A Dynamic Agency Based Asset Pricing Model with Production

Jincheng Tong and Chao Ying*

University of Minnesota

Jincheng Tong - Job Market Paper

Abstract

We 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.

JEL Code: E21, E23, E32, E44, G12

Keywords: Agency Frictions, Dynamic Optimal Contract, General Equilibrium, Heterogeneous Agents, Production-Based Asset Pricing, Value Premium, Investment, Incomplete Market

This draft: November 12, 2018

*Jincheng Tong (tongx160@umn.edu), Chao Ying (yingx040@umn.edu) are affiliated with the Carlson School of Management, University of Minnesota. We are particularly grateful to Hengjie Ai, Frederico Belo, Anmol Bhandari and Robert Goldstein for valuable guidance and encouragement. The authors would like to thank Philip Bond, Jaroslav Borovicka, Riccardo Colacito, Andrea Eisfeldt, Murray Frank, 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 for their helpful comments. The financial support from the Becker Friedman Institute at the University of Chicago is greatly acknowledged.

1 Introduction

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. Due to the owner-side lack of commitment, adverse shocks to workers' compensations are not fully insured. 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 due to a high adjustment cost parameter leads to large variation in capital prices but smooth aggregate investment, as 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 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

[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.

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 the 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. In this line of research, agents (or workers, as in our context) who realize large positive income shocks have a stronger incentive to renege on the contract because of higher outside values. We emphasize the prominence of firm-side incentive constraint that is different from these studies.

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 respects. 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 volatility of the risk-free interest

[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))

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

[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. (forthcoming), 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.

market clearing conditions. Similar procedures have 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.

2 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 to form an firm and firms start 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 γ and intertemporal elasticity of substitution (IES) ψ. The subjective discount factor β is also common across agents.

2.1 Technology

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

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

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

We assume that capital investment I made in the current period can be immediately put to use for producing output. Every period, firms first observe realizations of aggregate and firm-specific shocks. With the knowledge of their productivities, firms decide how much investment goods to acquire in the capital market and they can do so within the same period.

We do not assume one period time-to-build in the physical capital. As can be seen clearly in section 3, this assumption simplifies the solution to firm's recursive maximization problem by reducing the number of relevant state variables by one.

Output is produced according to a standard Cobb-Douglas production function with capital K_t and labor L_t as two production factors in 2. Z_t is firm's idiosyncratic productivity and $\mathbf{A}_t$ is productivity at the aggregate level that is identical to all firms. We assume that 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.

$$Y_t = \mathbf{A}_t \left(Z_t L_t\right)^{1-\alpha} K_t^{\alpha} \tag{2}$$

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

$$\mathbf{A}_t = \theta_t \mathbf{K}_t^{1-\alpha} \tag{3}$$

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 effect is akin to the capital externality model as in Romer (1986): a higher aggregate capital stock makes individual firms more productive, but firms do not internalize this effect of the capital input decision on the aggregate capital stock. As will be more clear the next section, this assumption essentially leads to the homotheticity properties of allocations in our model and all quantity variables can be scaled by the aggregate capital stock $\mathbf{K}$ in the current period which further simplifies our model. θ is a Markov process of exogenous productivity shocks and it can take two possible realizations $\{\theta_H, \theta_L\}$ with transition probability matrix π.

Firm's productivity S follows a random walk process with ε being firm-specific shock:

$$\log(Z_{t+1}) = \log(Z_t) + \varepsilon_{t+1} \tag{4}$$

where ε 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 θ, with the density function $f(\varepsilon|\theta)$. We use $(\theta_t, \varepsilon_t)$ to represent time t exogenous shocks for a firm. We denote the history of shocks up to time t as $(\theta^t, \varepsilon^t) = \{\theta_s, \varepsilon_s\}_{s=0}^t$.

Dividend paid by the firm 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. It's worth noting that firm-level investment does not incur other costs such as investment adjustment cost which is a common feature in the investment based asset pricing literature.

Firms and workers exit the economy if they receive an exogenous death shock that arrives with probability κ. For simplicity, we assume that within a match, 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 evaporate. κ fraction of new born firms replace firms that are hit by death shocks and exit the economy. We defer the discussion about entry and exit after we characterize firm's value functions in later section.

2.2 Contracting environment

When a firm starts up, it employs a worker by entering matching market and initiating a contractual relationship with that worker. Firms offer workers state-contingent compensation plan as a function of histories of firm-specific and aggregate shocks. Workers accept the offers if the life-time utility derived from the contract is 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. This point is made clear in the section where we discuss owners' utility maximization problem.

Let $\Lambda_{t,t+\tau}(\theta^{t+\tau}|\theta^t)$ denote the price of a claim to one unit of consumption in state $\theta^{t+\tau}$ denominated in state θ^t consumption numeraire. Also let $\mathbf{p}_t(\theta^t)$ be the price of one unit of capital at time 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:

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

Workers' utility is determined by the state-contingent compensation policies $\mathcal{W}_j$ of an contract that 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

$$\mathcal{R}_t U_{t+1} = \left(\mathbf{E}\left[U_{j,t+1}^{1-\gamma}(\mathcal{W}_j, \mathcal{K}_j|\theta^{t+1}, \varepsilon_j^{t+1})\middle|\theta^t, \varepsilon_j^t\right]\right)^{\frac{1}{1-\gamma}} \tag{5}$$

The key agency friction is that neither side of the contract can fully commit to it. Specifically, to smooth out workers' income fluctuations can be costly for firms. Owners have an incentive to terminate an existing labor relationship by shutting down firms and pursue its outside option. Once owner chooses to default, it loses firm's future cash flow stream and evades its obligations to insure workers. Owner seizes the full amount of capital and brings it to the capital market to trade.

We further assume that whether owner can sell capital on the market after default is random and it depends on the realization of a random variable η. With probability ρ, η takes the value of 1 meaning owner can sell the whole amount of capital they seize at the prevailing market price. In this case, owner's outside value is the market value of firm's capital. With probability $1-\rho$, $\eta = 0$ that represents the situation in which owner cannot sell capital at all. Therefore, if $\eta = 0$ owner's outside value is zero. The severity of the outside is closely related with the probability parameter ρ: a more lucrative outside due to a higher chance of selling the capital on market gives owner more incentive to default on the labor contract.

If worker decides to default, it loses a fraction of productivities or human capital and the firm-worker pair is dissolved. The interpretation is that part of worker's productivity or human capital is firm-specific. Once worker separates with firm, it loses the fraction of human capital that is attributable to the firm. After separation, worker can immediately match with another firm who offers an initial utility that depends on worker's productivity. We use $\underline{U}(\theta^t, \varepsilon_j^t)$ to denote the outside option of a worker after history $(\theta^t, \varepsilon_j^t)$. Worker accumulates sufficient amount of human capital in a firm that receives a sequence of positive shocks. With high level of human capital accumulated, the worker may be better off by exiting the existing labor relationship with its employer and matching with another firm, despite the human capital loss upon separation. The worker side constraint can be connected to theories of downside wage rigidity such as Harris and Holmstrom (1982) and solvency constraints and asset pricing such as Kehoe and Levine (1993) and Alvarez and Jermann (2000).

A worker with initial productivity S_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 which guarantee that optimal 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:

$$\max_{\mathcal{W},\mathcal{K}} V_0[\mathcal{W}, \mathcal{K}|\theta_0, \varepsilon_0] \tag{6}$$

$$U_0(\mathcal{W}, \mathcal{K}|\theta^0, \varepsilon^0) \geq U_0 \tag{7}$$

$$U_t(\mathcal{W}|\theta^t, \varepsilon^t) \geq \underline{U}_t(\theta^t, \varepsilon^t) \quad \forall(\theta^t, \varepsilon^t) \tag{8}$$

$$V_t(\mathcal{W}, \mathcal{K}|\theta^t, \varepsilon^t, \eta_t) \geq \eta_t \mathbf{p}_t(\theta^t) K_t \quad \forall(\theta^t, \varepsilon^t, \eta_t) \tag{9}$$

Equation 7 is the participation constraint for the worker that says the initial utility should be no less than workers' reservation utility U_0. Equations 8 and 9 correspond to limited commitment constraints on workers side and on firms side, respectively.

Workers are assumed to be hand-to-mouth in our model. Unlike standard incomplete market models in which agents build up wealth buffers and self insure by trading state non-contingent asset, workers have access to state-contingent insurance in the form of compensation contracts but limited commitment frictions restrict the set of feasible compensation contracts. Without limited commitment problem, workers would be completely insured with respect to their idiosyncratic shocks.

2.3 Recursive formulation of optimal contracting problem

In this section we recast firm's optimal contracting problem recursively, with the help of the joint distribution of promised utility U and firm productivity Z, $\Phi(U, Z)$ as a 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 Φ and aggregate state θ.

$$V(Z, K_{-1}, U|\Phi, \theta) = \max_{U', K, W} \left\{ \mathbf{A} Z^{1-\alpha} K^{\alpha} - W - \mathbf{p}(\Phi, \theta)(K - K_{-1}) + \mathbf{E}\Big[\Lambda'(\Phi', \theta'|\Phi, \theta) V(Z', K(1-\delta), U'|\Phi', \theta')\Big] \right\} \tag{10}$$

$$s.t. \quad U = \left\{(1-\beta) W^{1-\frac{1}{\psi}} + \beta \mathcal{R} U'^{1-\frac{1}{\psi}}\right\}^{\frac{1}{1-\frac{1}{\psi}}} \tag{11}$$

$$V(Z', K(1-\delta), U'|\Phi', \theta') \geq \eta' \mathbf{p}'(\Phi', \theta')(1-\delta) K \quad \forall(Z', \eta', \Phi', \theta') \tag{12}$$

$$U'(Z'|\Phi', \theta') \geq \underline{U}(Z'|\Phi', \theta') \quad \forall(Z', \Phi', \theta') \tag{13}$$

Firm's dividend D is simply output less compensation and cost of new investment, $D = \mathbf{A}Z^{1-\alpha}K^{\alpha} - W - \mathbf{p}(\Phi,\theta)(K - K_{-1})$. Notice that in firm's value function 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 choices of this dynamic programming problem. With this observation, we define the "cash flow value" of a firm $\tilde{V}$ which is the difference between firm's market value less the market value of firm's capital stock:

$$\tilde{V}(Z, U|\Phi,\theta) = V(Z, K_{-1}, U|\Phi,\theta) - \mathbf{p}(\Phi,\theta)K_{-1} \tag{14}$$

Intuitively, a firm is valuable because of the present value of firm's cash flow captured by $\tilde{V}$, and the market value of firm's capital stock which is the term $\mathbf{p}K_{-1}$.

