0879 to match the observed CEO pay to capital ratio of young firms. Two additional moments that we target in our calibration comprise the cross-sectional mean and median of firm growth rates. The calibrated parameter values and the set of moments that they are based on are summarized in Tables 2.1 and 2.2.

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

$$H(K) = \frac{\psi_0}{1+\psi_1} K^{1+\psi_1}, \tag{2.28}$$

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

In Specification 2, we retain the same parameter values as in Specification 1, including the productivity parameter, $\mathbf{z}$.[21] This leaves us with only one additional parameter to calibrate: the managers' endowment of home labor, $\bar{n}$. Because $\bar{n}$ affects managers' outside options relative to their compensation, we choose $\bar{n} = 0.011$ so that the implied average CEO compensation-to-capital ratio is the same under the two model-specifications.[22]

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

### 2.7.2 Size, Age, and Growth Dynamics

In this section, we discuss the cross-sectional distributions of key growth moments in the data and in the model. In what follows, we present average characteristics of quintile (or decile) portfolios sorted by firm size. We follow the standard sorting procedure in the data by assigning firms into portfolios according to their size using breakpoints

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

[21] Note that in Specification 2, the death rate is endogenous, therefore, we can no longer use equation (2.27) to calibrate $\mathbf{A}$.

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

based on the NYSE-traded firms. In the model, firms are sorted using breakpoints that are equally-spaced in log size. Portfolios are re-balanced at the annual frequency. We consider two measures of firm size in the data: the number of employees and gross capital. In the model, we employ a single size measure because, as discussed before, the two measures of size (capital and labor) are equivalent.[23]

We begin with the relationship between age and size. Figure 2.8 plots the median age across decile portfolios sorted on size. Panel (a) presents the data, and Panel (b) shows the implications of the two model specifications. Notice that the data features a nearly monotonic positive relationship between age and size. Smaller firms, on average, are significantly younger than larger firms. While a model with positive average growth rate is likely to generate a positive correlation between size and age, the monotonic pattern observed in the data is more challenging as shown in Panel (b).

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

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

Despite its simplicity, our model captures the distribution of firm age quite well. Luttmer (2010) documents that the median age of firms with more than $10,000$ employees in 2008 was about 75 years. In our model, this number is 93 for Specification 1 and 91 for Specification 2.[24] Luttmer (2010) shows that to jointly account for the

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

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

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

Table 2.3 shows the average investment rate (defined as annual investment divided by beginning-of-year capital stock) as a function of firm size. In the data, small firms invest at a higher rate of about 17% per year relative to large firms, which on average, invest at an about 9% rate. As the table shows, the difference in investment rates of large and small firms is strongly statistically significant. The right two columns show that both versions of our model are able to replicate the observed negative relationship between investment and firm size. The model-implied difference in investment rates of firms in the bottom and top quintiles is about 10%.[25]

Table 2.4 reports the average growth rates of size-sorted portfolios. Consistent with the previous literature (for example, Evans (1987a,b) and Hall (1987)) we find that in the data small firms grow at a significantly higher rate than large firms and this pattern is robust to different measures of firm size. On average, small firms grow by about 10% faster compared with large firms. Our model generates a similar cross-sectional variation in growth rates. The difference in average growth rates between small and large firms implied by the two model specifications is about 11-12%. As explained above, managers in small firms are poorly diversified due to the limited commitment on the shareholder side. As a result, small firms accelerate investment to grow out of the agency conflict. As firm size increases, managers' outside option becomes more attractive. Therefore, large firms have a strong incentive to reduce investment and curtail growth to limit managers' outside options.

Table 2.5 shows the fraction of dividend-paying firms in each size-sorted portfolio. In the data, large firms are much more likely to pay dividends to shareholders than small firms. In the bottom size quintile, on average, only one out of ten firms pays dividends. The fraction of dividend-paying firms increases to about 70-80% in the right tail of the

---

the median age of the largest firms that account for 27% of total employment of the economy is 93 years in Specification 1 and 91 years in Specification 2.

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

size distribution. Our model similarly implies a monotonic increase in the fraction of dividend-paying firms with size. Small firms in the model use most of their resources for investment and, therefore, tend not to pay dividends. These firms start distributing profits to shareholders as they grow and become less constrained.

Empirically, small firms are more likely to fail and exit the market than large firms. As Table 2.6 shows, on average, exit rates of small firms are about twice as high compared with large firms. The model predictions vary depending on the model specification. Under Specification 1, firm death rate does not depend on size because this version of the model has no endogenous exit. Under Specification 2, similar to the data, small firms exit more frequently relative to large firms. However, in the data, exit rates monotonically decrease with size, while in our second specification, death rates of medium and large firms are virtually identical. Note that for simplicity, we assume that, upon entrance, all firms have the same initial condition. Specification 2 of our model can easily generate a smooth variation of exit rates with size if we specify a smooth entrance density of new firms.[26]

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

### 2.7.3 Power Laws in Size, Dividends, and CEO Compensation

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

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

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

for some constant $\xi > 0$. The parameter $\xi$ is called the power-law exponent. The

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

complementary cumulative distribution function of $X$ is given by:

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

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

On average, the estimate of the power-law coefficient of firm size is about 1.26 when size is measured by the number of employees and about 1.09 if size is measured by market capitalization. The latter is very close to the estimates obtained using Census data. For example, Luttmer (2007) reports a power law estimate of 1.07; similar estimates are reported in Gabaix and Landier (2008). Overall, the goodness-of-fit test does not reject the power-law null especially in market capitalization, for which the p-value is above the conventional five-percent level in all but two sample years.[27]

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

Similar to the data, our calibrated model produces a power law in firm size and dividends with a slope close to one. In particular, the tail slope of firm size and dividend payout is 1.09 under Specification 1 and 1.04 under Specification 2.[28] Figure 2.9 provides a visual comparison of the tail behavior of the model-implied distribution and the empirical distribution constructed using one year of the sample data. The horizontal axis in the figure represents market capitalization (and re-scaled dividends) and the vertical axis shows the complementary cumulative distribution function, both are equally spaced on the log scale. In the data, market capitalization is represented by stars and dividends are represented by circles. The solid line is the power law in both size and dividends generated by Specification 1 of our model.[29] Under power law,

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

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

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

the log-log plot is a straight line with a slope equal to the negative exponent. As the figure shows, overall, the model-implied slope matches the slope observed in the data quite well.

As shown in Section 2.5, limited commitment on the manager side implies that CEO compensation in large firms is a linear function of the running maximum of $K^{\nu}$ and, therefore, obeys the same power law as $K^{\nu}$. That is, the optimal contract under limited commitment makes a power law in firm size translate into a power law in CEO compensation. Note that the power law in CEO pay is a limiting result that applies to firms in the right tail of the size distribution. Thus, our power law results for CEO compensation remain valid in the model with two-sided limited commitment.

In Figure 2.10, we show the complementary cumulative distribution function of CEO compensation in the data and in our model. The power law in CEO pay generated by Specification 1 of our model is represented by the solid line with a slope of (negative) 1.09, which is the same as that of firm size. Specification 2 is represented by the dashed line with a power-law coefficient of $\frac{1.04}{0.36} = 2.89$. The observed CEO compensation data (using data of 2000) is represented by circles. As the figure shows, the power-law exponent in the data is in-between the slopes implied by the two model specifications.

In Specification 1, because managers upon default are allowed to participate in the competitive labor market, capital is perfectly scalable and managers' outside options are proportional to firm size. As a result, the power law in CEO pay is the same as the power law in firm size. Under Specification 2, managers are excluded from labor market after default, and capital is no longer scalable. Therefore, managers' outside options are proportional to $K^{\alpha}$ and the power law in CEO pay is $\frac{1}{\alpha}$ times of the power law in firm size. As shown in Figure 2.10, Specification 1 overstates the thickness of the right tail of CEO pay while Specification 2 understates it. Note that we could always choose the parameters of the home production technology to match exactly the power law in CEO compensation data. We do not entertain this approach here as we think that the choice of the home production technology should be ultimately guided by micro-evidence that is not available. Our objective is not to match the data estimates exactly but rather to show that limited commitment on the manager side goes a long way in explaining the right-tail behavior of the CEO pay that we document in the data.

### 2.7.4 CEO Compensation and Firm Size

The cross-sectional distribution of CEO compensation is characterized by several stylized features. While the level of managerial pay increases with firm size, the increase in less than proportional. Therefore, CEO's of small firms have a significantly larger stake in their companies relative to managers of large firms. Table 2.7 shows the variation in the CEO pay-to-capital ratio across size-sorted portfolios. In the data, the median ratio falls from 5.9% for small firms to 0.4% for large firms when size is measured by the number of firm employees, and from 7.3% to 0.1% when size is measured by firm capital. The model features a similar cross-sectional pattern – CEO compensation scaled by firm size declines significantly as size increases under both specifications. In the model, the negative relationship between CEO pay-to-capital ratio and firm size is an implication of (imperfect) risk sharing. Risk sharing reduces sensitivity of CEO compensation to productivity shocks. Therefore, as firms become large, CEO compensation as a fraction of firm size declines.

Similar to the previous literature (see Gabaix (2009)), we find that in the data, the average elasticity of CEO compensation to size is about one-third and is very similar across alternative measures of firm size. Table 2.8 shows the elasticities of managerial compensation to market capitalization, the number of firm employees and capital. The elasticities are estimated in a panel regression of CEO pay on lagged firm size (both measured in logs), controlling for firm and time fixed effects. In the model, the magnitude of size elasticity of CEO compensation is 0.25 and 0.05 under Specifications 1 and 2, respectively. Note that the model-implied elasticities are smaller than in the data because limited commitment is the sole mechanism that generates a positive relationship between CEO pay and firm size. Other agency frictions (for example, moral hazard) are likely to increase the response of managerial compensation to size and help account for the magnitudes observed in the data.

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

Table 2.9 shows variation of size elasticity of CEO pay across quintile portfolios sorted on firm size.[30] As the table shows, the model-implied elasticity is close to zero for medium-size firms, and is quantitatively large for firms in the bottom and the top size portfolios. Under Specification 1, size elasticity of CEO compensation is equal to 0.42, 0.00 and 0.63 for firms in the bottom, middle and top portfolios, respectively. Under Specification 2, only small and large firms feature non-trivial elasticities of 0.08 and 0.17, respectively. In the data, the elasticity is likewise V-shaped. For example, when firm size is measured by market capitalization, the sensitivity of CEO pay to size falls from 0.29 (SE=0.04) for small firms to 0.17 (SE=0.08) for medium-size firms and increases to 0.38 (SE=0.04) for large firms. The cross-sectional pattern in size elasticities is similar if firm size is measured by the number of employees or firm capital.

