Delegated Investment, Q-Theory, and Firm Dynamics

May 31, 2011

Hengjie Ai and Rui Li[1]

Abstract

We present a dynamic general equilibrium model with heterogeneous firms. Owners of the firms delegate investment decisions to managers, whose consumption and investment are private information. We solve the optimal incentive compatible contracts and characterize the implied firm dynamics. Optimal risk sharing requires managers' equity share to decrease with firm size. This in turn implies that it is harder to prevent private benefit in larger firms, where managers have lower equity stake under the optimal contract. Consequently, smaller firms invest more, pay less dividends, and grow faster. We calibrate our model and evaluate the quantitative importance of the above mechanism in accounting for the empirical relationship between firm size and firms' investment and payout policies.

Keywords: Dynamic Contracting, General Equilibrium, Firm Dynamics, Q-Theory, Moral Hazard, Continuous-time Method.

JEL Code: C73, D51, D82, D92, E13, E22, G35.

[1]Ai (hengjie@duke.edu) is affiliated with the Fuqua School of Business at Duke University; Li (rli2@wisc.edu) is at the Department of Economics at University of Wisconsin-Madison. We thank Ravi Bansal, Adriano Rampini, Vish Viswanathan, and Noah Williams for useful conversations on this topic. All errors are our own responsibility. The most current version is available via http://faculty.fuqua.duke.edu/~hengjie/homepage/research/AL2011A.pdf.

Introduction

In this paper, we present a general equilibrium model with heterogeneous firms. Each firm in our model operates a neoclassical production technology that combines capital and labor to produce consumption goods. We also allow capital accumulation to be subject to adjustment cost as in Hayashi [1982]. We deviate from the neoclassical setting by incorporating separation of ownership and control in our framework. Shareholders do not have the technology to invest in the firms they own and have to delegate these decisions to managers. In addition, managers' consumption and investment decisions are private information. Capital accumulation is determined by not only managers' investment decisions but also unobservable and observable idiosyncratic shocks. The size of the total capital stock of firms therefore serves as a noised signal of managers' past actions, and predicts firms' future investment and payout policies under the optimal contract. We assume that both shareholders and managers have recursive utility with constant relative risk aversion and constant intertemporal elasticity of substitution (IES) and derive the optimal dynamic incentive contract. We close the model in general equilibrium and characterize the cross-section distribution of firms' characteristics. Our model thus provide a unified framework to study firms' investment and payout policies, the dynamics of firm size and firm growth, and production based asset pricing in general equilibrium setting.[2]

Our focus is to understand how the presence of moral hazard affects firms' investment and payout policies at the micro level and their implications on firm dynamics and aggregate quantities at the macro level. We find both unobservable and observable shocks to be important in understanding some of the well-documented empirical relationship between firm size and firms' investment and payout policies. Qualitatively, the presence of unobservable shocks allows our model to deviate from the neoclassical framework and its predictions on the relationship between firms' investment and valuation ratios (average Q). It also generates statistical predictability of firms' investment and payout decisions by their sizes. Quantitatively, we find that the predictions of our model are consistent with empirical evidence only when observable shocks account for a large fraction of idiosyncratic shocks at the firm level.

We first present a theoretical characterization of the optimal response of managerial compensation with respect to observable shocks. We show that the marginal cost of utility provision must not comove with observable shocks. To understand the intuition

[2] Ai and Li [2011], extends our model to a setting with aggregate uncertainty and long-run risks and study its implications on the cross-section of stock returns.

of the above result, note that the design of the continuation contract can be viewed as a process of assigning continuation utilities across time and states of nature. Optimality requires that the marginal cost of utility assignment is equalized across states of nature if utility is additively separable.[3] This property of the optimal contract generalizes to the case of recursive utility as long as an appropriate normalization of the utility function is chosen (so that aggregation of continuation utility is additive across states).

We next show that when observable shocks are large compared with unobservable shocks, the empirical predictions of our model are consistent with several stylized facts on the relationship between the size of firms and their investment and payout policies. Because managers are risk averse, the optimal incentive contract must trade off risk sharing against incentive provision. Risk sharing implies that the manager's equity share in the firm is a decreasing function of firms' performance, and therefore firm size. As managers own only a small equity stake in large firms, it is harder to provide incentives to invest. Consequently, smaller firms invest more, grow faster, and have higher pay-performance sensitivity.

Our model also offers a potential explanation of the lack of statistical power of average Q (the ratio of a firm's market value over the replacement cost of its total capital stock) in forecasting firms' investment in empirical works. In our model, in steady state, most of the firms reside in the region where average Q is a decreasing function of the promised utility to the manager. An increase in average Q lowers the continuation utility of the manager and has two offsetting effects. First, lower continuation utility is implemented by higher levels of investment and lower levels of consumption for the manager. This generates a positive relationship between average Q and investment. Second, lower continuation utility is also associated with a smaller equity share of the manager, which implies that it is harder to provide incentives for the manager to invest. We show that together they imply a non-monotonic relationship between average Q and investment. In our calibration we show average Q has little predictive power for investment, while cash flow does.