Replace the market value V with the cash flow value $\tilde{V}$ using equation 14, the firm's value maximization problem now becomes:

$$\begin{aligned}
&\tilde{V}(Z, U|\Phi,\theta) = \\
&\qquad \max_{U',K,W} \Bigg\{ \mathbf{A}Z^{1-\alpha}K^{\alpha} - W - \mathbf{p}(\Phi,\theta)K + \mathbf{E}\left[\Lambda'(\Phi',\theta'|\Phi,\theta)\tilde{V}(Z',U'|\Phi',\theta')\right]) \\
&\qquad\qquad + (1-\delta)\mathbf{E}\Big[\Lambda'(\Phi',\theta'|\Phi,\theta)\mathbf{p}'(\Phi',\theta')K\Big]\Bigg\}
\end{aligned} \tag{15}$$

subject to constraints 11, 12 and 13.

2.4 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 investment goods using a very 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 $\mathbf{I}(\Phi, \theta)$.

2.5 Owners' utility maximization problem

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

$$V_O(\Phi, \theta) = \max_{\mathbf{C}, S'(U,Z), S_I'} \Big\{(1-\beta)\mathbf{C}^{1-\frac{1}{\psi}} + \beta\mathcal{R}(V_O')^{1-\frac{1}{\psi}}\Big\}^{\frac{1}{1-\frac{1}{\psi}}} \tag{16}$$

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

$\mathbf{C}$ 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 $S_I = 1$.

The owner's income is derived from two sources: the total profit of each individual firm and the profit of the investment sector. Therefore, the α parameter in firm's production function that represents the output elasticity with respect to capital does not equal capital share in our economy because the owner also earn profits from the investment sector.

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 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, its 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 which is useful for our defining our notion of recursive competitive equilibrium

$$\mathbf{C}(\Phi, \theta) + \iint W(U, Z|\Phi, \theta)\Phi(U, Z)dUdZ + \mathbf{I}(\Phi, \theta) + H(\mathbf{I}, \mathbf{K})(\Phi, \theta) = \iint Y(U, Z|\Phi, \theta)\Phi(U, Z)dUdZ \tag{17}$$

The resource constraint says the total output of firms on the right hand side of equation 17, is split among the owner's consumption $\mathbf{C}(\Phi, \theta)$, workers' total compensation and total cost of investment $\mathbf{I}(\Phi, \theta) + H(\mathbf{I}, \mathbf{K})(\Phi, \theta)$.

[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 law of motion of total capital is

$$\iint K(U, Z|\Phi, \theta)\Phi(U, Z)dUdZ = \mathbf{K}(\Phi, \theta) \tag{18}$$

That is the total demand of capital equals to the total capital available in the economy. Total amount of capital available on the market consists of two components: firms' undepreciated capital and new investment made by the investment sector.

2.6 Recursive competitive equilibrium

In this section, we define a recursive competitive equilibrium for our economy. The homogeneity property resulting from assumptions on preference, productivity specification and production function enable us to construct an equilibrium that has two aggregate state variables (ϕ, θ). θ is the Markov state for aggregate productivity shock. ϕ is a one-dimensional density that summarizes firms' type. We show that the random walk assumption on firm-specific productivity allows us to reduce firms' distribution Φ to a one dimensional distribution ϕ.

Let $\tilde{V}(U, Z|\phi, \theta)$ be the value function that attains the optimum of the firm's maximization problem 15. In Appendix, we show that $\tilde{V}(U, Z|\phi, \theta)$ satisfies

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

for some function $\tilde{v}(\cdot|\phi, \theta)$ that represents normalized firm cash flow value. Intuitively, firm's franchise value is proportional to its idiosyncratic productivity Z because of the unit root assumption on firm-specific productivity and homogeneous property of production function. In addition, $\mathbf{K}$ represents the aggregate growth of the aggregate economy that is common to all firms. Similarly, we introduce the definition of normalized utility, normalized compensation and normalized capital stock:

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

Similarly, we introduce detrended aggregate state variables: $\mathbf{c} = \frac{\mathbf{C}}{\mathbf{K}}, \mathbf{i} = \frac{\mathbf{I}}{\mathbf{K}}$. Aggregate capital evolves according to $\mathbf{K}' = (1-\delta)\mathbf{K} + \mathbf{I}$ and this immediately implies that the growth rate of the economy is $\frac{aggk'}{\mathbf{K}} = (1 - \delta + \mathbf{i})$.

The normalized continuation utility u can be interpreted as a measure of a workers share in the firms valuation, because U is the total utility delivered by the compensation contract and Z is firm's productivity and hence closely related to firm's size. We express the optimal contracting problem recursively 15 in terms of normalized values as

$$\begin{aligned}\tilde{v}(u|\phi,\theta) = \max_{u',k,w}\Bigg\{ & \theta k^{\alpha} - w - \mathbf{p}(\phi,\theta)k + \mathbf{E}\Big[\Lambda'(\phi',\theta'|\phi,\theta)(1-\delta+\mathbf{i}(\phi,\theta))e^{\varepsilon'}\tilde{v}(u'|\phi',\theta')\Big] \\ & + k(1-\delta)\mathbf{E}\Big[\Lambda'(\phi',\theta'|\phi,\theta)\mathbf{p}'(\phi',\theta')\Big]\Bigg\}\end{aligned} \tag{19}$$

$$u = \Big\{(1-\beta)w^{1-\frac{1}{\psi}} + \beta\mathcal{R}(u')^{1-\frac{1}{\psi}}\Big\}^{\frac{1}{1-\frac{1}{\psi}}} \tag{20}$$

$$(1-\delta+\mathbf{i}(\phi,\theta))e^{\varepsilon'}\tilde{v}(u'|\phi',\theta') \geq (\eta'-1)(1-\delta)\mathbf{p}'(\phi',\theta')k \quad \forall(\phi',\eta',\theta',\varepsilon') \tag{21}$$

$$u'(u|\mathbf{c}',\theta') \geq \underline{u}(\phi',\theta') \tag{22}$$

Constraints 21 and 22 are the normalized counterparts of 12 and 13. Here we have assumed that the outside values for the worker $\underline{U}(\phi,\theta,Z\mathbf{K})$ can be expressed as $\underline{u}(\phi,\theta)Z\mathbf{K}$. This restriction is without loss of generality if workers retain a fraction of their labor productivity when they leave the firm. Equation 20 is the promise keeping constraint that guarantees the contract delivers the initial promised value. Solving the normalized problem 19 gives optimal investment decision $k(u,\phi,\theta)$, compensation policy $w(u,\phi,\theta)$ and continuation utility contingent on different values of η:$u'_{\eta'=0}(\varepsilon',u,\theta'|\phi,\theta)$ and $u'_{\eta'=1}(\varepsilon',u,\theta'|\phi,\theta)$.

Finally, we describe the construction of our aggregate state variable ϕ, 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 19 and recall that $\Phi(U,Z)$ denote the joint distri-

[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.

bution 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 17. To illustrate the construction of the normalized measure ϕ, we use the total labor compensation term in the aggregate resource constraint 17 as an example. We can re-write 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 17 now reads

$$\mathbf{c}(\phi,\theta) + \int w(u,\phi,\theta)\phi(u|\theta)du + \mathbf{i}(\phi,\theta) + \frac{h}{2}\mathbf{i}(\phi,\theta)^2 = \int k(u,\phi,\theta)^\alpha\phi(u|\theta)du \tag{23}$$

The capital market clearing condition 18 can be written as

$$\int k(u,\phi,\theta)\phi(u|\theta)du = 1 \tag{24}$$

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

$$\forall\tilde{u}\quad \phi'(\tilde{u}|\theta') = \int\phi(u|\theta)\int e^{\varepsilon'}f(\varepsilon'|\theta')\Big(\rho\mathcal{I}_{u'_{\eta'=1}(u,\theta',\varepsilon'|\phi,\theta)=\tilde{u}} + (1-\rho)\mathcal{I}_{u'_{\eta'=0}(u,\theta',\varepsilon'|\phi,\theta)=\tilde{u}}\Big) \tag{25}$$

where $\mathcal{I}$ is an indicator that takes the value one if the continuation utility for firm-specific shock ε' 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 κ 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 $\overline{u}(\phi, \theta)$. Competition in the matching market drives the initial value of the firm $\tilde{v}(\overline{u}(\phi, \theta), \phi, \theta)$ to be zero and initial utility $\overline{u}(\phi, \theta)$ is determined by this zero profit condition. This assumption ensures that the total measure of firm is always one and the distribution of firm-specific productivity is stationary.

If a worker defaults on the contract, it keeps a λ 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 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 ϕ, $\Gamma(\theta'|\phi, \theta)$; a set of prices: the SDF $\{\Lambda(\theta'|\phi, \theta)\}$[8], prices of capital $\{\mathbf{p}(\phi, \theta)\}$; 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 $u'_{\eta'=1}(u, \theta', \varepsilon|\phi, \theta)$, $u'_{\eta'=0}(u, \theta', \varepsilon|\phi, \theta)$; 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. 24 is satisfied.