The V-shaped elasticity of CEO pay with respect to firm size is consistent with our empirical evidence on power law. The power law coefficient for firm size in the data is about 1.1. A uniform elasticity of CEO pay with respect to firm size of $\frac{1}{3}$ would imply a power-law exponent of CEO compensation of $1.1 \times 3 = 3.3$ . However, the distribution of CEO compensation in the data has a fatter tail, with a power law coefficient of about 2.1, which is consistent with a higher than average size elasticity of CEO pay in large firms. Our model has important implications for the time-series dynamics of CEO compensation in response to variation in firm size. Note that while the high sensitivity of CEO pay to size is a feature shared by both small and large firms, its mechanism and, therefore, the dependance of managerial compensation on firm history are quite different. CEO compensation of small firms is sensitive to negative productivity shocks because small firms are more likely to run into a binding limited commitment constraint on the shareholder side. For small firms, a decline in firm size relative to its running minimum is likely to lead to a decline in CEO compensation. In contrast, managerial compensation of large firms is sensitive to positive productivity shocks. For large firms,

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

a sizable increase in firm size (an increase relative to its running maximum) is likely to result in an increase in CEO compensation.[31] We test these model-implied dynamics using panel regression analysis. We elaborate on the model predictions further below after we specify our regression and introduce relevant notation.

We run the following panel regression:

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

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

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

[31] As we show in Section 2.5, the running minimum and the running maximum of firm size are sufficient statistics for the optimal contract in models with one-sided limited commitment. In our full model with two-sided limited commitment, they can no longer summarize firm history completely but they do remain highly informative of binding limited-commitment constraints.

Panel A of Table 2.10 presents the estimates of our panel regression. First, notice that in the model, the coefficients on $\Delta^{Min-}k_{i,t} \cdot I_{i,t}^{Small}$ and $\Delta^{Max+}k_{i,t} \cdot I_{i,t}^{Large}$ are indeed large and positive: 0.53 and 0.76, respectively. For small firms that are declining and large firms that are growing, CEO compensation is highly sensitive to changes in firm size. The coefficients on the other two interaction terms are close to zero – CEO compensation does not change for small firms that experience a positive shock and large firms that realize a negative shock. These firms are moving away from either shareholder or manager constraints and have no need to adjust managerial compensation. Also, the coefficient on the leading term ($\Delta k_{i,t-1}$) is small, of about 0.05, due to the inelastic response of CEO compensation to changes in firm value for medium-size firms. The reported numbers in "Model" column correspond to Specification 1. They are computed using a large panel of simulated data and, as such, represent population values. Specification 2 produces a similar cross-sectional pattern in elasticities and, therefore, is not reported.

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

Our evidence is similar if we estimate elasticities by running a panel regression in levels. In Panel B of Table 2.10, we regress the log-level of CEO compensation ($c_{i,t}$) on its lag ($c_{i,t-1}$), the log of the running maximum of firm size ($k_{i,t-1}^{Max}$) and the four interaction terms. Notice that in the model, CEO compensation is more persistent than in the data because binding limited-commitment constraints are the only channel of variation in managerial pay. Importantly, the estimates in Panel B confirm the V-shaped pattern in size elasticities – in the model and in the data, under-performing small firms and out-performing large firms are characterized by significantly higher sensitivities of CEO compensation to size relative to their counterparts and relative to medium-size firms.

## 2.8 Conclusion

We present a general equilibrium model of firm dynamics. We start with a frictionless model with Arrow-Debreu contracts and illustrate how different forms of limited commitment on compensation contracts help explaining a wide range of empirical regularities in firms' investment, CEO compensation and dividend payout policies. We show that a simple model with two-sided limited commitment is consistent with key cross-sectional characteristics of firms' behavior.

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

Table 2.1: Calibrated Parameters

Table 2.1 presents the calibrated parameter values chosen to target the set of moments listed in Table 2.2.

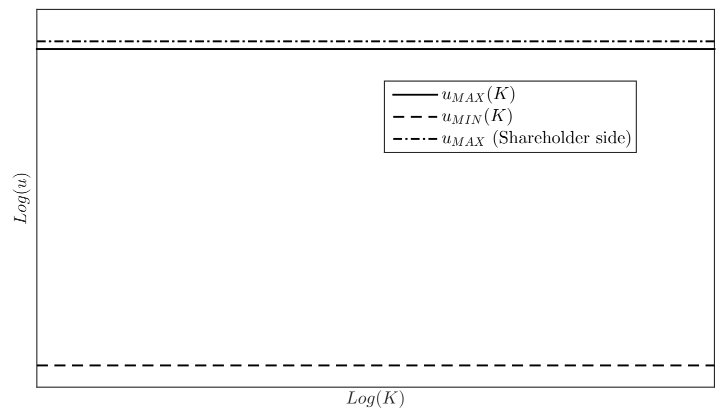
| Description | Notation | Value |
|---|---|---|
| Capital share | $\alpha$ | 0.36 |
| Marginal product of capital | $\mathbf{A}$ | 0.231 |
| Adjustment cost | $\phi$ | 5 |
| Volatility | $\sigma$ | 35% |
| Firm death rate | $\kappa$ | 5% |
| Effective depreciation | $\gamma + \delta$ | 12% |
| Manager initial utility | $\bar{U}$ | 0.0879 |

Table 2.2: Targeted Moments

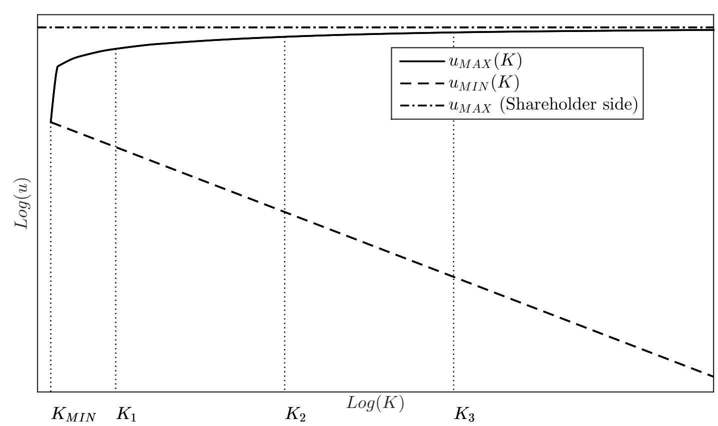
Table 2.2 shows the set of moments targeted in calibrating parameters listed in Table 2.1. We report data statistics and the corresponding moments implied by the two model specifications. The empirical estimates of the capital share and capital depreciation are taken from Kydland and Prescott (1982), and the investment-to-output ratio is computed using the National Income and Product Accounts data available on the Bureau of Economic Analysis website. All other moments are constructed using our sample data. In computing the CEO pay to capital ratio for young firms we consider firms that are less than five years old.