Q-theory still holds in our model, in the sense that the optimal contract requires the marginal cost of investment to equal its marginal benefit. In fact, profit maximization implies that this condition has to hold from the perspective of the shareholders, and incentive compatibility requires that the same holds for the managers. Average Q does not have strong predictive power for investment in our model because the relationship between average Q and marginal Q is non-monotonic.

[3]The same principal, when applied to utility assignment over time, leads to the famous "Inverse Euler Equation" of Rogerson [1985], Spear and Srivastava [1987] and Golosov et al. [2003].

We cast our model in continuous time and obtain simple characterizations of the optimal incentive contract and the implied firm dynamics. Like previous literature, for example, Sannikov [2008] and DeMarzo and Sannikov [2006], we find continuous time methods offer a convenient way to study the optimal contract. We extend their methodology to allow for recursive utility. We show that the optimal contract can be characterized as the solution to an ordinary differential equation with known boundary conditions. An additional advantage of our continuous time framework is that it allows a simple characterization of the cross-section distribution of firms' characteristics. By exploring the homogeneity property of firms' decision rules, we show that the two-dimensional distribution of firms can be summarized by a one-dimensional measure defined on the space of scaled continuation utilities of the agent. This measure obeys a version of the Kolmogorov forward equation whose solution can be easily calculated numerically.

The microeconomic foundation of our model is related to a large literature on firm dynamics with principal-agent problem. Earlier work includes Albuquerque and Hopenhayn [2004], Clementi and Hopenhayn [2006], Quadrini [2004]. Their main focus is to investigate investment behavior inside the firm under optimal long-term contracts in various contracting environments. The above papers typically abstract from capital accumulation on the firm side. Recently, DeMarzo et al. [2009],Biais et al. [2010], Philippon and Sannikov [2007], Clementi et al. [2008] and He [2009] incorporate firm growth and firm size dynamics in dynamic agency models. In DeMarzo et al. [2009],Biais et al. [2010], Philippon and Sannikov [2007], Clementi et al. [2008], the investment decision is made by the principal and is publicly observable. The agent's action does not directly affect firm growth. He [2009] presents a model in which the manager's hidden action affects the expected growth rate of the firm. As in DeMarzo et al. [2009],Biais et al. [2010], Philippon and Sannikov [2007], He [2009]'s model also focuses on the case of risk neutral agent. None of the above mentioned papers considers delegated investment problem with neoclassical production technology as we do.

In addition, this paper differs from the above literature in two significant ways. First, our model is general equilibrium with neoclassical production technologies. The general equilibrium framework allows us to tie some of the important assumptions in the above literature to the structural parameters of preferences and technologies in our framework. For example, partial equilibrium models typically specify an exogenous liquidation value of firms, and a difference between the discount rate of the principal and the agents. In our framework, the liquidation value of firms is determined endogenously by the negativity constraint of dividend payment. The difference between the discount rate of the principal and the agent is motivated by precautionary saving motives of the latter. The production technology in our model is governed by a few parameters, capital share, depreciation rate, volatility of firm level output, and total factor productivity at the macro level. This allows us to draw on the research in a large body of the macroeconomics literature to discipline our choice of these parameters in calibration. It also allows us to confront the implications of the model with both empirical evidence at the macro level, as well as those at the firm level. For example, our model relates the equilibrium distribution of firms' characteristics to the dynamics of individual firms' performance. Because the distribution of firms characteristics depends on the parameters of firms' technology, this provides additional discipline in choosing the structural parameters of model from micro-evidence in quantitative exercises.

The second distinguishing feature of our model is that managers' preferences are represented by recursive utility that separates risk aversion and intertemporal elasticity of substitution (IES). The decision theoretic literature, for example, Kreps and Porteus [1978], Epstein and Zin [1989], among others, has established convincingly that they are distinctive notions of economic behavior. The macroeconomics and asset pricing literature emphasizes that they are responsible for different aspects of the quantitative implications of general equilibrium models on macroeconomic quantities and asset prices, for example, Tallarini [2000], Bansal and Yaron [2004]. The generality of our framework allows us to calibrate the preference parameters of managers to the relevant micro-evidence in quantitative exercises. It also makes possible for more realistic economic settings to be studied in our model.

(Modified from this paragraph on)We demonstrate that the design of the optimal contract is quantitatively mostly affected by IES but not risk aversion. We find firm value to be monotonically decreasing in the manager's IES, holding risk aversion fixed. The presence of moral hazard lowers firm level investment. The extent to which investment is affected is mainly determined by the IES parameter of the manager. The intuition for this echoes a similar result obtained by Tallarini [2000] in a neoclassical model with a representative firm. The neoclassical investment technology is essentially a technology that trades off consumption over time, but not across states of nature. In fact, as pointed out by Cochrane [1993], the neoclassical technology is Leotiff across states. Consequently, investment in these models is mostly determined by the agent's attitude toward intertemporal substitution of consumption, not risk aversion. On the other hand, we find that the effect of risk aversion on firms' value is ambiguous. An increase in managers' risk aversion has two offsetting effects on firm value. First, it amplifies the utility loss due to incentive provision. Second, it also raises the agent's incentive to invest through precautionary saving motives, which implies a better alignment of principal and agent's interest. Depending on parameter values, either effect could dominate. Overall, quantitatively, the most important determinant of firm value is IES, not risk aversion.