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

- owner solves her consumption and portfolio choice problem 16 and the equilibrium SDF is consistent with owner's IMRS derived from the solution to problem 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 in 19.
- given the policy functions for continuation utilities, the law of motion of the normalized measure ϕ satisfies 25.
- the policy functions and the normalized measure ϕ satisfy the goods market clearing condition 17.

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

3 Implications of optimal contract

In this section, we show that optimal contract has several important implications that help us jointly account for asset prices and investment dynamics. First of all, optimal contract generates uninsurable tail risk in labor earnings as in Ai and Bhandari (2018), Constantinides and Ghosh (2015) and Schmidt (2015). Uninsurable tail risks channel amplifies the conditional volatility of the pricing kernel and our production economy inherits the success of this mechanism in endowment economy. Secondly, we show that optimal contract generate a form of operating leverage channel. Specifically, heterogeneity in the severity of binding firm-side constraint is translated into heterogeneity in firms' cash flow risk exposures and hence heterogeneity in expected stock returns. Thirdly, we show that in a production environment, in the presence of uninsurable tail risks, investment creates a novel benefit that is to alleviate the binding firm-side constraint and improve workers' risk sharing conditions. This force is particularly prominent for unproductive firms burdened with significant labor commitment because workers' risk sharing conditions are poor in these firms.

3.1 Uninsurable tail risk

With perfect risk sharing, continuation utility does not respond to idiosyncratic shocks and thus normalized continuation utility $u'(\varepsilon', u, \theta'|\phi, \theta)$ must be inversely proportional to the idiosyncratic shock $e^{\varepsilon'}$. In particular, an extremely negative shock in ε' pushes $u'(\varepsilon', u, \theta'|\phi, \theta)$ toward infinity. However, with firm side limited commitment constraint, owner of an unproductive firm has strong incentive to renege on the contract because of the low continuation value of the firm. On the other hand, the magnitude of earnings' losses depend on whether owners can sell capital on the market. In the state where the owner is able to sell capital with $\eta = 1$, owner has stronger incentive to renege on the contract and therefore workers need to take a significant compensation cut to raise owner's continuation values. If the realization of η is 0, the owner has less incentive to default on the contract and owner can still provide some insurance for workers against adverse idiosyncratic shocks. To demonstrate these properties formally, we prove the following two propositions

Proposition 1. *In the state where $\eta = 1$, there exists $\overline{\varepsilon}(u, \theta'|\phi, \theta)$ with $\frac{\partial \overline{\varepsilon}(u, \theta'|\phi, \theta)}{\partial u} > 0$ such that*

$$u'_{\eta=1}(\varepsilon', u, \theta'|\phi, \theta) = \overline{u}(\phi', \theta') \quad \forall \varepsilon' \leq \overline{\varepsilon}(u, \theta'|\phi, \theta)$$

$u'_{\eta=1}$ denotes the continuation utility contingent on that owner can sell the capital on the capital market. And $\overline{u}(\phi', \theta')$ is the maximum utility that the owner can provide.

By proposition 1, sufficiently low realizations of firm-specific shock ε' are not associated with changes in continuation utility u'. As a result, the unnormalized continuation utility, U' moves one to one with ε' and so do the levels of future labor compensations. Thus earnings fall proportionally with negative shocks for all shocks below the cutoff $\overline{\varepsilon}(u, \theta'|\phi, \theta)$. In the upper panel of figure 1, we display the normalized continuation utility u' as a function of size of firm-specific shock ε. It is clear that for any shock below the cutoff shock $\overline{\varepsilon}(u, \theta'|\phi, \theta)$, continuation utility is held constant. For shocks above this cutoff shock, normalized utility is inversely related with shock, reflecting the risk sharing function of the optimal contract.

The next proposition highlights the link between risk sharing and firm's investment choices.

Proposition 2. *In the state where $\eta = 0$, there exists a a cutoff $\hat{\varepsilon}(u, \theta'|\phi, \theta)$, a continuation utility associated with this cutoff $\hat{u}(u, \theta'|\phi, \theta)$ and firm's optimal investment decision $k(u|\phi, \theta)$ such that the firm-side limited commitment constraint binds as*

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

and for any $\varepsilon \leq \hat{\varepsilon}(u, \theta'|\phi, \theta)$, the continuation utility $u'_{\eta=0}(\varepsilon', u, \theta'|\phi, \theta)$ associated with this shock is given by

$$u'_{\eta=0}(\varepsilon', u, \theta'|\phi, \theta) = \tilde{v}^{-1}(e^{\hat{\varepsilon}(u,\theta'|\phi,\theta)-\varepsilon'}\tilde{v}(\hat{u}(u, \theta'|\phi, \theta)|\phi', \theta'))$$

Compared with proposition 1, workers' continuation utility for extremely adverse shock no longer falls one-to-one with the size of unfavorable shocks. Instead, proposition 2 shows that the continuation utility is still inversely related with the shock but unnormalized utility is no longer constant for large adverse shocks. That is, worker receives insurance when the firm-side limited commitment constraint binds but they would not get the full insurance as in the first best case. Lower panel of Figure 1 clearly illustrate this point. For shocks below the cutoff shock $\hat{\varepsilon}(u, \theta'|\phi, \theta)$, normalized continuation utility is decreasing with respect to shocks. However, the slope of utility to shocks is not as steep as in the unconstrained regions in which the realization of shock is above the cutoff shock.

Proposition 2 highlights the interaction between risk sharing and firm optimal investment. Figure 2 displays the effect of increasing capital stock on risk sharing. As firm increases its investment from k_1 to k_2, we observe that the cutoff shock decreases indicating that firm-side constraint in the state $\eta = 0$ is less likely to be binding. Therefore, a new benefit of investment emerges because investment also relaxes the binding firm-side constraint and improve worker's risk sharing. Note that this effect is only present when owner cannot sell firm capital she absconds with on the capital market.

3.2 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 intuition is much akin to the mechanism in Ai and Bhandari (2018) in an endowment economy. 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 Ai and Bhandari (2018) is that our model is setup in a production economy. This introduces two new economic forces compared with the literature on endogenous incomplete market and asset pricing: 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 under the optimal contract, by adjusting investment and production accordingly. Quantitatively, we show that despite the two counter forces that dampen the effect of uninsurable tail risks on the cyclicality of the owner's consumption share, our model is able to generate empirically plausible implications on asset prices and macroeconomic quantities.

3.3 Endogenous labor operating leverage

In figure 3b and 3a, 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 3a and 3b 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 θ_L than in θ_H. The steeper curvature of the compensation policy in θ_L reflects that marginal cost of providing insurance is higher for firms with equal committed value in θ_L than in θ_H.

3.4 Investment policy under optimal contract

In this section, we further investigate the relationship between investment and its effect on relaxing the binding firm-side constraint. First order condition with respect to investment k gives rise to the following optimality condition

$$\begin{aligned}\mathbf{p}(\phi,\theta) = \alpha\theta k(u,\phi,\theta)^{\alpha-1} + (1-\delta)\mathbf{E}\Big[\Lambda'(\phi',\theta'|\phi,\theta)\mathbf{p}'(\phi',\theta')\Big] \\ + (1-\delta)(1-\rho)\mathbf{E}\Big[\lambda'_{\eta'=0}(\varepsilon',u,\theta'|\phi,\theta)\Lambda'(\phi',\theta'|\phi,\theta)\mathbf{p}(\phi',\theta')\Big]\end{aligned} \tag{26}$$

As in standard neo-classical investment models, firms equalize marginal cost of investment and marginal benefit. At the margin, firms pay 1 unit of capital goods prices to acquire an additional unit of capital. On the benefit side, increasing capital stock by 1 unit simply endow them with more capital to sell next period. Also, firms produce more output at the rate of marginal product of capital. Those two channels are standard in neo-classical investment models.

The novel channel in our framework is the role of capital to relax the binding limited commitment problem. Let $\lambda'_{\eta'=0}(\varepsilon', u, \theta'|\phi, \theta)$ denote the Lagrangian multiplier associated with the binding firm-side constraint 21 in the state when owner cannot sell firm's capital. Consider a firm that receives a sequence of negative firm-specific shocks, such a firm face the situation that going forward, their constraints are binding in more states of the world. Risk sharing is poor for workers in these firms because workers are more likely to encounter with compensation cuts. To relax the binding firm-side constraint and improve risk sharing, these firms invest more relative to the first best case.

On the other extreme, consider a firm that is constantly hit by positive productivity shocks. In the first best world, these firms optimal investment decisions should be determined only by its productivity shocks and efficient capital allocation implies that a larger fraction of capital are in the hand of such productive firms. This is not the case with limited commitment problem. In equilibrium, there is always a group of firms that are constrained and their excessive demand for capital raises the price of capital in equilibrium. For unconstrained firms, this pecuniary externality effect affects their optimal capital decision because the equilibrium capital price is higher and they reduce their investment expenditures.

In figure 4, we visualize this economic mechanism. Equation 26 implies the following capital allocation rule without agency frictions: for a particular firm j with productivity Z_j

$$K_j^{FB} = \mathbf{K} Z_j \tag{27}$$

In the first best economy, firm capital is linear in aggregate capital and the proportion is its idiosyncratic productivity shock.In figure 3, we plot firms optimal investment policy functions as a function of normalized labor commitment. As explained, high labor commitment firms are firms that realize a sequence of unfavorable shocks and such firms overinvest than their first best investment while low u firms underinvest.

In the data, firm-level investment is positively autocorrelated. Standard convex adjustment cost of investment is a feature that delivers the positive autocorrelation of investment as in the data. Without agency frictions, first best capital allocation rule 27 implies investment is i.i.d over time and the autocorrelation of investment is zero. In our benchmark model with limited commitment problem, firm-level investment is also positive autocorrelated despite the absence of adjustment cost.