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

Figure 2.1: Domain of the State Variables in the $\ln K$–$\ln u$ Space

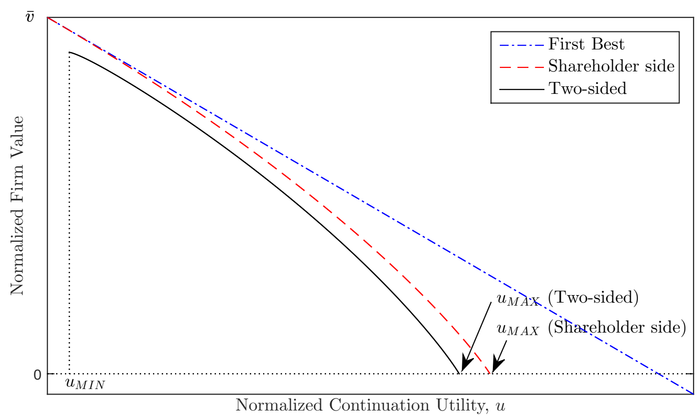



(a) Specification 1

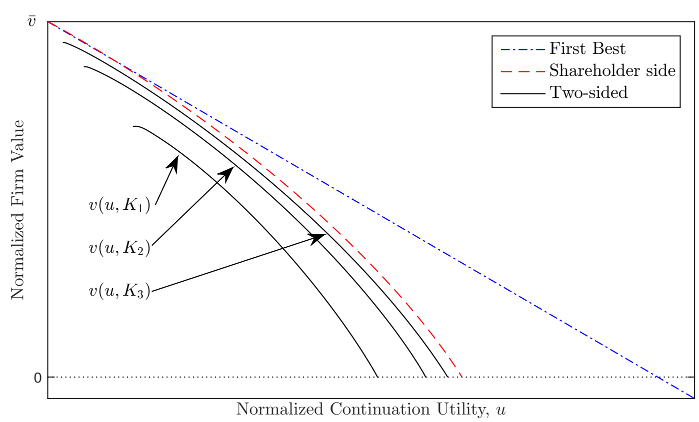



(b) Specification 2

Figure 2.2: Normalized Value Functions

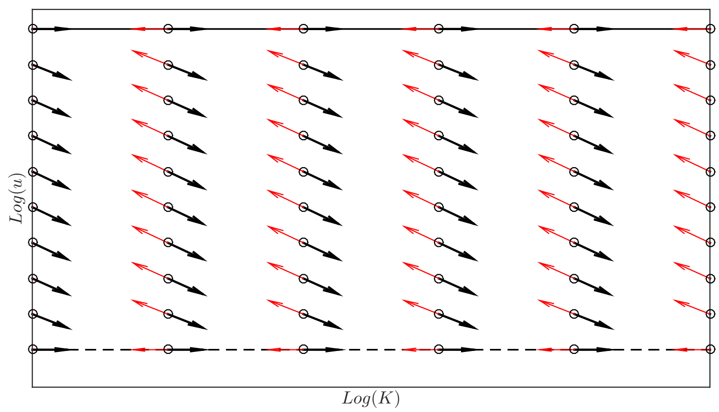



(a) Specification 1

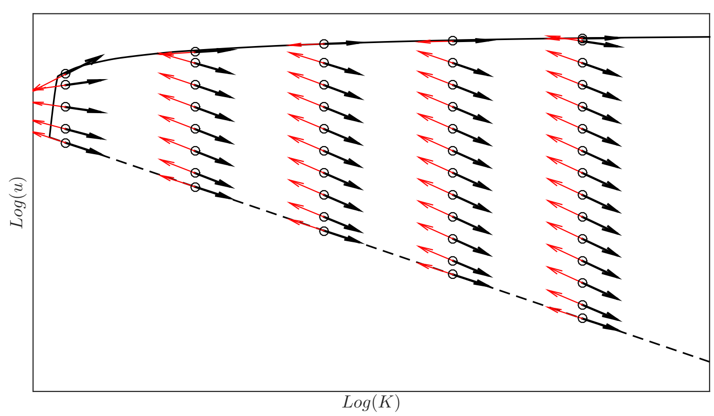



(b) Specification 2

Figure 2.3: Movement of State Variable upon Unexpected Brownian Motion Shocks

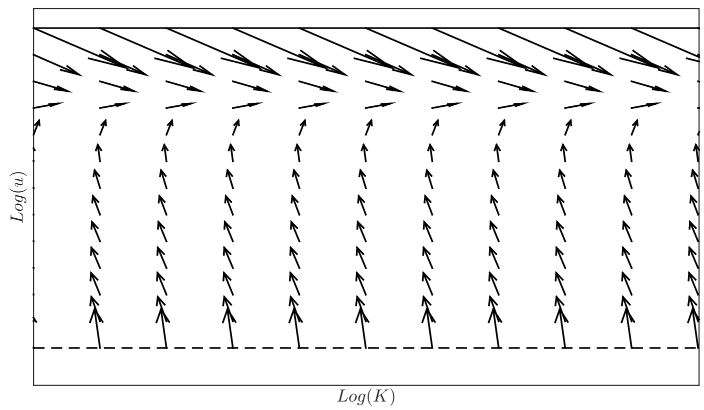



(a) Specification 1

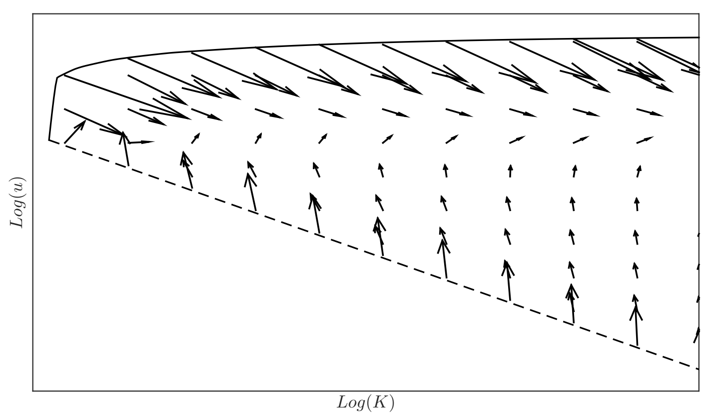



(b) Specification 2

Figure 2.4: The Expected Movement (Drift) of State Variables




(a) Specification 1




(b) Specification 2

Figure 2.5: Sample Path of CEO Compensation

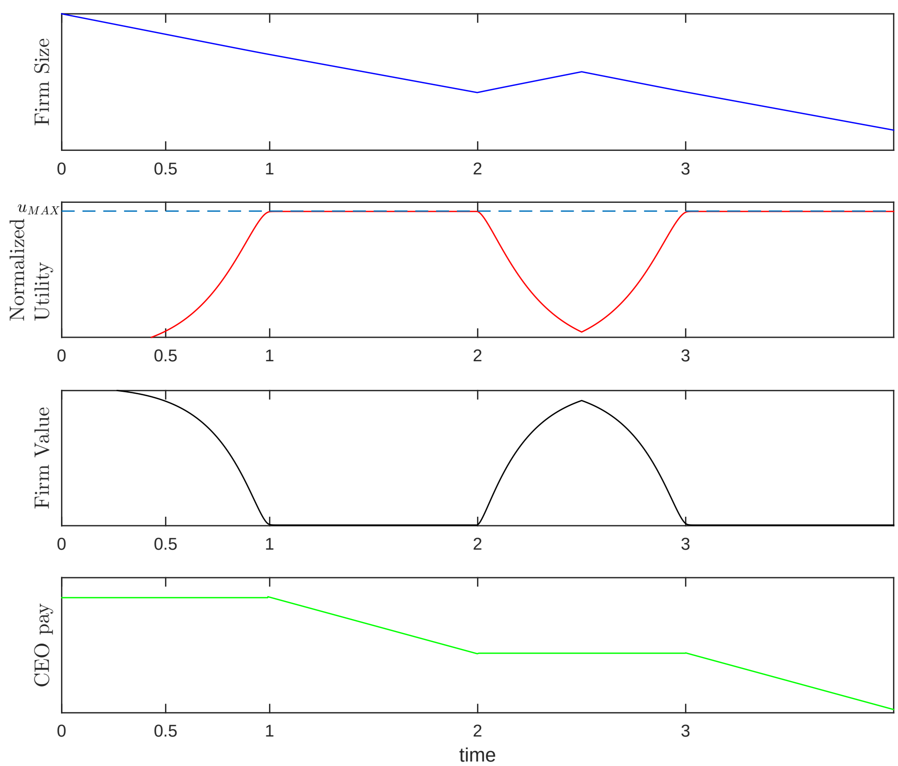


Figure 2.6: Sample Path of CEO Compensation

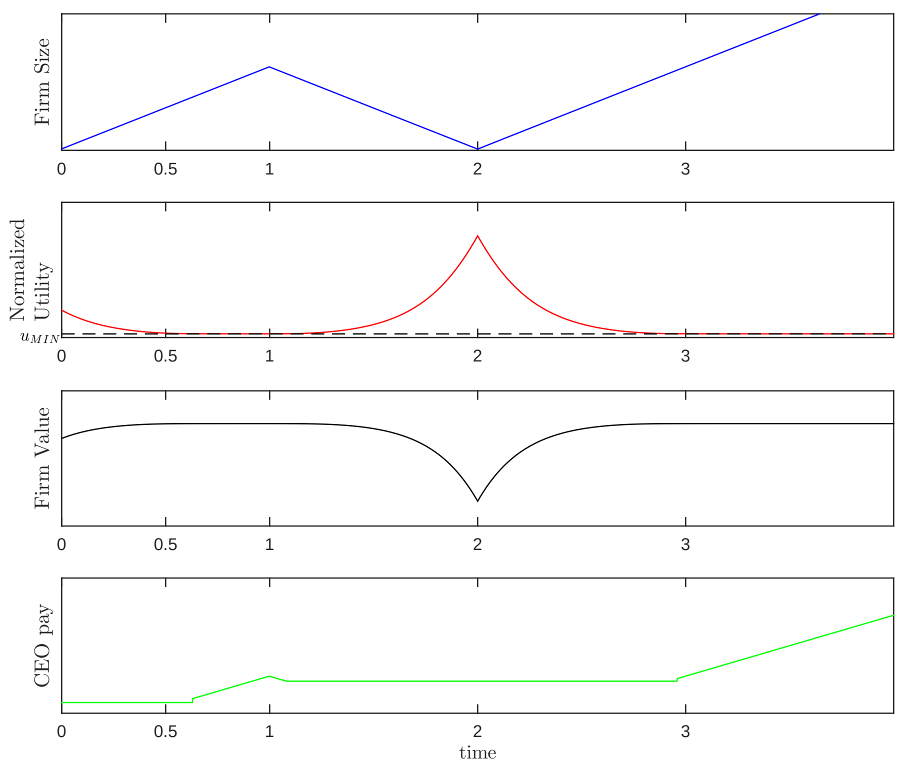
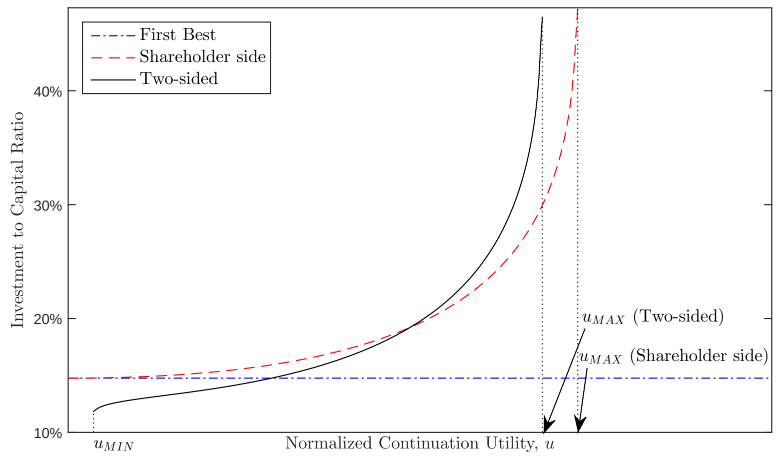
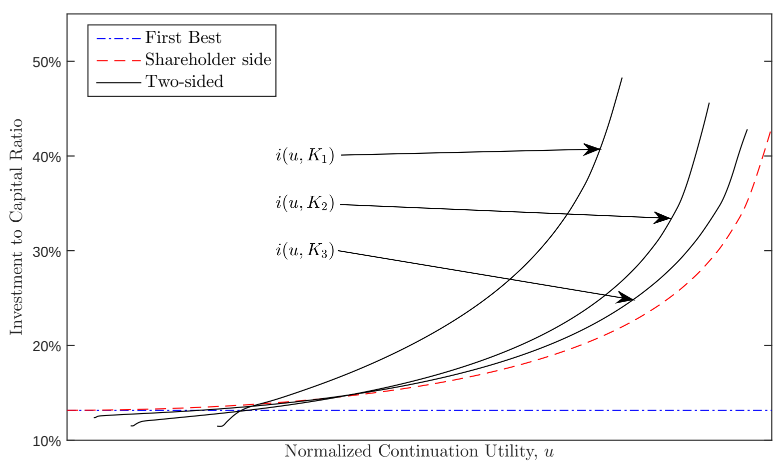


Figure 2.7: Investment Policy: Two-Sided Limited Commitment

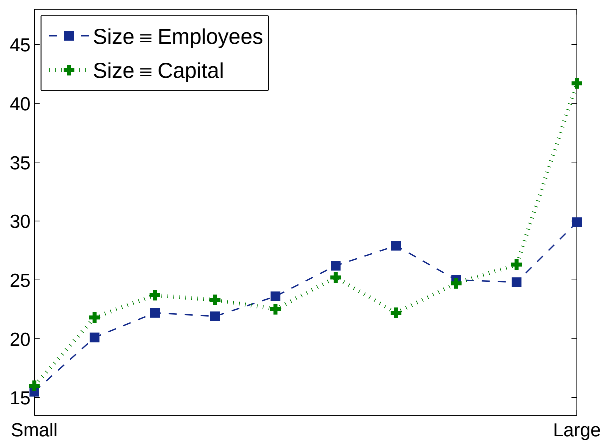



(a) Specification 1

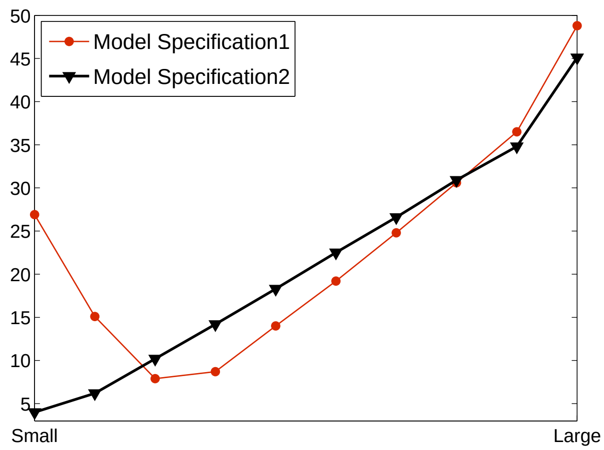



(b) Specification 2

Figure 2.8: Distribution of Age across Size




(a) Data




(b) Model

Figure 2.9: The Right Tail of the Distribution of Market Capitalization and Dividends