Our paper is also built on the literature on firm dynamics and the size distribution of firms. A robust feature of the data is that the size distribution of firms is stationary and resembles Pareto distribution despite the large amount of heterogeneity at the individual firm level. For an excellent recent survey of this literature, see Luttmer [2011]. Equilibrium models of firm dynamics includes Jovanovic [1982], Hopenhayn [1992], Klette and Kortum [2004], Luttmer [2007], and Arkolakis [2011], among others. None of the above papers study dynamic agency and delegated investment problems as we do.

The relationship between firm size and firms' investment and payout policies has been documented by many researchers. Mansfield [1962] is among the earliest to show that small firms grow faster. Evans [1987a] and Evans [1987b] argue that the size-growth relationship is robust to possible sample selection bias. Fama and French [2001] document that large firms are much more likely to pay dividend than small ones. The size distribution of firms has Pareto-like shape in our model. In addition, our model generates endogenously an inverse relationship between firm size and firm growth based on dynamic agency theory.

Our model relates to the large literature on "Q-theory" of investment. Abel and Eberly [1994] provide an excellent summary of this literature and develop a unified Q theory of investment in neoclassic settings. Recent papers that are most related to ours are DeMarzo et al. [2009], and Bolton et al. [2010]. DeMarzo et al. [2009] is a dynamic agency model where the productivity of firms depends on unobservable effort.Bolton et al. [2010] focus on firms' financing decisions and risk management decisions. None of the above papers consider delegated investment problem in general equilibrium as we do here.

On the empirical side, the lack of predictability of firms' investment by their valuation ratio (average Q) has been documented by many authors. For example, Fazzari et al. [1988], Stein [2003] among others. We show that the relationship between marginal Q and average Q is non-monotonic in our model thus provide a potential explanation for the above phenomena.

Our paper also bears connections with the literature that studies private information in capital accumulation in "real business cycle" models, for example, Bernanke and Gertler [1989], Carlstrom and Fuerst [1997] and Bernanke et al. [1999]. The focus of these papers is to understand how credit market conditions affect business cycle fluctuations, while our purpose is to understand the time series and cross sectional properties of firms' behavior. In addition, these authors use discrete time models and study one-period contracts. We solve the optimal dynamic contract using continuous time method.

The rest of the paper is planned as follows. We introduce the setup of the model in Section I. Section II characterizes the optimal dynamic contract for individual firms. Section III aggregates firms decisions, studies the cross-section distribution of firms' characteristics, and closes the model in general equilibrium. We present our calibration results in Section IV. Section V concludes the paper.

I Setup of the Model

A Preference and Technology

A.1 Preferences of Shareholders and Managers

There are a continuum of shareholders in the economy who can plant trees. A tree is a technology to combine labor and tree-specific capital to produce consumption goods, and to accumulate tree-specific capital over time. The knowledge to grow trees is known only to managers, a special type of agents who arrive at the economy in overlapping generations.[4] Shareholders, therefore have to delegate the investment decisions to the managers. To keep our language consistent with the principal-agent literature, we will use the terminology shareholders and principals, managers and agents interchangeably.

Both shareholders and managers have recursive preferences of the Kreps and Porteus [1978] type. In continuous time, they are represented by stochastic differential utility with constant relative risk aversion and constant IES. Shareholders and managers have the same time discount rate, risk aversion and IES, which are denoted as β, γ and ψ, respectively.[5] The shareholders are also endowed with one unit of labor which is supplied inelastically.

A.2 Production and Investment Technology

At any point of time, shareholders can plant a unit measure of trees. Trees are indexed by j and each tree is associated with a certain amount of tree-specific capital. We use $K_{j,t}$ to denote the amount of capital of tree j at time t. A tree combines capital and

[4] The overlapping generation aspect of our model follows the continuous time "perpetual youth" model of Blanchard [1985].

[5] Because we focus on the steady state where there is no economic growth or aggregate uncertainty, the specification of the IES and risk aversion parameters of the shareholders is only relevant for welfare comparisons, but has no effect on equilibrium asset prices.

labor to produce consumption goods via the standard Cobb-Douglas technology:

$$Y_{j,t} = z{K_{j,t}}^{\alpha}{N_{j,t}}^{1-\alpha},$$

where z is an economy-wide common productivity parameter, and $Y_{j,t}$ and $N_{j,t}$ denote the total output produced and labor employed by tree j at time t. We focus on the case in which z is a constant.