4 Numerical algorithm

One computational challenge in numerically solving our model is that the normalized measure ϕ is a state variable with infinite dimension. To make matters worse, to solve for the equilibrium needs to simultaneously solve for the two pricing function, capital and the SDF, using the capital market clearing condition 24 and goods market clearing condition 23. We use a procedure adopted from Krusell and Smith (1998) and replace ϕ with sufficient statistics that can accurately characterize the dynamics of ϕ over time and states. We further build on the numerical procedure in Krusell and Smith (1997) to jointly look for the market clearing prices for both markets.

The distribution ϕ enters the problem through its effect on the stochastic discount factor and market clearings. To approximate the law of motion of Γ and capital price $\mathbf{p}$, we guess 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 ϕ for two reasons. First, ϕ affects capital owners consumption, and this effect is completely summarized by $\mathbf{c}$. Second, ϕ is a sufficient statistic to forecast future prices. Similarly, capital price depends on ϕ because ϕ affects the supply of aggregate investment through its effect on owner's SDF. And again, this effect is captured by the dynamics of $\mathbf{c}$.

Our method is numerically efficient to clear the goods market 23 because given the law of motion of $\mathbf{c}$, the equilibrium stochastic discount factor is completely determined without solving for the optimal contract. Through market clearing conditions in equation 23, we observe that once we solve for optimal policy functions for firms investment and compensations, owner's consumption share is simply firms' total output less total labor shares and aggregate investment shares. This choice contrasts our algorithm from that in Krusell and Smith (1998) who use the first moment of the distribution of wealth as a sufficient statistic.

To solve for capital market clearing price is more complicated. 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 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. In appendix, we describe these steps in details in the appendix.

5 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 sample 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.

5.1 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 θ follow a two-state Markov Chain and θ can take two values $\{\theta_H, \theta_L\}$. The transition probability between different states is denoted as π. We label θ_L as recession states and θ_H as boom states. We choose the transition probability from boom to boom to be 0.9802 at a quarterly frequency in order to target the duration of U.S. economic booms which is about 12 years on average. Similarly, 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 α 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. Even with constant return to scale in production function, the parameter α cannot be interpreted as capital share as in the standard RBC for two reasons: first, profit of firms constitute only one component of owner's income because owner is also endowed with the claim to the profits of the investment goods sector. secondly, the standard RBC assumes competitive labor market and with constant return to scale technology labor share is counter-factually constant over business cycles. In our model, labor compensation is determined by the optimal risk sharing labor contract and we demonstrate that labor is counter-cyclical as in the data. We set the annual depreciation rate of physical capital δ to be 10% as in Kydland and Prescott (1982) and King and Rebelo

(1999).

Preference parameters including risk aversion γ, IES ψ and the discount factor β. IES is set to be 2 which is consistent with the long run risk literature following Bansal and Yaron (2004). All economics agents in our model have recursive preference and early resolution of uncertainty is preferred which share the similar preference assumptions of long run risk based models. However, the main difference with this literature in the parameterization of preference is that we require a much smaller 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.

In our model, firm-specific productivity follows a random walk process. Therefore, it is hard to directly discipline the stochastic firm-level productivity process using existing estimates on firm-specific productivity in the literature. We parameterize our idiosyncratic shock process using firm level sales data following the calibration strategy in Bloom (2009). 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, $\int e^{\varepsilon} f(\varepsilon|\theta) d\varepsilon = 1$. 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 $\varepsilon^{max} = \log \frac{1+\iota}{\iota}$.

As a baseline we will use moments from Salgado et al. (2017) that uses Compustat data to document the cyclical properties of firms' sales growth over business cycles for the sample 1963 to 2016 as our targets. The value of σ_θ is calibrated to match the cross-sectional standard deviation of annual sales growth.[9] The mixing probability p and the negative

[9] Colacito et al. (2016) adopt a different approach and highlight the importance of time-varying skewness in the distribution of expected growth prospects to explain several puzzling asset pricing facts. We follow Salgado et al. (2017) and use cross-sectional distributional moments of firms' sales growth to discipline our quantitative exercise.

Table 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	σ_θ	2.20%
transition probability from boom to boom (Quarterly)	$\pi(\theta_H \vert \theta_H)$	98.02%
transition probability from recession to recession (Quarterly)	$\pi(\theta_L \vert \theta_L)$	93.71%
production technology		
capital elasticity	α	0.22
aggregate adjustment cost	h	20
depreciation of physical capital	δ	10%
preference		
risk aversion	γ	4.5
Elasticity of intertemporal substitution	ψ	2
subjective discount factor	β	0.99
firm-specific shocks		
volatility of idiosynratic shock	σ_F	30.41%
probability of receiving tail shock: recession	ρ	7.76%
agency frictions		
present value of earning loss upon separation	μ	30%
probability of selling capital	η	96.06%

exponential parameter λ are chosen to match the negative Kelley Skewness of the annual sales growth in recession. Note that Salgado et al. (2017) also reports that in expansions, firms' annual sales growth distribution exhibit a right tail. Our model cannot speak to this fact because of the Gaussian shock structure in θ_H. Adding a positive exponential distribution component in boom does not affect our quantitative results.

Two key parameters, ρ which is the probability of selling capital on the market if owner defaults and λ that is workers' human capital loss upon separation, determine owner and workers' incentives to renege on the labor contract. For the parameter λ, 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 as reported in table 1. The probability of whether owner can sell capital, ρ controls owner's incentive to default on the contract and hence the severity of firm-side limited commitment friction. The form of punishment for the firm in default is the expected loss of its assets. One can think of this feature of the model to capture Chapter 11 renegotiation.

5.2 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 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.

Table 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 [10]. To clarify the effect of risk premium channel on investment, suppose the economy is in downturn and investment adjustment cost is high as in our calibration. 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 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 counteracting the dampening effect of adjustment cost on the reduction of investment expenditure.

5.3 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 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 compared to the endowment 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, agents are able to smooth out any risks in consumption once they have access to a production technology and produce their own consumption. A second

[10] 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.

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.

To disentangle the effect of these two channels on firm value, consider equation 14 that demonstrates that the value of a firm can be decomposed into two components: market value of firm capital $\mathbf{p}K_{-1}$ and the cash flow value $\tilde{V}$. Capital adjustment cost in the investment sector reduces the elasticity of supply of capital. Therefore, shifts in the aggregate demand for capital due to aggregate shocks are absorbed mostly by movements in the equilibrium price of capital $\mathbf{p}$. The variation in $\mathbf{p}$ induces variation in firm value, mostly through the market value of firm's capital term $\mathbf{p}K_{-1}$. Labor compensation is a key determinant of firm's franchise value $\tilde{V}$ because by definition $\tilde{V}$ equals the present discounted value of cash flow generated by a firm-worker pair. As a result, the endogenous labor leverage channel affects firm value mostly through its effect on the cash flow value term $\tilde{V}$. With these being said, the two channels also have confounding effects on firm's value. For example, the equilibrium price of capital affects firm's cash flow value because it matters for owner's outside values that control owner's incentives to default. Quantitatively, the main effects, rather than all confounding effects, are the major forces that shape the dynamics of firm value.

5.4 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 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 a little lower but roughly 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 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 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 [11]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 partly due to the higher cost of capital. As a result, low investment is associated with low stock market valuation and hence higher expected return going forward.

[11]By integrating equation 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.

5.5 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 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.

Table 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%

In the model, the difference in firm-level productivity drives the heterogeneity in 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 5 presents our main findings on these relevant characteristics. Prior ro portfolio formations, low productivity firms tend to be value firms with high BM

Table 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

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 engage in large disinvestment draws a clear distinction between our model and models that rely on costly irreversibility as the main mechanism to generate the value premium.

5.6 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), Anderson and Garcia-Feijo (2006) 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. In table 6, we first present the empirical evidence on the investment rate sorted

Table 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%

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 6.

5.7 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.

Table 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%
$\text{skewness}(\Delta \log(\text{sales}))\vert\text{recession}$	-2.12%	-1.15%
mean $\frac{W}{Y}$	53.13%	65.11%
$\sigma(\frac{W}{Y})$	65.50%	71.66%

Examining the first two rows of table 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 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 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 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 notoriously sparse 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.

6 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, to further illustrate the main mechanisms of our model and to contrast our model with existing literature, we present a set of tests. We develop 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. More importantly, it 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.

6.1 Investment decomposition framework

The intuition behind 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 28 provides a link between investment and cash flow and discount rate innovations when firm's marginal cost of investment equals its average Q [12].

$$\eta \ln(\frac{I_t}{K_t}) \approx \text{constant} + \mathbf{E}_t \left[\ln \frac{D_{t+1}}{K_{t+1}} + \sum_{j=1}^{\infty} (\rho^j \Delta \ln D_{t+j+1} - \rho^{j-1} \ln R_{t+j}) \right] \tag{28}$$

Equation 28 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 $\ln R_{t+j}$. ρ is a parameter that captures average level of stock market prices. η is an adjustment cost curvature parameter.

Equation 28 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

[12] see Kogan and Papanikolaou (2012) for detailed derivation of equation 28

returns.

$$\begin{aligned} N_{ik,t} &= \eta \ln \frac{I_t}{K_t} - \mathbf{E}_{t-1}\eta \ln \frac{I_t}{K_t} \\ N_{DR,t} &= (\mathbf{E}_t - \mathbf{E}_{t-1}) \sum_{j=1}^{\infty} \rho^{j-1} \ln R_{t+j} \\ N_{CF,t} &= (\mathbf{E}_t - \mathbf{E}_{t-1}) \left[\ln \frac{D_{t+1}}{K_{t+1}} + \sum_{j=1}^{\infty} \rho^j \Delta \ln D_{t+j+1} \right] \end{aligned} \tag{29}$$

Re-writing equation 28 with expressions for news component 29, we obtain the following key condition that is crucial to our empirical analysis:

$$N_{ik,t} = N_{CF,t} - N_{DR,t} \tag{30}$$

Condition 29 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.