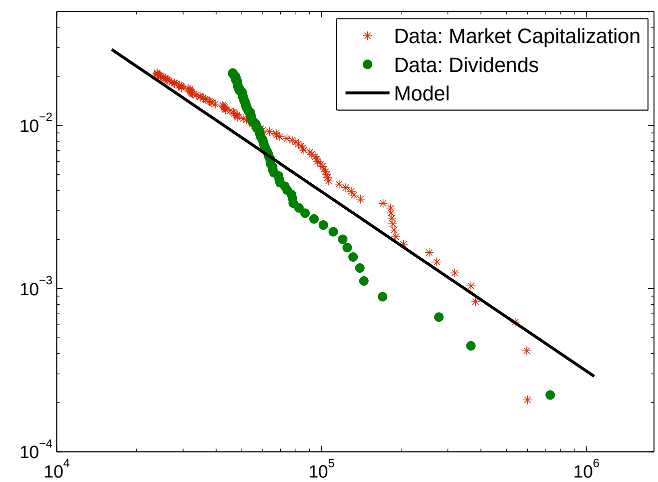



Figure 2.10: The Right Tail of the Distribution of CEO Compensation

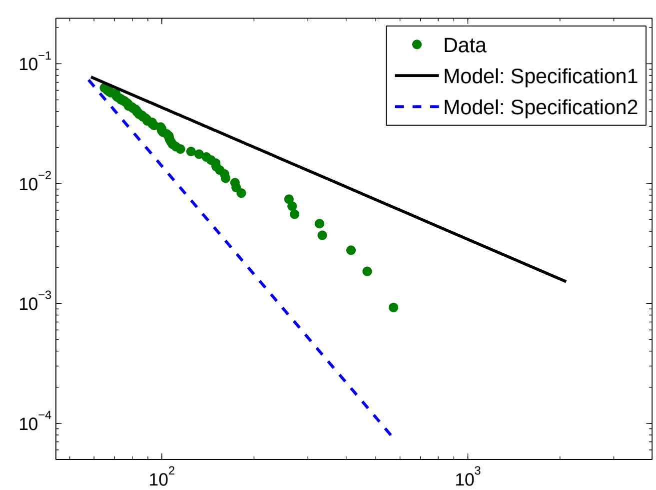


Table 2.3: Investment Rates

Table 2.3 presents the average investment-to-capital ratio of size-sorted portfolios in the data and in the model. Size in the data is measured by either the number of firm employees or gross capital. T-statistics for the difference between large and small firms based on the Newey and West (1987) estimator with four lags are reported in parentheses.

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

Table 2.4: Growth Rates

Table 2.4 presents the average capital growth of size-sorted portfolios in the data and in the model. Size in the data is measured by either the number of firm employees or gross capital. T-statistics for the difference between large and small firms based on the Newey and West (1987) estimator with four lags are reported in parentheses.

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

Table 2.5: Fraction of Dividend-Paying Firms

Table 2.5 presents the average fraction of dividend-paying firms for size-sorted portfolios in the data and in the model. Size in the data is measured by either the number of firm employees or gross capital. T-statistics for the difference between large and small firms based on the Newey and West (1987) estimator with four lags are reported in parentheses.

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

Table 2.6: Firm Exit Rates

Table 2.6 presents the average firm exit rate for size-sorted portfolios in the data and in the model. Size in the data is measured by either the number of firm employees or gross capital. T-statistics for the difference between large and small firms based on the Newey and West (1987) estimator with four lags are reported in parentheses.

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

Table 2.7: CEO Pay-to-Capital Ratio

Table 2.7 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 | | Model | Model |
|---|---|---|---|---|
| | Size≡Employees | Size≡Capital | Specification1 | Specification2 |
| Small | 0.059 | 0.073 | 0.138 | 0.088 |
| 2 | 0.021 | 0.019 | 0.087 | 0.042 |
| 3 | 0.013 | 0.009 | 0.036 | 0.017 |
| 4 | 0.008 | 0.005 | 0.017 | 0.007 |
| Large | 0.004 | 0.001 | 0.016 | 0.002 |
| Large−Small | −0.055 | −0.072 | −0.122 | −0.086 |
| | (−14.27) | (−14.78) | | |

Table 2.8: Elasticity of CEO Compensation to Firm Size

Table 2.8 shows the elasticity of CEO compensation to firm size. Elasticities are estimated in a panel regression of CEO pay on lagged firm size (both measured in logs), controlling for firm and time fixed effects. Size in the data is measured by either market capitalization, the number of firm employees or gross capital. In "Data" panel, we report the estimated elasticities, standard errors clustered by firm and time (in brackets), and regression $R^2$'s. The model statistics represent population numbers that are computed using a large panel of simulated data.

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

Table 2.9: Elasticity of CEO Compensation to Size Conditional on Firm Size

Table 2.9 shows variation in size elasticities of CEO compensation across size-sorted portfolios. Elasticities are estimated in a panel regression of CEO pay on lagged firm size (both measured in logs), controlling for firm and time fixed effects. Size in the data is measured by either market capitalization, the number of firm employees or gross capital. In "Data" panel, we report the estimated elasticities and clustered standard errors (in brackets). The model statistics represent population numbers that are computed using a large panel of simulated data.

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

Table 2.10: Dynamics of CEO Compensation

Table 2.10 presents the dynamics of CEO compensation in the data and in the model. Panel A shows the estimates from a panel regression of the log-growth of CEO compensation ($\Delta c_t \equiv \log \frac{C_t}{C_{t-1}}$) on the log-growth of firm size ($\Delta k_{t-1} \equiv \log \frac{K_{t-1}}{K_{t-2}}$), and four interaction terms that correspond to opposite changes in size realized by small and large firms.† For each firm, we compute the change in size at time $t$ relative to its running minimum and maximum: $\Delta^{Min} k_t \equiv \log \frac{K_t}{K_{t-1}^{Min}}$ and $\Delta^{Max} k_t \equiv \log \frac{K_t}{K_{t-1}^{Max}}$, where $K_{t-1}^{Min}$ and $K_{t-1}^{Max}$ are respectively minimum and maximum of firm size observed in the five years prior to year $t$. We focus separately on negative and positive changes in firm size: $\Delta^{Min-} k_t$ and $\Delta^{Max-} k_t$ represent declines, and $\Delta^{Min+} k_t$ and $\Delta^{Max+} k_t$ represent increases in firms size. $I_t^{Small}$ and $I_t^{Large}$ are size dummies that select firms in the bottom and the top decile of size distribution, respectively. In Panel B, we consider a specification in levels, where we regress the log-level of CEO compensation ($c_t$) on its lag ($c_{t-1}$), the log of the running maximum of firm size ($k_{t-1}^{Max}$) and the four interaction terms. In the data, we run a panel regression with firm and time fixed effects and report the estimates and the corresponding t-statistics (in parentheses). Size in the data is measured by market capitalization. The model statistics represent population numbers that are computed using a large panel of simulated data under Model Specification 1.

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

# References

Hengjie Ai and Anmol Bhandari. Asset pricing with endogenously uninsurable tail risks. *Working Paper*, 86(4):1383–1430, 2018.

Hengjie Ai and Rui Li. Investment and ceo compensation under limited commitment. *Journal of Financial Economics.*, 116:452–472, 2015.

Hengjie Ai, Dana Kiku, Rui Li, and Jincheng Tong. A unified model of firm dynamics with limitedcommitment and assortative matching. *Working Paper*, 2018.

Rui Albuquerque and Hugo A. Hopenhayn. Optimal dynamic lending contracts with imperfect enforceability. *Review of Economic Studies*, 71:285–315, 2004.

Fernando Alvarez and Urban J. Jermann. Efficiency, equilibrium, and asset pricing with risk of default. *Econometrica*, 68(4):775–797, 2000.

Fernando Alvarez and Urban J. Jermann. Quantitative Asset Pricing Implications of Endogenous Solvency Constraints. *Review of Financial Studies*, 14(4):1117–1151, 2001.

Daniel Andrei, William Mann, and Nathalie Moyen. Why did the q theory of investment start working? *forthcoming, Journal of Financial Economics*, 2018.

Manuel Arellano and Olympia Bover. Another look at the instrumental variable estimation of error-components models. *Journal of Econometrics*, 68(1):29–51, 1995.

Andrew Atkeson and PatrickJ. Kehoe. Modeling and measuring organization capital. *Journal of Political Economy*, 113(5):1026–1053, 2005. doi: 10.1086/431289. URL `https://doi.org/10.1086/431289`.

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

Andrew Atkeson and Robert Lucas. Efficiency and equality in a simple model of efficient unemployment insurance. *Journal of Economic Theory*, 66(1):64 – 88, 1995. ISSN 0022-0531.

Ravi Bansal and Amir Yaron. Risks for the Long Run: A Potential Resolution of Asset Pricing Puzzles. *Journal of Finance*, 59(4):1481–1509, 2004.

Suleyman Basak and Domenico Cuoco. An equilibrium model with restricted stock market participation. *The Review of Financial Studies*, 11(2):309–341, 1998.

Frederico Belo and Jianfeng Yu. Government investment and the stock market. *Journal of Monetary Economics*, 60(3):325–339, 2013.

Frederico Belo, Xiaoji Lin, and Santiago Bazdresch. Labor hiring, investment, and stock return predictability in the cross section. *Journal of Political Economy*, 122(1): 129–177, 2014.

Jonathan B. Berk, Richard C. Green, and Vasant Naik. Optimal investment, growth options, and security returns. *The Journal of Finance*, 54(5):1553–1607, 1999.

Bruno Biais, Thomas Mariotti, Guillaume Plantin, and Jean-Charles Rochet. Dynamic security design. *Working Paper, IDEI*, 2004.

Bruno Biais, Thomas Mariotti, Jean-Charles Rochet, and Stéphane Villeneuve. Large risks, limited liability, and dynamic moral hazard. *Econometrica*, 78:73–118, 2010a.

Bruno Biais, Thomas Mariotti, and Jean-Charles Rochet Stéphane Villeneuve. Large risks, limited liability, and dynamic moral hazard. *Econometrica*, 78:73–118, 2010b.

Nicholas Bloom. The impact of uncertainty shocks. *Econometrica*, 77(3):623–685, 2009.

Michele Boldrin, Lawrence J. Christiano, and Jonas D. M. Fisher. Habit persistence, asset returns, and the business cycle. *The American Economic Review*, 91(1):149–166, 2001.

Patrick Bolton, Neng Wang, and Jinqiang Yang. Liquidity and risk management: Coordinating investment and compensation policies. *Working Paper, Columbia University*, 2016.

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

John Y. Campbell. A variance decomposition for stock returns. *The Economic Journal*, 101(405):157–179, 1991.