Let W_t denote the real wage at time t, the total operating profit of tree j at time t is

$$P(K_{j,t}) = \max_{N_{j,t}} \left\{ z{K_{j,t}}^{\alpha}{N_{j,t}}^{1-\alpha} - W_t N_{j,t} \right\}. \tag{1}$$

Tree-specific capital can be accumulated according to the following investment technology:

$$dK_{j,t} = K_{j,t}\left[-\delta dt + \sigma^T dB_{j,t}\right] + I_{j,t}dt, \tag{2}$$

where $I_{j,t}$ is the total investment chosen by the manager of the tree. δ is the depreciation rate of capital common among all trees. The term $B_{j,t}$ is a 2×1 vector of standard Brownian motions independent across firms, and

$$\sigma^T dB_j = \sigma_u dB_{u,j} + \sigma_o dB_{o,j}.$$

In the above expression, $\sigma, \sigma_o > 0$ are constants. The Brownian motion $B_{u,j}$ is unobservable to all except the manager who operates the tree, and the Brownian motion $B_{o,j}$ is common knowledge in the economy.

A.3 Information and Managerial Compensation

At any point in time t, given the total amount of capital stock $K_{j,t}$, which is observable to all, the manager makes observable decisions on labor hiring $N_{j,t}$. After the wage bill $W_t N_{j,t}$ is paid, managers hand in the total operating profit, $P(K_{j,t})$ to shareholders. Total operating profit is divided among dividend payment to shareholders, $D_{j,t}$, compensation to the manager, $C_{j,t}$, investment in the firm, and a cost of capital adjustment of the form $\Omega(I_{j,t}, K_{j,t})$:

$$D_{j,t} + C_{j,t} + I_{j,t} + \Omega(I_{j,t}, K_{j,t}) = P(K_{j,t}). \tag{3}$$

Shareholders do not have the technology of transforming the tree-j specific capital stock into consumption goods. Therefore the maximum amount of dividend is $P(K_{j,t})$. We also assume only tree-j specific output can be used for accumulation of tree-j specific capital and the consumption of the manager of the tree. This implies that managerial consumption and investment cannot exceed total operating profit: $C_{j,t} + I_{j,t} + \Omega(I_{j,t}, K_{j,t}) \leq$

$P(K_{j,t})$[6]. In what follows, we will call the sum $C_{j,t} + I_{j,t} + \Omega(I_{j,t}, K_{j,t})$ the *retained earnings* of firm j. The constraints on dividend payment and retained earnings can be summarized as:

$$0 \le D_{j,t} \le P(K_{j,t}). \tag{4}$$

Because $P(K_{j,t})$ is observable, a given level of dividend payout $D_{j,t}$ can be implemented as long as it satisfies the constraint in (4).

The consumption and investment decisions of the manager are private information observable only to managers themselves. Therefore although the principal is free to chose any dividend policy as long as (4) is satisfied, the division of retained earnings between managerial compensation, $C_{j,t}$, and investment in the firm, $I_{j,t}$ is up to the manager. Given $P(K_{j,t}) - D_{j,t}$, the manager is free to choose any consumption and investment that satisfies

$$C_{j,t} + I_{j,t} + \Omega(I_{j,t}, K_{j,t}) = P(K_{j,t}) - D_{j,t}.$$

In particular, it is possible to choose negative levels of investment. In the absence of adjustment cost, this means that the manager can always maintain a positive level of consumption by setting $I_j(t) < 0$, even if total retained earnings is set to zero by the principal.

In general, we assume a quadratic adjustment cost:

$$\Omega(I, K) = \frac{1}{2}\omega\left(\frac{I}{K} - i^*\right)^2, \text{ where } \frac{1}{2i^*} > \omega \ge 0. \tag{5}$$

The specification $\omega = 0$ corresponds to the case with no capital adjustment cost. The assumption $\omega < \frac{1}{2i^*}$ guarantees that the agent can always obtain a positive amount of consumption by setting $I_j(t)$ negative for all possible dividend policies. We assume there is a minimum level of investment-to-capital ratio, that is,

$$\frac{I_{j,t}}{K_{j,t}} \ge -B,$$

for some large real number $B > 0$. This is merely a technical assumption that ensures that managers' optimization problems have compact choice sets and therefore well-defined solutions. In what follows, we will chose B to be large enough so that the above constraint never binds.

[6]Equivalently, our model can be interpreted as each tree j produce a tree-specific intermediate good, which can be used to produce final consumption good for shareholders using a constant return to scale technology (for example, that of the Dixit and Stiglitz [1977] type). Managers have the technology of transforming tree-specific output into tree-specific capital, or his own private consumption, while shareholders do not. To keep notation simple, we do not spell out the details of these specification.

A.4 Entry and Exit of Firms

Managers arrive at the economy continuously. Upon arrival, a manager is endowed with the technology to operate a tree, and an outside option that delivers a reservation utility U_0[7]. A unit measure of managers arrive at the economy at each point in time. The reservation utility of newly arrived managers is distributed according to a continuous density Λ with support $[U_L, U_H]$.