Classic Q-theory predicts that Tobin's Q, which is the ratio between the market value of a firm's capital and its replacement cost, should perfectly summarizes a firm's investment opportunities. It is well understood that classic Q-theory is confronted with difficulties when researchers bring this key prediction of the theory to the data. Our variance decomposition requires much less stringent assumptions than the conditions under which investment is perfectly correlated with Tobin's Q. In fact, the only structural assumption underlying our procedure is the consistency condition 30 with the help of some forecasting model based on which economic agents form expectations on variables of interest. In our framework, news to investment and stock market returns can be directly computed using a forecasting VAR model. We impose condition 30 and isolate news to cash flow component. This approach has several advantages. First of all, in empirical works, researchers tie a firm's investment in physical capital to the market value of the firm normalized by a firm's physical capital stock. The cash flow generated by a firm could come from other business activities irrelevant to capital investment, or other production factors[13] Furthermore, Tobin's Q in the data could also be contaminated by measurement errors [14]. All these problems could render a weak connection between investment in physical capital and market value of a firm., much the opposite to what Q-theory predicts. On the contrary, our method does not require measures of cash flow in the data. By imposing consistency condition 30, our method is able to isolate the cash flow news that is driven by physical capital investment only, as long as news to investment rate and cost of capital are predicted accurately.

6.2 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.

[13] Merz and Yashiv (2007) address the importance of labor production factor to understand variation of aggregate Tobin's Q. Peters and Taylor (2017) shows that incorporating intangible capital improves the empirical performance of investment and Q relationship. Belo et al. (2018) decomposes firm value to physical capital, labor and intangible capital.

[14] See Erickson and Whited (2000) for example.

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

$$z_{t+1} = c + \Gamma z_t + e_{t+1} \tag{31}$$

where z_{t+1} is a m-by-1 state vector. c and Γ 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 31, 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,

$$N_{ik,t+1} = \mathbf{e}_1' u_{t+1} \tag{32}$$

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

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

$$N_{DR,t+1} = \mathbf{e}_3' \rho \Gamma (I - \rho \Gamma)^{-1} e_{t+1} \tag{33}$$

where Γ 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 ρ is a parameter that depends on average pd ratio at the aggregate level or firm-level. Previous literature find that setting ρ 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 ρ and find similar patterns that different values of ρ 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 31.

The cash flow news can be indirectly computed by consistency condition 30, 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.

6.3 VAR data

The choice of state variables is crucial in empirical implementation of cross-sectional variance decomposition of firm investment. The literature that attempts to understand the stock return variation shows that a minor change in the state variables can dramatically affect the conclusion drawn from the return decomposition, see for example Chen and Zhao (2009). In our implementation, we do not consider the effect of different choices of state variables as experimented in Chen and Zhao (2009). We only 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.

6.4 Firm-level evidence

We report the firm-level variance decomposition results in Panel A of table 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. On average, 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

Table 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%

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.

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 8 presents these results and draws a clear distinction between our model and adjustment cost models.

In our model, we first generate a time series for 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. To be consistent with the data, at every time period, we only look at firms that have existed in our simulated sample for 10 years. 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 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.

6.5 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 9, we report the variance decomposition results. The first row in the data panel presents the variance decomposition of aggregate investment rate obtained using

Table 9: Aggregate investment variance decomposition

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).

	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%

equation 34. 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 because high investment rate seem to forecast lower real cash flow growth. As a result, 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.

7 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: continuation utility when the owner can/cannot sell firm's capital. This figure plots the normalized continuation utility u (y-axis) as a function of idiosyncratic shock (x-axis). The upper panel corresponds to continuation utility provision when the firm owner cannot sell capital on the market after default while the lower panel plots the other possibility. In the upper panel, $\overline{\varepsilon}$ denotes the cutoff shock below which worker receives no insurance and unnormalized utility falls one to one with unfavorable shock realized. In the lower panel, $\hat{\varepsilon}$ is the cutoff shock below which worker is only partially insured and $\hat{u}$ is the continuation utility at $\hat{\varepsilon}$.

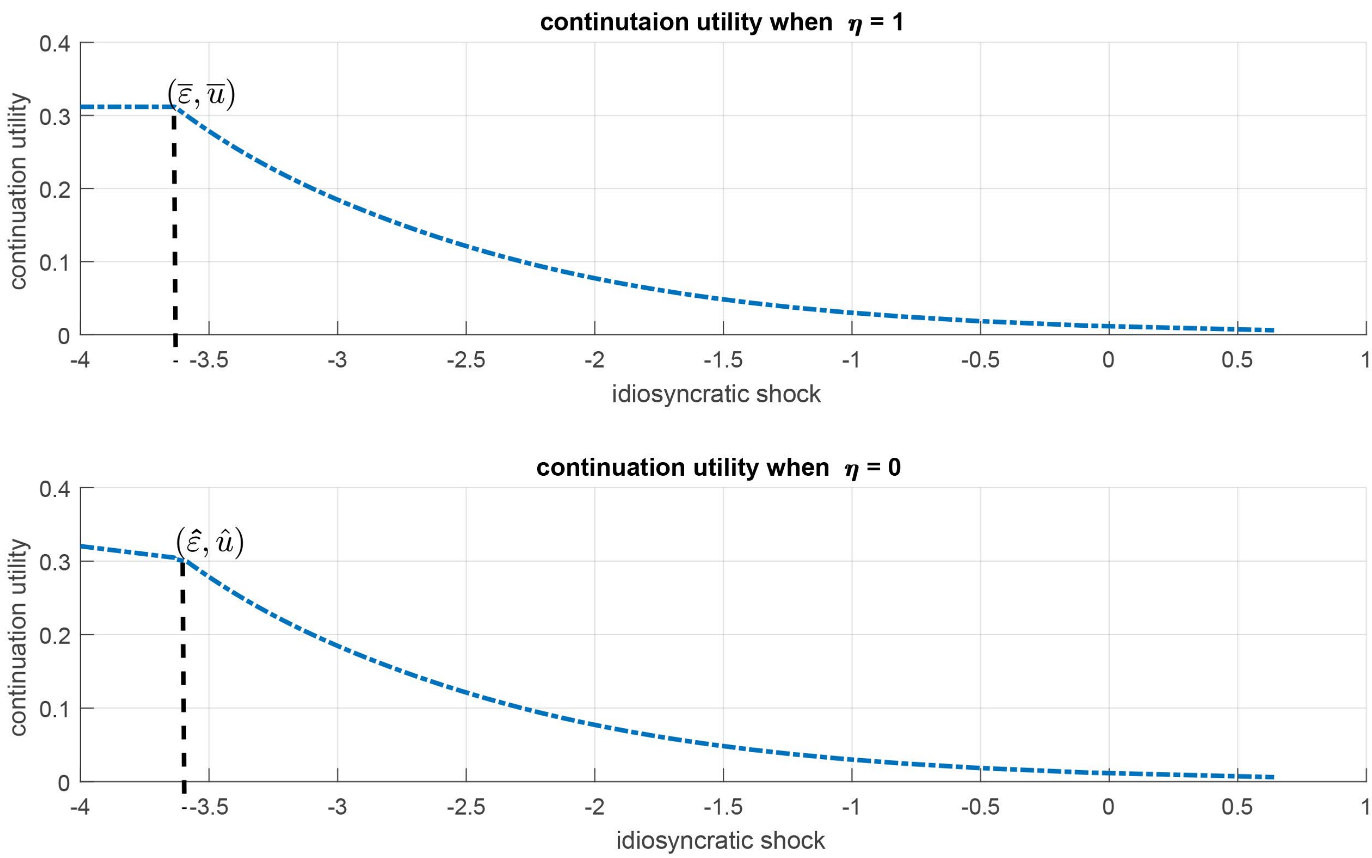


Figure 2: the effect of investment on risk sharing. This figure plots the normalized continuation utility u (y-axis) as a function of idiosyncratic shock (x-axis). Holding all else equal, we compare the cutoff shocks below which firm-side constraint 21 binds in the state where the owner can sell capital on the market, under a low value of investment (upper panel) and a high value of investment (lower panel). An increase in investment relaxes the binding firm-side constraint and worker is less likely ($\hat{\varepsilon}_2 < \hat{\varepsilon}_1$) to be constrained by firm-side constraint with a higher investment.