John Y Campbell and Robert J Shiller. The dividend-price ratio and expectations of future dividends and discount factors. *Review of financial studies*, 1(3):195–228, 1988.

John Y. Campbell and Tuomo Vuolteenaho. Bad beta, good beta. *The American Economic Review*, 94(5):1249–1275, 2004.

John Y. Campbell, Stefano Giglio, and Christopher Polk. Hard times. *The Review of Asset Pricing Studies*, 3(1):95–132, 2013.

Murray Carlson, Adlai Fisher, and Ron Giammarino. Corporate investment and asset price dynamics: Implications for the cross-section of returns. *The Journal of Finance*, 59(6):2577–2603, 2005.

Christopher D. Carroll. The method of endogenous gridpoints for solving dynamic stochastic optimization problems. *Economics Letters*, 91(3):312–320, 2006.

Andrew Y. Chen. External habit in a production economy: A model of asset prices and consumption volatility risk. *The Review of Financial Studies*, 30(8):2890–2932, 2017.

Long Chen and Xinlei Zhao. Return decomposition. *The Review of Financial Studies*, 22(12):5213–5249, 2009.

YiLi Chien and Hanno Lustig. The market price of aggregate risk and the wealth distribution. *Review of Financial Studies*, 23(4):1596–1650, April 2010.

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

Gian Clementi and Berardino Palazzo. Investment and the cross-section of equity returns. *forthcoming, Journal of Finance*, 2018.

John H. Cochrane. Production-based asset pricing and the link between stock returns and economic fluctuations. *The Journal of Finance*, 46(1):209–237, 1991.

John H. Cochrane. Presidential address: Discount rates. *The Journal of Finance*, 66 (4):1047–1108, 2011.

George Constantinides and Darrell Duffie. Asset pricing with heterogenous consumers. *Journal of Political Economy*, 104(2):219–240, 1996.

George Constantinides and Anisha Ghosh. Asset pricing with countercyclical household consumption risk. *The Journal of Finance*, 72(1):415–460, 2015.

Thomas Cooley, Ramon Marimon, and Vincenzo Quadrini. Aggregate consequences of limited contract enforceability. *Journal of Political Economy*, 112:817–847, 2004.

Thomas Cooley, Ramon Marimon, and Vincenzo Quadrini. Risky investments with limited commitment. *NBER Working Paper No. 19594*, 2013.

Alexandre Corhay, Howard Kung, and Lukas Schmid. Competition, markups and predictable returns. 2017.

Mariano Massimiliano Croce. Long-run productivity risk: A new hope for production-based asset pricing? *Journal of Monetary Economics*, 66:13–31, 2014.

Steven J. Davis and Till von Wachter. Recessions and the Costs of Job Loss. *Brookings Papers on Economic Activity*, 2011(2):1–72, 2011.

Peter DeMarzo and Michael Fishman. Agency and optimal investment dynamic. *Review of Financial Studies*, 20:151–188, 2007.

Peter DeMarzo and Yuliy Sannikov. Optimal security design and dynamic capital structure in a continuous-time agency model. *Journal of Finance*, 61:2681–2724, 2006.

Peter DeMarzo and Yuliy Sannikov. Learning, termination and payout policy in dynamic incentive contracts. *Working Paper, Princeton University*, 2012.

Andres Donangelo. Untangling the value premium with labor shares. *Working Paper*, 2018.

Andres Donangelo, François Gourio, Matthias Kehrig, and Miguel Palacios. The cross-section of labor leverage and equity returns. *Journal of Financial Economics*, (2017-22), 2018.

Alex Edmans and Xavier Gabaix. Executive compensation: A modern primer. *Journal of Economic Literature*, forthcoming, 2016.

Redouane Elkamhi and Chanik Jo. Time-varying market participation: A potential resolution of asset pricing puzzles. 2018.

David Evans. The relationship between firm growth, size, and age: Estimates for 100 manufacturing industries. *Journal of Industrial Economics*, 35:567–581, 1987a.

David Evans. Tests of alternative theories of firm growth. *Journal of Political Economy*, 95:657–674, 1987b.

Jack Favilukis and Xiaoji Lin. Wage rigidity: A quantitative solution to several asset pricing puzzles. *Review of Financial Studies*, 29:148–192, 2016.

Xavier Gabaix. Power laws in economics and finance. *Annual Review of Economics*, 1: 255–293, 2009.

Xavier Gabaix and Augustin Landier. Why has ceo pay increased so much? *Quarterly Journal of Economics*, 123:49–100, 2008.

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

Robert Goldstein and Fan Yang. Asset pricing implications from wealthy shareholder consumption and net payout. *Working Paper*, 2015.

Francisco Gomes, Alexander Michaelides, and Valery Polkovnichenko. Fiscal policy and asset prices with incomplete markets. *The Review of Financial Studies*, 26(2): 531–566, 2013.

Joao Gomes, Leonid Kogan, and Lu Zhang. Equilibrium cross section of returns. *Journal of Political Economy*, 111(4):693–732, 2003.

Myron Gordon. Dividends, earnings and stock prices. *Review of Economics and Statistics*, 41:99–105, 1959.

Francois Gourio. Disaster risk and business cycles. *American Economic Review*, 102 (6):2734–66, May 2012.

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

Fatih Guvenen. A parsimonious macroeconomic model for asset pricing. *Econometrica*, 77(6):1711–1750, 2009.

Bronwyn Hall. The reslationship between firm size and firm growth in the US manufacturing sector. *The journal of industrial economics*, 35:583–606, 1987.

Robert E. Hall. The stock market and capital accumulation. *American Economic Review*, 91(5):1185–1202, December 2001.

Milton Harris and Bengt Holmstrom. A theory of wage dynamics. *Review of Economic Studies*, XLIX:315–333, 1982.

Michael Harrison. *Brownian Motion and Stochastic Flow Systems.* John Wiley and Sons, 1985.

Fumio Hayashi. Tobin's marginal q and average q: A neoclassical interpretation. *Econometrica*, 1982.

Zhiguo He. Optimal executive compensation when firm size follows geometric brownian motion. *Review of Financial Studies*, 22:859–892, 2009.

Ayse Imrohoroglu and Selale Tuzel. Firm-level productivity, risk, and return. *Management Science*, 60(8):2073–2090, 2014.

Ravi Jagannathan and Zhenyu Wang. The conditional capm and the cross-section of expected returns. *The Journal of Finance*, 51(1):3–53, 1996.

Urban J. Jermann. Asset pricing in production economies. *Journal of Monetary Economics*, 41(2):257–275, April 1998.

Georg Kaltenbrunner and Lars A. Lochstoer. Long-run risk through consumption smoothing. *The Review of Financial Studies*, 23(8):3190–3224, 2010.

Timothy J. Kehoe and David K. Levine. Debt-constrained asset markets. *Review of Economic Studies*, 60:865–888, 1993.

Robert G. King and Sergio T. Rebelo. Chapter 14 resuscitating real business cycles. In *Handbook of Macroeconomics*, volume 1 of *Handbook of Macroeconomics*, pages 927 – 1007. Elsevier, 1999.

Nobuhiro Kiyotaki and John Moore. Credit cycles. *Journal of Political Economy*, 105: 211–284, 1997.

Narayana Kocherlakota. Implications of efficient risk sharing without commitment. *Review of Economic Studies*, 63:595–609, 1996.

Leonid Kogan and Dimitris Papanikolaou. Economic activity of firms and asset prices. *Annual Review of Financial Economics*, 4(1):361–384, 2012.

Leonid Kogan and Dimitris Papanikolaou. Firm characteristics and stock returns: The role of investment-specific shocks. *The Review of Financial Studies*, 26(11):2718–2759, 2013.

Leonid Kogan, Dimitris Papanikolaou, and Noan Stoffman. Winners and losers: Creative destruction and the stock market. *Journal of Political Economy*, 2018.

Dirk Krueger and Hanno Lustig. When is market incompleteness irrelevant for the price of aggregate risk (and when is it not)? *Journal of Economic Theory*, 145(1):1–41, 2010.

Per Krusell and Anthony A. Smith. Income And Wealth Heterogeneity, Portfolio Choice, And Equilibrium Asset Returns. *Macroeconomic Dynamics*, 1(02):387–422, June 1997.

Per Krusell and Anthony A Smith. Income and wealth heterogenity in the macroeconomy. *Journal of Political Economy*, 106(5):867–896, 1998.

Howard Kung and Lukas Schmid. Innovation, growth, and asset prices. *The Journal of Finance*, 70(3):1001–1037, 2015.

Finn E. Kydland and Edward C. Prescott. Time to build and aggregate fluctuations. *Econometrica*, 50(6):1345–1370, 1982.

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

Xiaoji Lin and Lu Zhang. The investment manifesto. *Journal of Monetary Economics*, 60(3):351–366, 2013.

Guido Lorenzoni and Karl Walentin. Financial frictions, investment and tobin's q. *NBER Working Paper No. 13092*, 2007.

Hanno Lustig, Chad Syverson, and Stijn Van Nieuwerburgh. Technological change and the growing inequality in managerial compensation. *Journal of Financial Economics*, 99:601–627, 2011.

Erzo Luttmer. Selection, growth, and the size distribution of firms. *Quarterly Journal of Economics*, 122:1103–1144, 2007.

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

Erzo Luttmer. On the mechanics of firm growth. *The Review of Economic Studies*, 78: 1042–1068, 2011.

Christopher J. Malloy, Tobais J. Moskowitz, and Annete Vissing-Jorgensen. Long-run stockholder consumption risk and asset returns. *The Journal of Finance*, 64(6):2427–2479, 2009.

Gregory Mankiw. The Equity Premium and The Concentration of Aggregate Shocks. *Journal of Financial Economics*, 17:211–219, 1986.

N Gregory Mankiw and Stephen P Zeldes. The consumption of stockholders and non-stockholders. *Journal of Financial Economics*, 29:97–112, 1991.

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

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

Stephen Nickell. Biases in dynamic models with fixed effects. *Econometrica*, 49(6): 1417–1426, 1981.

Dimitris Papanikolaou. Investment shocks and asset prices. *Journal of Political Economy*, 119(4):639–685, 2011.

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

Adriano Rampini and S Viswanathan. Collateral and capital structure. *Working paper, Duke University*, 2012.

William Rogerson. Repeated moral hazard. *Econometrica*, 53:69–76, 1985.

Paul M. Romer. Increasing returns and long-run growth. *Journal of Political Economy*, 94(5):1002–1037, 1986.

K. Geert Rouwenhorst. Asset pricing implications of equilibrium business cycle models. In *Frontiers of Business Cycle Research*. Princeton University Press, 1995.

Sergio Salgado, Fatih Guvenen, and Nicholas Bloom. Skewed business cycles. *Working Paper*, 2017.

Yuliy Sannikov. A continuous-time version of the principal-agent problem. *Review of Economic Studies*, 75:957–984, 2008.