At each point in time, the principal offers a contract to every newly arrived manager that specifies the initial amount of capital of the tree to be delegated to the manager and future payment to the manager as a function of the history the performance of the tree. If the offer is accepted, the principal will plant the tree and delegate the tree to the manager until he exits the economy. Planting a tree with initial amount of capital K_0 costs exactly K_0 amount of consumption good.

We assume that the principal has all the bargaining power. In this case principal always promises utility level U_0 as long as it is profitable to plant the tree. Managers will accept the contract and work for the tree until he exits the economy. Let $K_0(U_0)$ denote the amount of initial capital of the tree delegated to a manager with reservation utility U_0. The total amount of resources used to plan new trees is therefore

$$\int_{[U_L, U_H]} K_0(U)\,\Lambda(U)\,dU.$$

We assume that market is complete, so that shareholders can fully diversify idiosyncratic shocks in dB_t^i and $dB_{o,t}^i$. However, because of the moral hazard problem, the only way that managers can credibly participate in the credit market is though the compensation contracts that tie their consumption to the performance of the tree that he operates.

Tree managers receive health shocks at Poisson rate κ per unit of time. Once hit by the health shock, the manager exit the economy, and the tree along with the tree-specific capital evaporates. Under the above specification, a unit measure of trees enter the economy at each point in time. Given the death rate κ, law of large numbers imply that the total measure of trees in the economy in steady state is $\frac{1}{\kappa}$.

We think of each tree as a firm and use the terminology trees and firms interchangeably from now on. In our setup, each firm is identified by the firm-specific technology and the contractual relationship between the shareholders and the manager of the firm.

[7] The reservation utility can be endogenized by assuming that managers have access to an alternative technology that delivers exactly the life time utility U_0. We do not spell out the details here as in equilibrium, this option is never chosen.

A.5 Linearity of the Operating Profit Function

We first obtain a key simplification of the model: because the production function is constant return to scale, and there is no frictions in the labor market, optimal allocation of labor will equalize the marginal product of capital across firms of all sizes. This the Hayashi [1982]'s result and is summarized in the following proposition.

Proposition 1. (Linearity of the Operating Profit Function) *The operating profit function $P(K)$ is linear. In steady-state, the operating profit function is given by;*

$$P(K) = AK,$$

where the constant A is the marginal product of capital:

$$A = \alpha z \mathbf{K}^{\alpha - 1}, \tag{6}$$

and the bold faced letter $\mathbf{K}$ stands for the total capital stock in the economy in steady state.

Proof. See Appendix. □

It is important to note that there is no moral hazard problem in production decisions, as they are publicly observable. This allows us to apply Hayashi [1982]'s result and obtain linearity of the profit function. As we will show in the rest of the paper, this is a key simplification. Since manager's preferences are homothetic, the above result allows us to reduce the dimensionality of the contracting problem, as well as the description of the cross-section distribution of firms in computing aggregate quantities. We now turn to manager and shareholder's maximization problems.

B Shareholders' and Managers' Optimization Problem

B.1 Feasibility of Plans

At each point in time t, shareholders choose a dividend policy $D_{j,t}$, investment policy $I_{j,t}$ and compensation to the manager, $C_{j,t}$ for each firm j. We use superscript to denote the history of the realizations of a stochastic process up to time t. For example, K_j^t refers to the realizations of the firm-j specific capital from time 0 to time t, $\{K_j(s)\}_{s=0}^t$. The consumption, investment, and dividend payment policies can only be functions of observables, that is, the history, $\left\{K_j^t, B_{j,o}^t\right\}$, denoted by H_j^t. For $s \geq t$, we use the notation $H_j^s \geq H_j^t$ to denote a history that follows H_j^t. For convenience, we suppress the firm subscript j. Formally, a plan is a triple, $\left\{C_t\left(H^t\right), I_t\left(H^t\right), D_t\left(H^t\right)\right\}_{t=0}^{\infty}$ that specifies managerial consumption, investment, and dividend payout as a function of the history, $H^t = \left\{K_j^t, B_{j,o}^t\right\}$.

Given a plan, the manager chooses his consumption and investment policies $\left\{\tilde{C}_t, \tilde{I}_t\right\}_{t=0}^{\infty}$ conditioning on his information, which contains the history of the Brownian motions, B_u^t and B_o^t, as well as that of the manager's own past actions. We define a strategy of the manager to be an consumption and investment policy, $\left\{\tilde{C}_t, \tilde{I}_t\right\}_{t=0}^{\infty}$ adapted to the filtration generated by by the Brownian motions, B_u and B_o. Because dividend payment is observable to the principal, a strategy is feasible given a plan, $\left\{C_t\left(H^t\right), I_t\left(H^t\right), D_t\left(H^t\right)\right\}_{t=0}^{\infty}$, if it is adapted to the manager's information set and satisfies

$$\tilde{C}_t + \tilde{I}_t = P\left(K_t\right) - D_t\left(H^t\right) = C_t\left(H^t\right) + I_t\left(H^t\right), \text{ for all } t, \text{ and } H^t. \tag{7}$$

That is, given a plan, feasibility of manager's strategy requires it be consistent with the dividend policy prescribed by the plan. Since consumption and investment policy are unobservable to the principal, managers can choose alternative strategies other than that prescribed by the plan in order to maximize his utility.