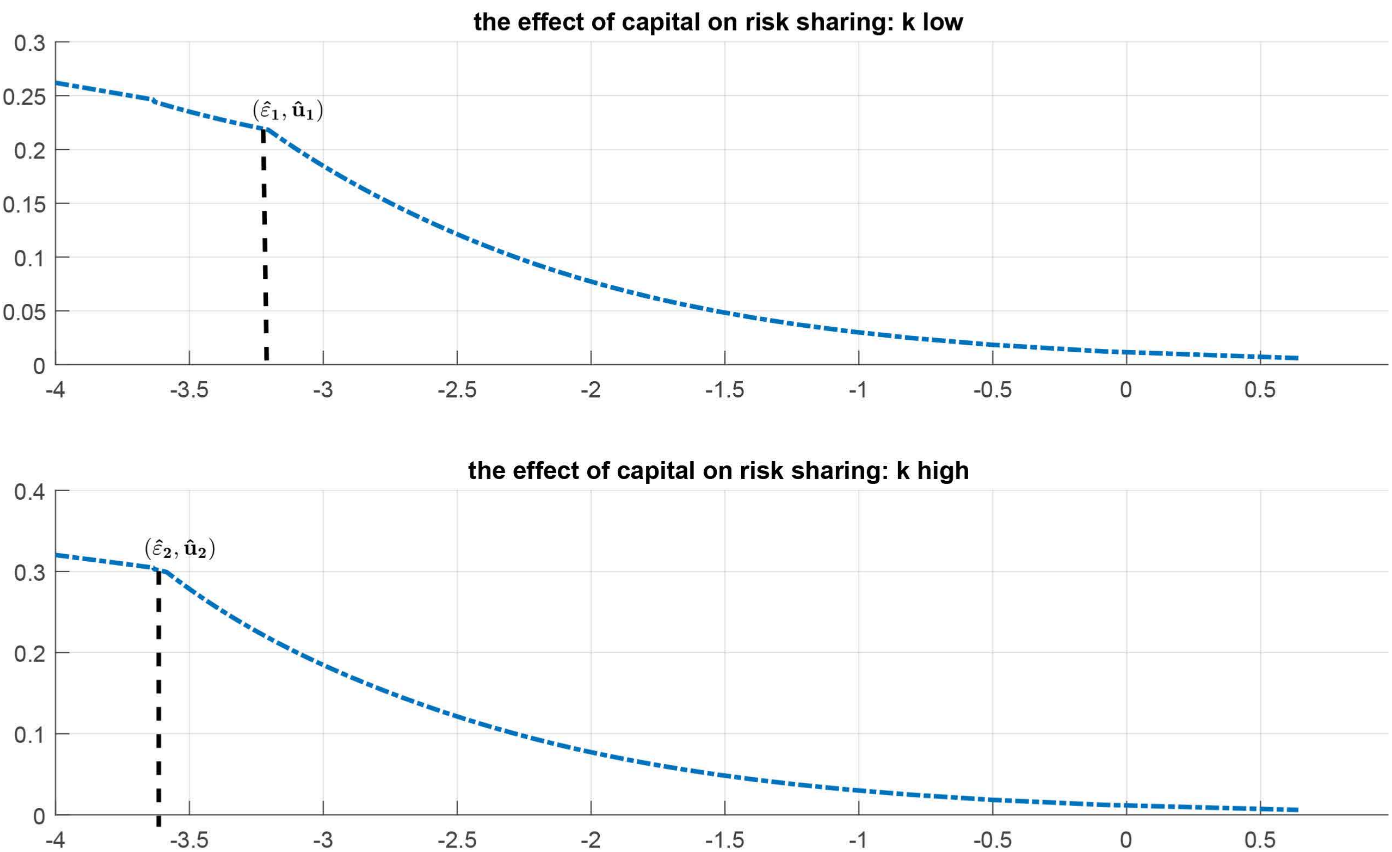
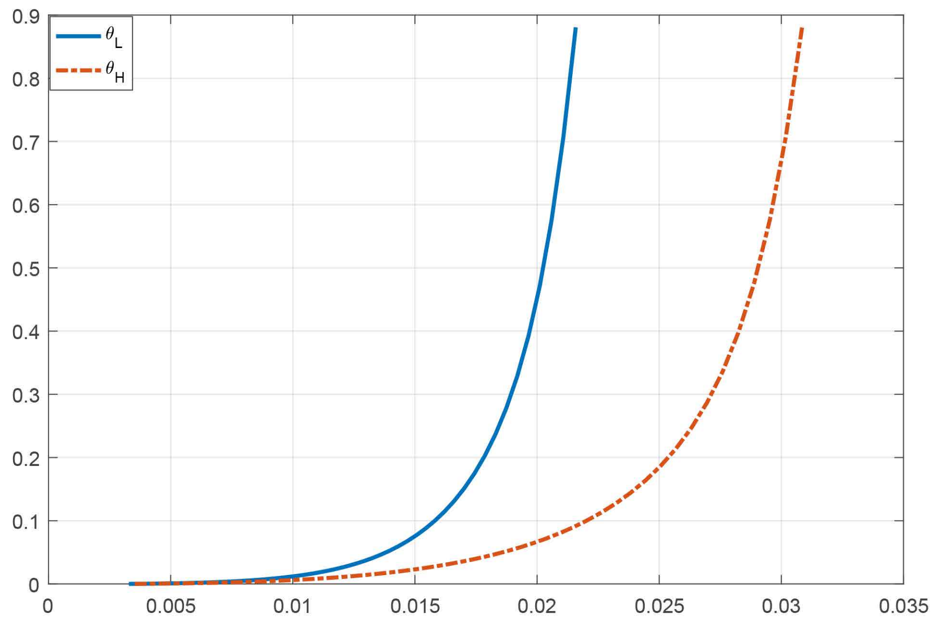


Figure 3: (a) normalized compensation w (y-axis) as a function of u (x-axis) . (b) normalized firm value $\tilde{v}$ (y-axis) as a function of u (x-axis).

(a) firm-level labor share



(b) normalized value function

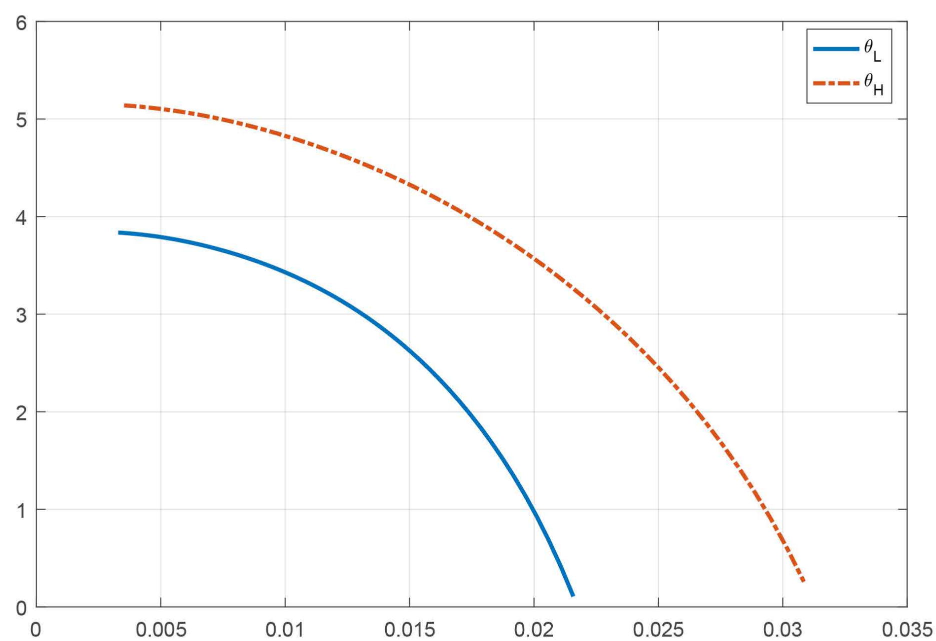
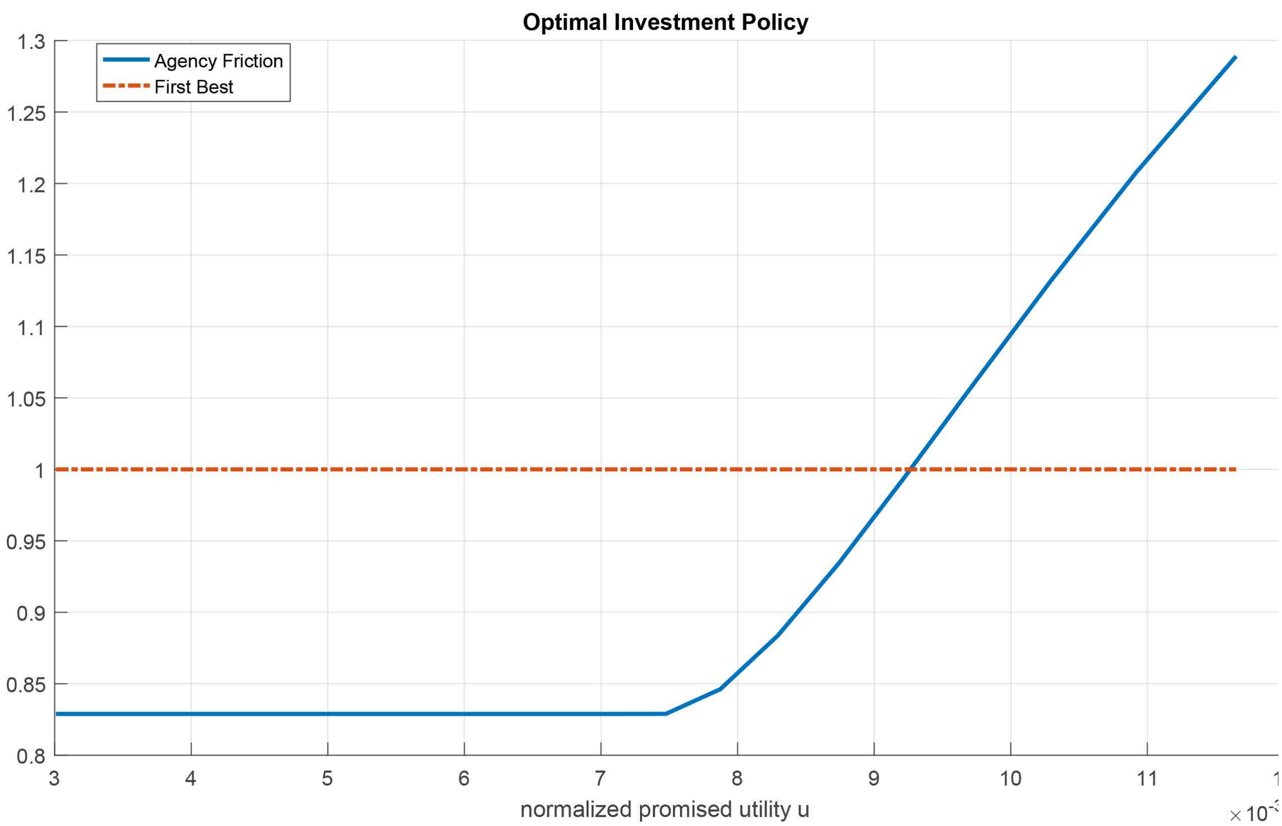


Figure 4: policy function for investment. This figure plots normalized firm investment (y-axis) as a function of normalized continuation utility u (x-axis) with and without agency frictions.