Lukas Schmid. A quantitative dynamic agency model of financing constraints. *Working paper, Duke University*, 2008.

Lawrence Schmidt. Climbing and falling off the ladder : Asset pricing implications of labor market event risk. *Working Paper*, 2015.

Marko Terviö. The difference that ceos make: An assignment model approach. *American Economic Review*, 98:642–668, 2008.

Sheridan Titman, K. C. John Wei, and Feixue Xie. Capital investments and stock returns. *The Journal of Financial and Quantitative Analysis*, 39(4):677–700, 2004.

Tuomo Vuolteenaho. What drives firm-level stock returns? *The Journal of Finance*, 57 (1):233–264, 2002.

Jessica A. Wachter and Motohiro Yogo. Why do household portfolio shares rise in wealth? *The Review of Financial Studies*, 23(11):3929–3965, 2010.

Thomas Winberry. Lumpy investment, business cycles, and stimulus policy. *Working Paper*, 2018.

Lu Zhang. The value premium. *The Journal of Finance*, 60(1):67–103, 2005.

# Appendix

## A. Data

In this section, we discuss aggregate variables we use for the estimation of aggregate VAR model.

### Data used for estimating aggregate VAR

The first variable variable is investment to capital ratio. Corporate investment is the seasonally-adjusted private nonresidential fixed investment provided by the FRED website of the St. Louis Fed. Since the stock of private capital time series is not available at quarterly frequency. We follow Cochrane (1991) and construct the time series of the investment to capital ratio using the following formula derived from the perpetual inventory method:

$$Ik_t = \frac{I_t}{I_{t-1}} \frac{IK_{t-1}}{1 - \delta + IK_{t-1}} \tag{2.29}$$

The initial value of the investment rate is set to be the steady-state level. Given its initial value, the entire time series of investment rate can be computed using equation 2.29. See Cochrane (1991) and Belo and Yu (2013) for more details on this method.

The second variable is the quarterly growth rate of net corporate dividend which is available on the FRED website of the St. Louis Fed. The series we include in our VAR is the corporate profits after tax with inventory valuation and capital consumption adjustment. We deflate this series by its the price index for gross value added by the U.S. business sector.

The third variable in the aggregate VAR is the market excess return which is the difference between log return on the CRSP value-weighted stock market index and the log risk-free rate. For risk-free rate, we use Treasury bills with three month maturity provided by CRSP.

We add other variables that have been demonstrated to predict our key variables of interests to our VAR systems. We further include 1) log price-earnings ratio (PE). 2) the term yield computed as the difference between the log yield on the 10-year U.S. constant maturity bond and the log yield on the 3-month U.S. Treasury bill. 3) the default spread that is the difference between the log yield on Moody's BAA and AAA bonds. 4) the small-stock value spread that is available on Kenneth French's website. Since the construction of these forecasting variables are standard, we refer the interested readers to Vuolteenaho (2002), Campbell and Vuolteenaho (2004) and Campbell et al. (2013).

**Data used for estimating firm-level panel VAR**

The sample of data we use to estimate panel VAR model is from the Compustat database for the period from 1978 quarter 1 to 2017 quarter 4. We exclude firms not incorporated in the US or not traded on either. To guarantee that our results are not driven by firms entry and exit, we exclude firms who exist in our sample period for less than 10 years. We further exclude firms with SIC code from 4900 to 4999, 6000 to 6999 or greater equal to 9000. We impose further exclusion criterion: 1) we winsorize the investment rate distribution every quarter. We winsorize investment rate observations that are above 99.5% percentiles or below 0.5%. 2) We discard firm observations that missing investment rate. Even if one firm reports PPENTQ for two adjacent time periods, we do not extrapolate the PPENTQ that is missing in between two observations. To assign industry-specific depreciation rate to each firm in our sample, we match 3 digits NAICS code to the BEA implied depreciation estimates. However, BEA reports depreciation estimates on different capital class. Since in Compustat data, there is no distinction between structure vs equipment, we take the average estimates of depreciation across different types of assets, for each industry.

We use two measures of firm-level cash flow. The first measure is ROE, calculated as net income dividend by the last period of book equity. I drop firm year observations if a firm's ROE is less than $-100\%$ because log transformations are not possible for variables less than 0. Vuolteenaho (2002) justifies the use of ROE as a measure of firm-level cash flow by loglinearizing the accounting clean-surplus identify.

Our second measure of cash flow is a firm's marginal product of capital. We measure the firm's capital stock, as the (net of depreciation) value of property, plant and equipment and firm revenue, as reported sales. We measure the marginal product of capital as the difference between log sales and log capital stock. In theory, a firm's marginal product of capital should also depend on a constant capital elasticity parameter. We ignore this constant term here since it does not play any role in our analysis.

### Firm-level investment and labor share

The sample period to compute firm-level quantities covers the same range as the data used for estimating VAR models. The construction of firm-level investment rate data used to calibrate the model is identical to that for estimating firm-level panel VAR model. We define firm-level labor share as labor compensation to firm value added ratio which can be directly mapped into compensation to output ratio in our model. For the labor compensation we use the Compustat item Total Staff Expense (XLR). We focus on value added that contains the contribution of labor and owned physical capital of the firm only because in our model firm's production requires labor and physical capital as the only inputs. Value added is computed as the Operating Income Before Depreciation and Amortization (Compustat item OIBDP) plus labor expenses. Donangelo (2018) also includes changes in inventories as part of firm value added. We do not consider inventories in our empirical measure of value added because in our model all goods produced are sold contemporaneously.

### Numerical solution

We now discuss how we solve the model numerically. Note that the optimal contracting problem 1.15 is not stationary and we must first detrend it and rewrite it in terms of stationary quantities. We rewrite the normalized problem 1.19 here for explaining our numerical solution more clearly.

$$\tilde{v}(u|\mathbf{c},\theta) = \max_{u',k,w} \Bigg\{ \theta k^{\alpha} - w - \mathbf{p}(\mathbf{c},\theta)k + \mathbf{E}\Big[\Lambda'(\mathbf{c}'|\mathbf{c},\theta)(1-\delta+\mathbf{i}(\mathbf{c},\theta))e^{\varepsilon'}\tilde{v}(u'|\mathbf{c}',\theta')\Big] + k(1-\delta)\mathbf{E}\Big[\Lambda'(\mathbf{c}'|\mathbf{c},\theta)\mathbf{p}(\mathbf{c}',\theta')\Big] \Bigg\} \tag{2.30}$$

$$s.t. \quad u = \Big\{(1-\beta)w^{1-\frac{1}{\psi}} + \beta\mathcal{R}(u')^{1-\frac{1}{\psi}}\Big\}^{\frac{1}{1-\frac{1}{\psi}}} \tag{2.31}$$

$$e^{\varepsilon'}\tilde{v}(u'|\mathbf{c}',\theta') + \frac{\mathbf{p}(\mathbf{c}',\theta')(1-\delta)}{1-\delta+\mathbf{i}(\mathbf{c},\theta)}k \geq 0 \quad \forall(\varepsilon',\theta'), \quad \text{if } \eta = 0 \tag{2.32}$$

$$e^{\varepsilon'}\tilde{v}(u'|\mathbf{c}',\theta') \geq 0 \quad \forall(\varepsilon',\theta'), \quad \text{if } \eta = 1 \tag{2.33}$$

$$u'(u,\varepsilon'|\mathbf{c}',\theta') \geq \underline{u}(\mathbf{c}',\theta')\forall(\varepsilon',\theta') \tag{2.34}$$

Note that we separate the two cases on whether owner can sell capital into two conditions 2.33 and 2.32.

The numerical solution then is obtained in two steps: given a forecasting rule for the law of motion of owner's consumption share and capital price as a function of aggregate state $\theta$ and owner's consumption share, we solve for the normalized 2.30. Secondly, we introduce an algorithm based on Krusell and Smith (1997) to simultaneously look for equilibrium prices that satisfies market clearing conditions 1.23 and 1.24.

**Solving the optimal contracting problem 2.30**

1. Start with an initial guess of the law of motion of $\mathbf{c}$, $\Gamma_{\mathbf{c}}(\theta'|\theta,\mathbf{c})$

$$\log \mathbf{c}' = \alpha(\theta,\theta') + \beta(\theta,\theta')\log \mathbf{c} \tag{2.35}$$

and a capital pricing function $\Gamma_{\mathbf{p}}(\theta,\mathbf{c})$

$$\log \mathbf{p} = a(\theta) + b(\theta)\log \mathbf{c} \tag{2.36}$$

2. Given $\Gamma_{\mathbf{c}}(\theta'|\theta,\mathbf{c})$ and $\Gamma_{\mathbf{p}}(\theta,\mathbf{c})$, we jointly solve for SDF denoted as $\Lambda(\theta'|\mathbf{c},\theta)$ and aggregate investment denoted as $\mathbf{i}(\mathbf{c},\theta)$ on a set of grid point for $\mathbf{c}$, using the expression of SDF and the optimality condition of the investment good sector:

$$\begin{aligned} \Lambda(\theta'|\mathbf{c},\theta) &= \beta(1-\delta+\mathbf{i}(\mathbf{c},\theta))^{-\gamma}\left\{\frac{\mathbf{c}'(\theta',\theta,\mathbf{c})}{\mathbf{c}}\right\}^{-\frac{1}{\psi}}\left\{\frac{v'_O}{\mathcal{R}v'_O}\right\}^{\frac{1}{\psi}-\gamma} \\ 1+h\mathbf{i}(\mathbf{c},\theta) &= \mathbf{E}\Big[\Lambda(\theta'|\mathbf{c},\theta)\mathbf{p}(\mathbf{c}',\theta')\Big] \end{aligned} \tag{2.37}$$

where $v'_O$ is the normalized owner's continuation value function $v_O = \frac{V_O}{\mathbf{K}}$ and $\mathcal{R}v'_O$ is the associated certainty equivalent that is defined in equation 1.5.

3. We adopt value function iteration algorithm and iterate on the Bellman operator 2.30. We now describe a procedure to compute a set of policy functions including labor compensation function $w(u,\mathbf{c},\theta)$, optimal firm investment $k(u,\mathbf{c},\theta)$, continuation utility function contingent on $\eta'=1$, $u'_{\eta'=1}(\varepsilon',u,\mathbf{c},\theta)$, continuation utility

function contingent on $\eta' = 0$, $u'_{\eta'=0}(\varepsilon', u, \mathbf{c}, \theta)$, the Lagrangian Multipliers associated with binding firm-side constraint [32] 2.32, $\lambda'_{\eta'=0}(\varepsilon', u, \mathbf{c}, \theta)$ and value function $\tilde{v}(u, \mathbf{c}, \theta)$.