Managers evaluate the utility of a consumption profile according to Kreps and Porteus [1978] preference represented by stochastic differential utility (Epstein and Zin [1989], Duffie and Epstein [1992]). We will use the notation $U_t\left(\left\{\tilde{C}_s\right\}_{s\geq t}^{\infty}\right)$ to denote the date-t continuation utility of a manager associated with the continuation strategy from time t, $\left\{\tilde{C}_s\right\}_{s=t}^{\infty}$. We will use the short-hand notation U_t whenever there is no confusion. In this case, for all $s > t$, U_s satisfies:

$$dU_s = [-f(\tilde{C}_s, U_s) - \frac{1}{2}\mathcal{A}(U_s)||\sigma_{U,s}||^2]ds + \sigma_{U,s}^T dB\left(s\right), \tag{8}$$

where $\sigma_{U,s}$ is a 2×1 vector of the diffusion coefficient on the Brownian motions $dB_{u,s}$ and $dB_{o,s}$. We adopt the convenient normalization with $\mathcal{A}(U) = 0$(Duffie and Epstein [1992]), and denote $\bar{f}$ the normalized aggregate. Under this normalization, $\bar{f}(C, U)$ is:

$$\bar{f}(C,U) = \frac{\beta+\kappa}{1-1/\psi} \frac{C^{1-1/\psi} - \left(\left(1-\gamma\right)U\right)^{\frac{1-1/\psi}{1-\gamma}}}{\left(\left(1-\gamma\right)U\right)^{\frac{1-1/\psi}{1-\gamma}-1}} \tag{9}$$

Here we assume $\gamma, \psi \neq 1$. The extension to $\gamma = 1$ and/or $\psi = 1$ is straightforward. In the above expression, β is the discount rate of the manager and $\beta + \kappa$ is the effective discount rate accounting for death. In this notation, the lifetime discounted utility of the agent is given by:

$$U_t\left(\left\{\tilde{C}_s\right\}_{s=t}^{\infty}\right) = E_t\left[\int_t^{\infty} \bar{f}\left(\tilde{C}_s, U_s\right) ds\right],$$

where the conditional expectation is take with respect to manager's information set. Formally, it is conditioned on the filtration generated by the two dimensional Brownian motion $\{B_u(t), B_o(t)\}_{t=0}^{\infty}$.

Because the consumption and investment decisions are not observable to the principal, given a plan $\{C_t(H^t), I_t(H^t), D_t(H^t)\}_{t=0}^{\infty}$, the manager chooses, $\left\{\tilde{C}_t, \tilde{I}_t\right\}_{t=0}^{\infty}$, from the set of all feasible investment and consumption strategies, to maximizes $U_0\left(\left\{\tilde{C}_t\right\}_{t=0}^{\infty}\right)$. A plan, $\{C_t(H^t), I_t(H^t), D_t(H^t)\}_{t=0}^{\infty}$, is *incentive compatible* if for any H^t, $C_t(H^t)$ and $I_t(H^t)$ are the optimal choice of the manager. Equivalently, a plan is incentive compatible if for any t, and any history $H^s \geq H^t$ that follows H^t,

$$U_t\left(\{C_s(H^s), I_s(H^s)\}_{s=t}^{\infty}\right) \geq U_t\left(\left\{\tilde{C}_s, \tilde{I}_s\right\}_{s=t}^{\infty}\right), \tag{10}$$

for any feasible continuation strategy, $\left\{\tilde{C}_s, \tilde{I}_s\right\}_{s=t}^{\infty}$ that satisfies condition (7). A plan $\{C_t(H^t), I_t(H^t), D_t(H^t)\}_{t=0}^{\infty}$ is said to be *feasible* if it satisfies the resource constraint in (3), the non-negativity constraints on dividend payment and retained earnings, (4), the incentive compatibility condition in (10) and participation constraint

$$U_0\left(\{C_t\}_{t=0}^{\infty}\right) = U_0.$$

The objective of the shareholders is to chose among the set of feasible plans to maximize profit, which we turn now to.

B.2 Shareholder Value Maximization

We focus on the steady-state of the economy where the aggregate consumption of shareholders is constant. This implies that the risk-free interest rate, r equals the discount rate of the shareholders:

$$r = \beta.$$

Shareholders choose a plan, $\{C_t(H^t), I_t(H^t), D_t(H^t)\}_{t=0}^{\infty}$, to maximize the present value of firms:

$$E\left[\int_0^{\infty} e^{-rt} D_t dt\right] \tag{11}$$

subject to the feasibility constraint.