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.

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.

Christopher W. Anderson and Luis Garcia-Feijo. Empirical evidence on capital investment, growth options, and security returns. *The Journal of Finance*, 61(1):171–194, 2006.

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 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. ISSN 08939454, 14657368. URL http://www.jstor.org/stable/2646048.

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.

Frederico Belo, Vito Gala, Juliana Salomao, and Maria Ana Vitorino. Decomposing firm value. *Working Paper*, 2018.

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.

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.

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 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.

Riccardo Colacito, Eric Ghysels, Jinghan Meng, and Wasin Siwasarit. Skewness in expected macro fundamentals and the predictability of equity returns: Evidence and theory. *The Review of Financial Studies*, 2016.

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.

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.

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.

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

Timothy Erickson and Toni Whited. Measurement error and the relationship between investment and q. *Journal of Political Economy*, 108(5):1027–1057, 2000.

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

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.

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

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

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. *The Review of Economic Studies*, 49(3):315–333, 1982.

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

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. ISSN 00221082, 15406261. URL http://www.jstor.org/stable/2329301.

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.

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*, forthcoming.

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.

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

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.

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 nonstockholders. *Journal of Financial Economics*, 29:97–112, 1991.

Monika Merz and Eran Yashiv. Labor and the market value of the firm. *The American Economic Review*, 97(4):1419–1431, 2007.

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.

Ryan H. Peters and Lucian A. Taylor. Intangible capital and the investment-q relation. *Journal of Financial Economics*, 123(2):251 – 272, 2017. ISSN 0304-405X.

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

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

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

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:

$$Ik_t = \frac{I_t}{I_{t-1}} \frac{IK_{t-1}}{1-\delta+IK_{t-1}} \tag{34}$$

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 34. 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 estimiates 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 firms 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.

B. Proof of propositions

C. Numerical solution

We now discuss how we solve the model numerically. Note that the optimal contracting problem 15 is not stationary and we must first detrend it and rewrite it in terms of stationary quantities. We rewrite the normalized problem 19 here for explaining our numerical solution more clearly.

$$\tilde{v}(u|\mathbf{c},\theta) = \max_{u',k,w} \Bigg\{ \theta k^{\alpha} - w - \mathbf{p}(\mathbf{c},\theta)k + \mathbf{E}\Big[\Lambda'(\mathbf{c}'|\mathbf{c},\theta)(1-\delta+\mathbf{i}(\mathbf{c},\theta))e^{\varepsilon'}\tilde{v}(u'|\mathbf{c}',\theta')\Big] \\ + k(1-\delta)\mathbf{E}\Big[\Lambda'(\mathbf{c}'|\mathbf{c},\theta)\mathbf{p}(\mathbf{c}',\theta')\Big] \Bigg\} \tag{35}$$

$$s.t. \quad u = \Big\{(1-\beta)w^{1-\frac{1}{\psi}} + \beta\mathcal{R}(u')^{1-\frac{1}{\psi}}\Big\}^{\frac{1}{1-\frac{1}{\psi}}} \tag{36}$$

$$e^{\varepsilon'}\tilde{v}(u'|\mathbf{c}',\theta') + \frac{\mathbf{p}(\mathbf{c}',\theta')(1-\delta)}{1-\delta+\mathbf{i}(\mathbf{c},\theta)}k \geq 0 \quad \forall(\varepsilon',\theta'), \quad \text{if } \eta = 0 \tag{37}$$

$$e^{\varepsilon'}\tilde{v}(u'|\mathbf{c}',\theta') \geq 0 \quad \forall(\varepsilon',\theta'), \quad \text{if } \eta = 1 \tag{38}$$

$$u'(u,\varepsilon'|\mathbf{c}',\theta') \geq \underline{u}(\mathbf{c}',\theta')\forall(\varepsilon',\theta') \tag{39}$$

Note that we separate the two cases on whether owner can sell capital into two conditions 38 and 37.

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 θ and owner's consumption share, we solve for the normalized 35. Secondly, we introduce an algorithm based on Krusell and Smith (1997) to simultaneously look for equilibrium prices that satisfies market clearing conditions 23 and 24.

Solving the optimal contracting problem 35

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

$$\log \mathbf{c}' = \alpha(\theta,\theta') + \beta(\theta,\theta')\log \mathbf{c} \tag{40}$$

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

$$\log \mathbf{p} = a(\theta) + b(\theta)\log \mathbf{c} \tag{41}$$

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:

$$\begin{aligned}\Lambda(\theta'|\mathbf{c}, \theta) &= \beta(1-\delta + \mathbf{i}(\mathbf{c}, \theta))^{-\gamma}\Big\{\frac{\mathbf{c}'(\theta', \theta, \mathbf{c})}{\mathbf{c}}\Big\}^{-\frac{1}{\psi}}\Big\{\frac{v'_O}{\mathcal{R}v'_O}\Big\}^{\frac{1}{\psi}-\gamma} \\ 1 + h\mathbf{i}(\mathbf{c}, \theta) &= \mathbf{E}\Big[\Lambda(\theta'|\mathbf{c}, \theta)\mathbf{p}(\mathbf{c}', \theta')\Big]\end{aligned} \tag{42}$$

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 5.

3. We adopt value function iteration algorithm and iterate on the Bellman operator 35. 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 [15] 37, $\lambda'_{\eta'=0}(\varepsilon', u, \mathbf{c}, \theta)$ 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:

$$\begin{aligned}&\Lambda(\theta'|\mathbf{c},\theta) = \\ &\qquad \beta e^{-\gamma\varepsilon'}(1-\delta + \mathbf{i}(\mathbf{c}, \theta))^{-\gamma}\Big\{\frac{w(u'_{\eta'=1}(u, \varepsilon', \mathbf{c}', \theta'), \mathbf{c}', \theta')}{w}\Big\}^{-\frac{1}{\psi}}\Big\{\frac{u'_{\eta'=1}(u, \varepsilon', \mathbf{c}', \theta')}{\mathcal{R}u'}\Big\}^{\frac{1}{\psi}-\gamma}\end{aligned} \tag{43}$$

[15] Note that we can also compute the Lagrangian Multipliers associated with binding firm-side constraint 38 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.

$$\Lambda(\theta'|\mathbf{c},\theta) = \beta e^{-\gamma\varepsilon'}(1-\delta+\mathbf{i}(\mathbf{c},\theta))^{-\gamma}\left\{\frac{w(u'_{\eta'=0}(u,\varepsilon',\mathbf{c}',\theta'),\mathbf{c}',\theta')}{w}\right\}^{-\frac{1}{\psi}}\left\{\frac{u'_{\eta'=0}(u,\varepsilon',\mathbf{c}',\theta')}{\mathcal{R}u'}\right\}^{\frac{1}{\psi}-\gamma} \tag{44}$$

(b) **continuation utility provision when 39 binds**: when the worker-side constraint binds, we have the following risk sharing conditions for both scenarios $\eta' = 1$ and $\eta' = 0$

$$\Lambda(\theta'|\mathbf{c},\theta) \geq \beta e^{-\gamma\varepsilon'}(1-\delta+\mathbf{i}(\mathbf{c},\theta))^{-\gamma}\left\{\frac{w(\underline{u}(\mathbf{c}',\theta'))}{w}\right\}^{-\frac{1}{\psi}}\left\{\frac{\underline{u}(\mathbf{c}',\theta')}{\mathcal{R}u'}\right\}^{\frac{1}{\psi}-\gamma} \tag{45}$$

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 38 binds**

$$\Lambda(\theta'|\mathbf{c},\theta) \geq \beta e^{-\gamma\varepsilon'}(1-\delta+\mathbf{i}(\mathbf{c},\theta))^{-\gamma}\left\{\frac{w(\overline{u}(\mathbf{c}',\theta'),\mathbf{c}',\theta')}{w}\right\}^{-\frac{1}{\psi}}\left\{\frac{\overline{u}(\mathbf{c}',\theta')}{\mathcal{R}u'}\right\}^{\frac{1}{\psi}-\gamma} \tag{46}$$

$\overline{u}(\mathbf{c}',\theta')$ is the level of continuation utility such that the value function attains 0 as in constraint 38. We ignore the Lagrangian Multiplier on the binding firm side constraint 38 because it does not interact with other equilibrium conditions.

(d) **continuation utility when 37 binds**

$$\Lambda(\theta'|\mathbf{c},\theta)(1+\lambda'_{\eta'=0}(u,\varepsilon',\mathbf{c}',\theta')) = \beta e^{-\gamma\varepsilon'}(1-\delta+\mathbf{i}(\mathbf{c},\theta))^{-\gamma}\left\{\frac{w(u'_{\eta'=0}(u,\varepsilon',\mathbf{c}',\theta'),\mathbf{c}',\theta')}{w}\right\}^{-\frac{1}{\psi}}\left\{\frac{u'_{\eta'=0}(u,\varepsilon',\mathbf{c}',\theta')}{\mathcal{R}u'}\right\}^{\frac{1}{\psi}-\gamma} \tag{47}$$

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 37 introduces a new marginal benefit of investment which is present in equation 48.