- we first present the set of optimality conditions

  (a) **continuation utility provision in the interior**: we first characterize the continuation utility provision none of the firm-side constraints or worker-side constraint bind:

$$\begin{aligned} &\Lambda(\theta'|\mathbf{c}, \theta) = \\ &\beta e^{-\gamma\varepsilon'}(1 - \delta + \mathbf{i}(\mathbf{c}, \theta))^{-\gamma} \times \\ &\left\{\frac{w(u'_{\eta'=1}(u, \varepsilon', \mathbf{c}', \theta'), \mathbf{c}', \theta')}{w}\right\}^{-\frac{1}{\psi}} \left\{\frac{u'_{\eta'=1}(u, \varepsilon', \mathbf{c}', \theta')}{\mathcal{R}u'}\right\}^{\frac{1}{\psi}-\gamma} \end{aligned} \tag{2.38}$$

$$\begin{aligned} &\Lambda(\theta'|\mathbf{c}, \theta) = \\ &\beta e^{-\gamma\varepsilon'}(1 - \delta + \mathbf{i}(\mathbf{c}, \theta))^{-\gamma} \times \\ &\left\{\frac{w(u'_{\eta'=0}(u, \varepsilon', \mathbf{c}', \theta'), \mathbf{c}', \theta')}{w}\right\}^{-\frac{1}{\psi}} \left\{\frac{u'_{\eta'=0}(u, \varepsilon', \mathbf{c}', \theta')}{\mathcal{R}u'}\right\}^{\frac{1}{\psi}-\gamma} \end{aligned} \tag{2.39}$$

  (b) **continuation utility provision when 2.34 binds**: when the worker-side constraint binds, we have the following risk sharing conditions for both scenarios $\eta' = 1$ and $\eta' = 0$

$$\Lambda(\theta'|\mathbf{c}, \theta) \geq \beta e^{-\gamma\varepsilon'}(1 - \delta + \mathbf{i}(\mathbf{c}, \theta))^{-\gamma} \left\{\frac{w(\underline{u}(\mathbf{c}', \theta'))}{w}\right\}^{-\frac{1}{\psi}} \left\{\frac{\underline{u}(\mathbf{c}', \theta')}{\mathcal{R}u'}\right\}^{\frac{1}{\psi}-\gamma} \tag{2.40}$$

  Note that we ignore the Lagrangian Multiplier for the worker-side constraint since it is not useful for the following computational steps.

  (c) **continuation utility when 2.33 binds**

$$\Lambda(\theta'|\mathbf{c}, \theta) \geq \beta e^{-\gamma\varepsilon'}(1 - \delta + \mathbf{i}(\mathbf{c}, \theta))^{-\gamma} \left\{\frac{w(\overline{u}(\mathbf{c}', \theta'), \mathbf{c}', \theta')}{w}\right\}^{-\frac{1}{\psi}} \left\{\frac{\overline{u}(\mathbf{c}', \theta')}{\mathcal{R}u'}\right\}^{\frac{1}{\psi}-\gamma} \tag{2.41}$$

[32] Note that we can also compute the Lagrangian Multipliers associated with binding firm-side constraint 2.33 in the state of $\eta' = 1$. As we will explain soon, solving for this multiplier is not useful because this multiplier does not interfere with the determination of other variables.

$\overline{u}(\mathbf{c}', \theta')$ is the level of continuation utility such that the value function attains 0 as in constraint 2.33. We ignore the Lagrangian Multiplier on the binding firm side constraint 2.33 because it does not interact with other equilibrium conditions.

(d) **continuation utility when 2.32 binds**

$$\begin{aligned}
&\Lambda(\theta'|\mathbf{c}, \theta)(1 + \lambda'_{\eta'=0}(u, \varepsilon', \mathbf{c}', \theta')) = \\
&\beta e^{-\gamma\varepsilon'}(1 - \delta + \mathbf{i}(\mathbf{c}, \theta))^{-\gamma}\left\{\frac{w(u'_{\eta'=0}(u, \varepsilon', \mathbf{c}', \theta'), \mathbf{c}', \theta')}{w}\right\}^{-\frac{1}{\psi}} \times \\
&\left\{\frac{u'_{\eta'=0}(u, \varepsilon', \mathbf{c}', \theta')}{\mathcal{R}u'}\right\}^{\frac{1}{\psi}-\gamma}
\end{aligned} \tag{2.42}$$

where $\lambda'_{\eta'=0}(u, \varepsilon', \mathbf{c}', \theta') \geq 0$ and the policy function for this multiplier needs to be solved for because relaxing the binding firm side constraint 2.32 introduces a new marginal benefit of investment which is present in equation 2.43.

(e) **optimality condition for investment decision**

$$\begin{aligned}
\mathbf{p}(\mathbf{c}, \theta) - \mathbf{E}\Big[\Lambda(\theta'|\mathbf{c}, \theta)\mathbf{p}(\mathbf{c}', \theta')\Big] &= \alpha\theta k(u, \mathbf{c}, \theta)^{\alpha-1} \\
+ (1 - \delta)\mathbf{E}\Big[\Lambda(\theta'|\theta, \mathbf{c})\lambda'_{\eta'=0}&(u, \varepsilon', \mathbf{c}', \theta')\mathbf{p}(\mathbf{c}', \theta')\Big]
\end{aligned} \tag{2.43}$$

- Grid points: the relevant state variables are $\mathbf{c}, \theta$ at the aggregate level and firm's promised continuation value $u$. We set up 20 grid points in the ergodic distribution of $\mathbf{c}$. Value and policy functions are smooth along the $\mathbf{c}$ dimension so adding more grid points to $\mathbf{c}$ does not change our results at all.

  Instead of setting grid point directly on firm level promised value $u$, we create grid points on the cutoff shocks that make 2.33 binds exactly with zero Lagrangian Multiplier, so as to capture the occasionally-binding nature of these constraints. This is essentially the idea of endogenous grid method as in Carroll (2006).

  In particular, let the grid points be denoted as $\{\overline{\varepsilon}_j^{\eta'=1}(\theta_L|\mathbf{c}, \theta)\}_{j=1,\ldots,nE}, \forall\theta \in \{\theta_L, \theta_H\}, \forall\mathbf{c}$ where $nE$ denotes the total number of grid points. That is, for each current aggregate shock $\theta$ and each point on the discretized space of $\mathbf{c}$, $\overline{\varepsilon}_j^{\eta'=1}(\theta_L|\mathbf{c}, \theta)$ satisfies 2.41 with a zero Lagrangian Multiplier, that is

$\forall\theta \in \{\theta_L, \theta_H\}$

$$\Lambda(\theta_L|\mathbf{c},\theta) = \beta e^{-\gamma\overline{\varepsilon}_j^{\eta'=1}(\theta_L|\mathbf{c},\theta)}(1-\delta+\mathbf{i}(\mathbf{c},\theta))^{-\gamma}\Big\{\frac{w(\overline{u}(\mathbf{c}',\theta_L),\mathbf{c}',\theta_L)}{w}\Big\}^{-\frac{1}{\psi}}\Big\{\frac{\overline{u}(\mathbf{c}',\theta_L)}{\mathcal{R}u'}\Big\}^{\frac{1}{\psi}-\gamma} \tag{2.44}$$

We now explain the advantage of indexing firm-level state variable $u_j$ with the cutoff shock defined above. Given the cutoff shock and the continuation utility associated with this shock, we can analytically construct the entire policy function for $u'_{\eta'=1}(\varepsilon',\cdots)$ for any realizations of idiosyncratic shocks.

$$u'_{\eta'=1}(\varepsilon',u_j,\mathbf{c}',\theta_L)\begin{cases} = \overline{u}(\mathbf{c}',\theta_L), \quad \text{if } \varepsilon' \leq \overline{\varepsilon}_j^{\eta'=1}(\theta_L|\mathbf{c},\theta) \\ \text{is determined by equation 2.46}, \quad \text{otherwise} \\ = \underline{u}(\mathbf{c}',\theta_L), \quad \text{if } \varepsilon' \geq \overline{\varepsilon}_j^{\eta'=1,A}(\theta_L|\mathbf{c},\theta) \end{cases} \tag{2.45}$$

$$\begin{aligned} &w(\overline{u}(\mathbf{c}',\theta_L),\mathbf{c}',\theta_L)^{-\frac{1}{\psi}}\overline{u}(\mathbf{c}',\theta_L)^{\frac{1}{\psi}-\gamma}e^{-\gamma(\overline{\varepsilon}_j^{\eta'=1}(\theta_L|\mathbf{c},\theta)-\varepsilon')} \\ &= w(u'_{\eta'=1}(\varepsilon',u_j,\mathbf{c}',\theta_L),\mathbf{c}',\theta_L)^{-\frac{1}{\psi}}u'_{\eta'=1}(\varepsilon',u_j,\mathbf{c}',\theta_L)^{\frac{1}{\psi}-\gamma} \end{aligned} \tag{2.46}$$

In the above equation, $\overline{\varepsilon}_j^{\eta'=1,A}(\theta_L|\mathbf{c},\theta)$ refers to the cutoff shock on worker side limited commitment constraint. At this level of shock, continuation utility achieves the lowest level that binding 2.34 can support and its associated multiplier is zero. Such a cutoff shock can be computed by comparing the definition of cutoff shocks in ratio form:

$$\overline{\varepsilon}_j^{\eta'=1,A}(\theta_L|\mathbf{c},\theta) = \overline{\varepsilon}_j^{\eta'=1}(\theta_L|\mathbf{c},\theta) + \frac{1}{\gamma\psi}\log\Big(\frac{w(\overline{u}(\mathbf{c}',\theta_L),\mathbf{c}',\theta_L)}{w(\underline{u}(\mathbf{c}',\theta_L),\mathbf{c}',\theta_L)}\Big) + (1-\frac{1}{\gamma\psi})\log\Big(\frac{\overline{u}(\mathbf{c}',\theta_L)}{\underline{u}(\mathbf{c}',\theta_L)}\Big) \tag{2.47}$$

Notice that for each point $\overline{\varepsilon}_j^{\eta'=1}(\theta_L|\mathbf{c},\theta)$, there is a correspond cutoff shock $\overline{\varepsilon}_j^{\eta'=1}(\theta_H|\mathbf{c},\theta)$. Such a cutoff shock can be pinned down by comparing the risk sharing conditions from $\theta$ to $\theta_L$ and that from $\theta$ to $\theta_L$ using the definition of cutoff shocks 2.44, on the $j^{th}$ point, that is

$$\begin{aligned}\overline{\varepsilon}_j^{\eta'=1}(\theta_H|\mathbf{c},\theta) = \overline{\varepsilon}_j^{\eta'=1}(\theta_L|\mathbf{c},\theta) &+ \frac{1}{\gamma}\log(\frac{\Lambda(\theta_L|\mathbf{c},\theta)}{\Lambda((\theta_H|\mathbf{c},\theta))}) \\ &+ \frac{1}{\gamma\psi}\log\Big(\frac{w(\overline{u}(\mathbf{c}',\theta_L),\mathbf{c}',\theta_L)}{w(\overline{u}(\mathbf{c}',\theta_H),\mathbf{c}',\theta_H)}\Big) \\ &+ \Big(1-\frac{1}{\psi}\Big)\log\Big(\frac{\overline{u}(\mathbf{c}',\theta_L)}{\overline{u}(\mathbf{c}',\theta_H)}\Big)\end{aligned} \tag{2.48}$$

Using the same step as in 2.45, we can solve the continuation utilities $u'_{\eta'=1}(\varepsilon', u_j, \mathbf{c}', \theta_H)$.