B.3 Recursive Formulation

We state without proof here that the history can be summarized by promised utility to the manager, U, and the current level of firm's capital stock, K.[8] We can therefore formulate the contract in a recursive fashion taking (K, U) as state variables. Following the

[8] For formal justification of this result, see Abreu et al. [1990]and Spear and Srivastava [1987].

language of Atkeson and Lucas [1992], we refer to such as allocation rules, to distinguish them from the plans and strategies we discussed above. In this formulation, we think of the entire history of managers' consumption-investment decisions and the realizations of the Brownian motion shocks $B_{o,t}$ and $B_{u,t}$ as being summarized by the pair of state variables (K, U). The principal chooses dividend payment, consumption, investment, and rules of assigning continuation utilities as functions of the state variables (K, U). With a slight abuse of notation, we denote $C(K, U)$, $I(K, U)$ and $D(K, U)$ as rules of assigning managerial consumption, investment, and dividend payout as a function of the promised utility and firm's capital stock. Let $\mathcal{K}$ denote the space of possible realizations of capital, and $\mathcal{U}$ denote the space of promised utilities, $C(K, U)$, $I(K, U)$ and $D(K, U)$ are all functions that maps the state space to the real line: $C : \mathcal{K} \times \mathcal{U} \to \mathbb{R}$, $I : \mathcal{K} \times \mathcal{U} \to \mathbb{R}$, $D : \mathcal{K} \times \mathcal{U} \to \mathbb{R}$. In this formulation, the resource constraint (3) is written as:

$$C(K, U) + I(K, U) + D(K, U) = AK. \tag{12}$$

Non-negativity of dividend and retained earnings requires

$$0 \leq D(K, U) \leq AK. \tag{13}$$

The state variable U is the amount of continuation utility has to be delivered to the manager through the continuation of a contract. Because the manager's utility can be represented in the recursive form as in (8) and (9), we think of the principal as delivering the continuation utility via the following procedure. At each point in time, for a given level of promised utility U, the principal allocate the manager's continuation utility over time and states by choosing an instantaneous consumption flow, $C(K, U)$, a rate of change of continuation utility with respect to shocks in the unobservable Brownian motion, $dB_{u,t}$, denoted $G(K, U)$, and a rate of change of continuation utility with respect to the observable Brownian motion, $dB_{u,t}$, denoted $H(K, U)$. A continuation contract from time t on delivers the utility level U_t as long as

$$dU_s = -\bar{f}(C(K_s, U_s), U_s)ds + G(K_s, U_s)\sigma_u dB_{u,s} + H(K_s, U_s)\sigma_o dB_{o,s}, \tag{14}$$

for all $s > t$, and

$$\lim_{T \to \infty} E_s\left[e^{-\beta T} U_T\right] = 0 \tag{15}$$

for all $s \geq t$. The policy functions $G : \mathcal{K} \times \mathcal{U} \to \mathbb{R}$ and $H : \mathcal{K} \times \mathcal{U} \to \mathbb{R}$ describes the rule of assigning continuation utilities based on realizations of the Brownian shocks.

Given an allocation rule, continuation utility is assigned according to (14). Since the principal does not observe the realizations of the Brownian motion B_u, he has to infer it from equilibrium outcomes. In particular, along the equilibrium path, $\sigma_u dB_{u,t}$ can be computed through the law of motion of capital:

$$\sigma_u dB_{u,t} = \frac{dK(t)}{K(t)} - \left[\frac{I(K_t, U_t)}{K_t} - \delta\right] dt - \sigma_o dB_{o,t}, \tag{16}$$

and the principal assigns utility according to:

$$\begin{aligned} dU_t = -\bar{f}(C(K_t, U_t), U_t)dt + G(K_t, U_t)\frac{1}{K_t}\{dK_t - [I(K_t, U_t) - \delta K_t + K_t \sigma_o dB_{o,t}]\} \\ + H(K_t, U_t)\sigma_o dB_{o,t}. \end{aligned} \tag{17}$$

Consider an agent who has conformed to the allocation rule prescribed by the principal up to time t, with promised utility level U_t, and the size of the firm K_t. The agent, however, is contemplating on deviating from the allocation rule $C(K_t, U_t)$ and $I(K_t, U_t)$ by choosing an alternative action, $\tilde{C}_t$, $\tilde{I}_t$. Under the alternative action, the law of motion of capital will be:

$$dK_t = K_t\left[\left(\frac{\tilde{I}_t}{K_t} - \delta\right) dt + \sigma_u dB_{u,t} + \sigma_o dB_{o,t}\right]. \tag{18}$$

Because the principal assigns utility according to (17), the law of motion of the actual utility delivered to the agent is:

$$\begin{aligned} dU_t = \left[-\bar{f}(\tilde{C}_t, U_t) + \frac{G(K_t, U_t)}{K(t)}\left[\tilde{I} - I(K_t, U_t)\right]\right] dt \\ + G(K_t, U_t)\sigma_u dB_{u,t} + H(K_t, U_t)\sigma_o dB_{o,t}. \end{aligned} \tag{19}$$