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

$$\begin{aligned}\mathbf{p}(\mathbf{c},\theta) - \mathbf{E}\Big[\Lambda(\theta'|\mathbf{c},\theta)\mathbf{p}(\mathbf{c}',\theta')\Big] =& \alpha\theta k(u,\mathbf{c},\theta)^{\alpha-1} \\ &+ (1-\delta)\mathbf{E}\Big[\Lambda(\theta'|\theta,\mathbf{c})\lambda'_{\eta'=0}(u,\varepsilon',\mathbf{c}',\theta')\mathbf{p}(\mathbf{c}',\theta')\Big]\end{aligned} \tag{48}$$

- 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 38 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 $\{\bar{\varepsilon}_j^{\eta'=1}(\theta_L|\mathbf{c},\theta)\}_{j=1,\dots,nE}, \forall\theta \in \{\theta_L,\theta_H\}, \forall\mathbf{c}$ where nE denotes the total number of grid points. That is, for each current aggregate shock θ and each point on the discretized space of $\mathbf{c}$, $\bar{\varepsilon}_j^{\eta'=1}(\theta_L|\mathbf{c},\theta)$ satisfies 46 with a zero Lagrangian Multiplier, that is $\forall\theta \in \{\theta_L,\theta_H\}$

$$\Lambda(\theta_L|\mathbf{c},\theta) = \beta e^{-\gamma\bar{\varepsilon}_j^{\eta'=1}(\theta_L|\mathbf{c},\theta)}(1-\delta+\mathbf{i}(\mathbf{c},\theta))^{-\gamma}\left\{\frac{w(\overline{u}(\mathbf{c}',\theta_L),\mathbf{c}',\theta_L)}{w}\right\}^{-\frac{1}{\psi}}\left\{\frac{\overline{u}(\mathbf{c}',\theta_L)}{\mathcal{R}u'}\right\}^{\frac{1}{\psi}-\gamma} \tag{49}$$

  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.

$$u'_{\eta'=1}(\varepsilon', u_j, \mathbf{c}', \theta_L) \begin{cases} = \overline{u}(\mathbf{c}', \theta_L), \quad \text{if } \varepsilon' \leq \overline{\varepsilon}_j^{\eta'=1}(\theta_L|\mathbf{c}, \theta) \\ \text{is determined by equation 51}, \quad \text{if } \varepsilon' \in (\overline{\varepsilon}_j^{\eta'=1}(\theta_L|\mathbf{c}, \theta), \overline{\varepsilon}_j^{\eta'=1,A}(\theta_L|\mathbf{c}, \theta)) \\ = \underline{u}(\mathbf{c}', \theta_L), \quad \text{if } \varepsilon' \geq \overline{\varepsilon}_j^{\eta'=1,A}(\theta_L|\mathbf{c}, \theta) \end{cases} \tag{50}$$

$$w(\overline{u}(\mathbf{c}', \theta_L), \mathbf{c}', \theta_L)^{-\frac{1}{\psi}} \overline{u}(\mathbf{c}', \theta_L)^{\frac{1}{\psi}-\gamma} e^{-\gamma(\overline{\varepsilon}_j^{\eta'=1}(\theta_L|\mathbf{c},\theta)-\varepsilon')} \\ = w(u'_{\eta'=1}(\varepsilon', u_j, \mathbf{c}', \theta_L), \mathbf{c}', \theta_L)^{-\frac{1}{\psi}} u'_{\eta'=1}(\varepsilon', u_j, \mathbf{c}', \theta_L)^{\frac{1}{\psi}-\gamma} \tag{51}$$

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 39 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:

$$\overline{\varepsilon}_j^{\eta'=1,A}(\theta_L|\mathbf{c}, \theta) = \overline{\varepsilon}_j^{\eta'=1}(\theta_L|\mathbf{c}, \theta) + \frac{1}{\gamma\psi} \log\left(\frac{w(\overline{u}(\mathbf{c}', \theta_L), \mathbf{c}', \theta_L)}{w(\underline{u}(\mathbf{c}', \theta_L), \mathbf{c}', \theta_L)}\right) + (1 - \frac{1}{\gamma\psi}) \log\left(\frac{\overline{u}(\mathbf{c}', \theta_L)}{\underline{u}(\mathbf{c}', \theta_L)}\right) \tag{52}$$

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 θ to θ_L and that from θ to θ_L using the definition of cutoff shocks 49, on the j^{th} point, that is

$$\overline{\varepsilon}_j^{\eta'=1}(\theta_H|\mathbf{c}, \theta) = \overline{\varepsilon}_j^{\eta'=1}(\theta_L|\mathbf{c}, \theta) + \frac{1}{\gamma} \log(\frac{\Lambda(\theta_L|\mathbf{c}, \theta)}{\Lambda((\theta_H|\mathbf{c}, \theta))}) + \frac{1}{\gamma\psi} \log\left(\frac{w(\overline{u}(\mathbf{c}', \theta_L), \mathbf{c}', \theta_L)}{w(\overline{u}(\mathbf{c}', \theta_H), \mathbf{c}', \theta_H)}\right) \\ + \left(1 - \frac{1}{\psi}\right) \log\left(\frac{\overline{u}(\mathbf{c}', \theta_L)}{\overline{u}(\mathbf{c}', \theta_H)}\right) \tag{53}$$

Using the same step as in 50, 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 38, $\overline{\varepsilon}_j^{\eta'=1}(\theta'|\mathbf{c},\theta)$, 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 37 binds with a zero Lagrangian Multiplier. However, the utility, cutoff shock and Lagrangian Multipliers for binding constraint 37 also interferes with the optimal investment decisions. Therefore, we jointly solve for $\Big\{k(u_j,\mathbf{c},\theta), \hat{\varepsilon}_j^{\eta'=0}(\theta'|\mathbf{c},\theta), \hat{u}_j^{\eta'=0}(\theta'|\mathbf{c},\theta), u'_{\eta'=0}(u_j,\varepsilon',\mathbf{c}',\theta'), \lambda'_{\eta'=0}(u_j,\varepsilon',\mathbf{c}',\theta'), \forall\varepsilon' \leq \hat{\varepsilon}_j^{\eta'=0}(\theta'|\mathbf{c},\theta)\Big\}$ 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 48 that depends on all five quantities, the binding firm-side constraint 37 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 37 for $\Big\{u'_{\eta'=0}(u_j,\varepsilon',\mathbf{c}',\theta'), k(u_j,\mathbf{c},\theta), \hat{\varepsilon}_j^{\eta'=0}(\theta'|\mathbf{c},\theta), \forall\varepsilon' \leq \hat{\varepsilon}_j^{\eta'=0}(\theta'|\mathbf{c},\theta)\Big\}$, equation 54 for $\Big\{\lambda'_{\eta'=0}(u_j,\varepsilon',\mathbf{c}',\theta'), u'_{\eta'=0}(u_j,\varepsilon',\mathbf{c}',\theta'), \forall\varepsilon' \leq \hat{\varepsilon}_j^{\eta'=0}(\theta'|\mathbf{c},\theta)\Big\}$, equation 55 for $\Big\{\hat{\varepsilon}_j^{\eta'=0}(\theta'|\mathbf{c},\theta), \hat{u}_j^{\eta'=0}(\theta'|\mathbf{c},\theta)\Big\}$.[16]

$$\lambda'_{\eta'=0}(u_j,\varepsilon',\mathbf{c}',\theta') = e^{-\gamma\overline{\varepsilon}_j^{\eta'=1}(\theta'|\mathbf{c},\theta)+\gamma\varepsilon'}\left\{\frac{w(\overline{u}(\mathbf{c}',\theta'),\mathbf{c}',\theta')}{w(u'_{\eta'=0}(u_j,\varepsilon',\mathbf{c}',\theta'),c',\theta')}\right\}^{-\frac{1}{\psi}}\left\{\frac{\overline{u}(\mathbf{c}',\theta')}{u'_{\eta'=0}(u_j,\varepsilon',\mathbf{c}',\theta')}\right\}^{\frac{1}{\psi}-\gamma} - 1 \tag{54}$$

$$\hat{\varepsilon}_j^{\eta'=0}(\theta'|\mathbf{c},\theta) = \overline{\varepsilon}_j^{\eta'=1}(\theta'|\mathbf{c},\theta) + \frac{1}{\gamma\psi}\log\left(\frac{w(\overline{u}(\mathbf{c}',\theta'),\mathbf{c}',\theta')}{w(\hat{u}_j^{\eta'=0}(\theta'|\mathbf{c},\theta),\mathbf{c}',\theta')}\right) + (1-\frac{1}{\gamma\psi})\log\left(\frac{\overline{u}(\mathbf{c}',\theta')}{\hat{u}_j^{\eta'=0}(\theta'|\mathbf{c},\theta')}\right) \tag{55}$$

Equation 54 and 55 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 35 is to solve for this system of nonlinear equations described above. We use various of

[16] 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.

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 48.

- with all policy functions obtained in previous steps, we use the definition of cutoff shock 49 to update the policy function for labor compensation w on each set of grid points. We then use the promise keeping constraint 36 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 35.

$$\begin{aligned}\tilde{v}^{\mathcal{CMK}}(u|\mathbf{c},\theta,p) = \max_{u',k,w}\Big\{&\theta k^{\alpha} - w - pk + \mathbf{E}\Big[\Lambda'(\mathbf{c}'|\mathbf{c},\theta)(1-\delta+\mathbf{i}(\mathbf{c},\theta))e^{\varepsilon'}\tilde{v}(u'|\mathbf{c}',\theta')\Big] \\ &+ k(1-\delta)\mathbf{E}\Big[\Lambda'(\mathbf{c}'|\mathbf{c},\theta)\mathbf{p}(\mathbf{c}',\theta')\Big]\Big\}\end{aligned} \tag{56}$$

$$s.t. \quad 36, 37, 38, 39$$

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 41. Also the law of motion for owner' consumption share is perceived as equation 40. Since price of capital deviates from the assumed price function 41 for only one period, the continuation value of the firm coincides with the value function in problem 35.

This problem will generate investment decisions, $k^{\mathcal{CMK}}(u, \mathbf{c}, \theta|p)$ and continuation utilities (for both realizations of η' and we abstract from the subscript that represents the state of η' 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 41 and law of motion 40.

1. solve the optimal contract problem 35 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 θ_t, θ_{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 56 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 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 23.

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

(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[17].

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 41 and 40. 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.

[17]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.