- We now discuss the computation to solve for the rest of policy functions. To begin with, for cutoff shock of the binding firm side constraint 2.33, $\overline{\varepsilon}_j^{\eta'=1}(\theta'|\mathbf{c},\theta)$, there is a related cutoff shock, $\hat{\varepsilon}_j^{\eta'=0}(\theta'|\mathbf{c},\theta)$, together with an associated continuation utility $\hat{u}_j^{\eta'=0}(\theta'|\mathbf{c},\theta)$ such that 2.32 binds with a zero Lagrangian Multiplier. However, the utility, cutoff shock and Lagrangian Multipliers for binding constraint 2.32 also interferes with the optimal investment decisions. Therefore, we jointly solve for

$$\Big\{k(u_j,\mathbf{c},\theta), \hat{\varepsilon}_j^{\eta'=0}(\theta'|\mathbf{c},\theta), \hat{u}_j^{\eta'=0}(\theta'|\mathbf{c},\theta), u'_{\eta'=0}(u_j,\varepsilon',\mathbf{c}',\theta'), \lambda'_{\eta'=0}(u_j,\varepsilon',\mathbf{c}',\theta')\Big\}$$

for some endogenously determined $u_j(\mathbf{c},\theta)$ which we explain later, using the following system of nonlinear equations that include the optimal investment decision 2.43 that depends on all five quantities, the binding firm-side constraint 2.32 for $\Big\{\hat{\varepsilon}_j^{\eta'=0}(\theta'|\mathbf{c},\theta), \hat{u}_j^{\eta'=0}(\theta'|\mathbf{c},\theta), k(u_j,\mathbf{c},\theta)\Big\}$, the binding firm-side constraint 2.32 for $\Big\{u'_{\eta'=0}(u_j,\varepsilon',\mathbf{c}',\theta'), k(u_j,\mathbf{c},\theta), \hat{\varepsilon}_j^{\eta'=0}(\theta'|\mathbf{c},\theta)\Big\}$, equation 2.49 for $\Big\{\lambda'_{\eta'=0}(u_j,\varepsilon',\mathbf{c}',\theta'), u'_{\eta'=0}(u_j,\varepsilon',\mathbf{c}',\theta')\Big\}$, equation 2.50 for $\Big\{\hat{\varepsilon}_j^{\eta'=0}(\theta'|\mathbf{c},\theta), \hat{u}_j^{\eta'=0}(\theta'|\mathbf{c},\theta)\Big\}$.[33]

$$\begin{aligned}&\lambda'_{\eta'=0}(u_j,\varepsilon',\mathbf{c}',\theta') = \\ &e^{-\gamma\overline{\varepsilon}_j^{\eta'=1}(\theta'|\mathbf{c},\theta)+\gamma\varepsilon'}\Big\{\frac{w(\overline{u}(\mathbf{c}',\theta'),\mathbf{c}',\theta')}{w(u'_{\eta'=0}(u_j,\varepsilon',\mathbf{c}',\theta'),c',\theta')}\Big\}^{-\frac{1}{\psi}}\Big\{\frac{\overline{u}(\mathbf{c}',\theta')}{u'_{\eta'=0}(u_j,\varepsilon',\mathbf{c}',\theta')}\Big\}^{\frac{1}{\psi}-\gamma} - 1\end{aligned} \tag{2.49}$$

[33] This procedure is described at conceptual level. In fact, with some efforts we can further simplify this nonlinear equation system. More details are available upon request.

$$\hat{\varepsilon}_j^{\eta'=0}(\theta'|\mathbf{c},\theta) = \overline{\varepsilon}_j^{\eta'=1}(\theta'|\mathbf{c},\theta) + \frac{1}{\gamma\psi}\log\Big(\frac{w(\overline{u}(\mathbf{c}',\theta'),\mathbf{c}',\theta')}{w(\hat{u}_j^{\eta'=0}(\theta'|\mathbf{c},\theta),\mathbf{c}',\theta')}\Big)$$
$$+ (1-\frac{1}{\gamma\psi})\log\Big(\frac{\overline{u}(\mathbf{c}',\theta')}{\hat{u}_j^{\eta'=0}(\theta'|\mathbf{c},\theta')}\Big) \tag{2.50}$$

Equation 2.49 and 2.50 are derived by comparing the risk sharing condition of the cutoff shock grid point and the constraint agents.

With all the policy function solved, we can use the same tricks as in previous step and construct the continuation utility function $u'_{\eta'=0}(u_j,\varepsilon',\mathbf{c}',\theta'),\forall\varepsilon'$.

The most important step to solve for the dynamic programming problem 2.30 is to solve for this system of nonlinear equations described above. We use various of numerical tools and parallel computing techniques to accelerate this procedure. We use Intel Fortran Complier that is bundled with Parallel Computational toolbox such as OpenMP. 1D cubic spline and 2D Bspline Interpolation methods from the commercial IMSL numerical library are implemented whenever interpolating function values is required. We use an efficient numerical integration method which is the QAGS method to evaluate the integration in equation 2.43.

- with all policy functions obtained in previous steps, we use the definition of cutoff shock 2.44 to update the policy function for labor compensation $w$ on each set of grid points. We then use the promise keeping constraint 2.31 to back out the level of current promised utility associated with each cutoff shock. We keep iterating until the Bellman equation converges at the criteria of $10^{-7}$.

### Finding equilibrium prices

We now explain our modification of Krusell and Smith (1997) algorithm to find the equilibrium prices for capital and consumption numeraire. It is useful to first to define a problem built on 2.30.

$$\tilde{v}^{\mathcal{CMK}}(u|\mathbf{c},\theta,p) = \max_{u',k,w} \Big\{ \theta k^{\alpha} - w - pk + \mathbf{E}\Big[\Lambda'(\mathbf{c}'|\mathbf{c},\theta)(1-\delta+\mathbf{i}(\mathbf{c},\theta))e^{\varepsilon'}\tilde{v}(u'|\mathbf{c}',\theta')\Big] + k(1-\delta)\mathbf{E}\Big[\Lambda'(\mathbf{c}'|\mathbf{c},\theta)\mathbf{p}(\mathbf{c}',\theta')\Big] \Big\} \tag{2.51}$$

$$s.t. \quad 2.31, 2.32, 2.33, 2.34$$

Thus, firms make optimal choices based on an arbitrary current value $p$ for the capital price. Firms take the current price of capital to equal $p$ and perceive future capital prices to be given by the function 2.36. Also the law of motion for owner' consumption share is perceived as equation 2.35. Since price of capital deviates from the assumed price function 2.36 for only one period, the continuation value of the firm coincides with the value function in problem 2.30.

This problem will generate investment decisions, $k^{\mathcal{CMK}}(u,\mathbf{c},\theta|p)$ and continuation utilities (for both realizations of $\eta'$ and we abstract from the subscript that represents the state of $\eta'$ shock to simplify notations), $u'^{\mathcal{CMK}}(u,\varepsilon',\mathbf{c}',\theta'|p)$. We are ready to describe the simulation procedure to update the price function 2.36 and law of motion 2.35.

1. solve the optimal contract problem 2.30 using the algorithm described in the previous section. Obtain a set of policy functions $w(u,\mathbf{c},\theta), k(u,\mathbf{c},\theta)$ and value function $\tilde{v}(u,\mathbf{c},\theta)$.

2. simulation stage

   (a) at a time point $t$ of a simulation path, we have a distribution of firms $\phi_t(u)$, $\mathbf{c}_t$, simulated aggregate state $\theta_t, \theta_{t+1}$ and a predicted value for owner's consumption share $x_{t+1=\Gamma_{\mathbf{c}}(\mathbf{c}_t,\theta_t,\theta_{t+1})}$

   (b) use a robust root-find algorithm (we use Brent's method) and find the equilibrium price of capital $p_t^*$ that clears the capital market:

   $$\int k^{\mathcal{CMK}}(u,\mathbf{c}_t,\theta_t|p_t^*)\phi_t(u)du = 1$$

   In particular, Brent's method generates different values of $p$. At each trial value $p$, we solve problem 2.51 for policy function $k^{\mathcal{CMK}}(u,\mathbf{c}_t,\theta_t|p)$. We then evaluate the integral using the existing measure of firms $\phi_t(u)$. If market is clear, we stop. Otherwise, we move to the next trial value and repeat this step. The market clearing error is set to be $10^{-7}$.

(c) Step b) also provides us with the continuation utility function under market clearing price $p_t^*$, $u'^{\mathcal{CMK}}(u, \varepsilon', \mathbf{c}_{t+1}, \theta_{t+1}|p_t^*)$. By equation 1.25, we use the continuation functions and simulate a large number of idiosyncratic shocks to construct the measure of firms in the next period, $\phi_{t+1}(u)$.

(d) compute the goods market clearing owner's consumption share $\mathbf{c}_{t+1^{MC}}$ via the goods market clearing condition 1.23.

$$\begin{aligned}\mathbf{c}_{t+1}^{MC} =& \theta_{t+1} \int k(u, \mathbf{c}_{t+1}, \theta_{t+1})^{\alpha} \phi_{t+1}(u) du \\ & - \int w(u, \mathbf{c}_{t+1}, \theta_{t+1}) \phi_{t+1}(u) du - \mathbf{i}(\mathbf{c}_{t+1}, \theta_{t+1}) - \frac{h}{2} \mathbf{i}(\mathbf{c}_{t+1}, \theta_{t+1})^2\end{aligned}$$

(e) obtain long time series (T=5000 observations in our experiment) for $\{\mathbf{c}_t^{MC}\}_{t=1}^{T}$, $\{p_t^*\}_{t=1}^{T}$. Fit a new price function $\Gamma_{\mathbf{p}}$, law of motion $\Gamma_{\mathbf{c}}$ and replace the old one[34] .

3. If the pricing function and law of motion converge, we stop. The stopping criteria, which is the maximum absolute difference in forecasting and pricing coefficients between two adjacent iteration, is set to be $10^{-5}$. We also check the regression $R^2$ for both converged 2.36 and 2.35. The minimum $R^2$ for both specifications is close to 99.85%. Otherwise, continue from step 1 with updated pricing function and law of motion.

[34] We use the dampening trick that is standard in solving heterogeneous agent model. We assume that the updated price and law of motion are convex combination of the recent obtained one and those from last iteration. The weight on the recent obtained rules is set to be 0.1.