Let $F(K, U)$ be the maximum utility can be achieved by the agent give the promised utility U and size of the firm, K, $F(K, U)$ has to satisfy the HJB of the agent's utility maximization problem:

$$\begin{aligned} &\max_{\tilde{C},\tilde{I}} \left\{\bar{f}\left(\tilde{C}, U\right) + \mathcal{L}^{\tilde{C},\tilde{I}} F(K, U)\right\} \\ \text{subject to} \quad & \tilde{C} + \tilde{I} \leq C(K, U) + I(K, U), \end{aligned} \tag{20}$$

where $\mathcal{L}^{\tilde{C},\tilde{I}}$ is the infinitesimal generator of the controlled Markov process, (18) and (19) under the decision rule $\tilde{C}, \tilde{I}$, that is,

$$\mathcal{L}^{\tilde{C},\tilde{I}}\mathcal{F}(K, U) = \lim_{\Delta \to 0} \frac{1}{\Delta} E^{\tilde{C},\tilde{I}}\left[F(K_{t+\Delta}, U_{t+\Delta}) - F(K_t, U_t)\right| U_t = U, K_t = K].$$

The notation $E^{\tilde{C},\tilde{I}}$ emphasizes that the law of motion of U_t and K_t follow (18) and (19) under the decision rule $\tilde{C}, \tilde{I}$. Incentive compatibility requires that $C(U, K)$ and $I(C, K)$

is a solution to the HJB in (20) and the value function of the manager coincides with promised utility:

$$F(K,U) = U.$$

In this formulation, an allocation rule is said to be feasible if it satisfies conditions (12) and (13), and is incentive compatible. Finally, note that equation (17) makes it clear that the marginal utility of an additional unit capital to the agent is $\frac{G(K_t,U_t)}{K_t}$. This leads to an interpretation of the incentive compatibility constraint in terms of the neoclassical Q-theory, which we will discuss in the next section.

C Definition of Equilibrium

We will focus on the competitive equilibrium of the economy where the claims to equity of all firms are traded. We focus on the steady-state equilibrium in which aggregate quantities are constant. Thanks to the recursive formulation discussed above, firms can be identified by the state variables (U,K). Let $\Phi(U,K)$ denote the density of the distribution of firms' type in steady-state. The equilibrium allocation consists of the choice of dividend payment, managerial compensation, investment, continuation utility assignment rules, and amount of labor hired for firms of all types: $\{D(U,K), C(U,K), I(U,K), G(U,K), H(U,K), N(U,K)\}_{(K,U)\in\mathcal{K}\times\mathcal{U}}$, the initial capital of newly built firms, K_0, and the total amount of consumption of shareholders, $\mathbf{C}^P$. Equilibrium prices include wage rate W, the price of equity for firms of each type, $\{V(U,K)\}_{U,K}$, and equilibrium interest rate r. A competitive equilibrium is a list of equilibrium allocations and prices such that

1. Maximization of operating profit on the production market: given the equilibrium wage, firms of all types choose $N(U,K)$ to maximize operating profit as in (1).

2. Shareholder value maximization: given the equilibrium interest rate, for firms of all types, the allocation $D(U,K), C(U,K), I(U,K), G(U,K), H(U,K)$ maximize the value of equity defined in (11) among all feasible allocations.

3. Intertemporal maximization of shareholders: equilibrium interest rate is constant with shareholder's intertemporal optimization, that is, $r = \beta^*$ in steady-state.

4. Labor market clearing: the total amount of labor hired by all firms sum up to the total labor endowment, 1:

$$\int N(U,K)\,\Phi(U,K)\,dU dK = 1 \tag{21}$$

5. Product market clearing: total consumption of the shareholders and managers, and total investment in existing firms and in creating new firms sum up to total output:

$$\begin{aligned}\mathbf{C}^P + \int_{[U_L,U_H]} K_0(U)\,\Lambda(U)\,dU + \int \left[C(U,K) + I(U,K)\right]\Phi(U,K)\,dU\,dK \\ = \int zK^{\alpha} N(U,K)^{1-\alpha}\,\Phi(U,K)\,dU\,dK \quad (22)\end{aligned}$$

We solve for the optimal contract for a single firm in the next section, and close the model in general equilibrium and characterize the distribution $\Phi(U,K)$ in Section IV.

II Optimal Contract

A Incentive Compatibility

In this section, we characterize the optimal contract between shareholders and managers. First, we show that the incentive compatibility constraint can be written as a constraint on the sensitivity of the continuity utility with respect to the unobservable Brownian motion, $G(K,U)$. This is summarized by the following lemma.

Lemma 1. (Characterization of Incentive Compatibility) *An allocation rule*

$$\{C(K,U), I(K,U), D(K,U), G(K,U), H(K,U)\}$$

is incentive compatible if for all $(K,U) \in \mathcal{K} \times \mathcal{U}$*,*

$$\frac{\partial}{\partial C}\bar{f}(C(K,U),U)\left[1 + \Omega_I(I(K,U),K)\right] = \frac{G(K,U)}{K}. \quad (23)$$

The above condition highlights the intertemporal choice problem of the manager. The term $\frac{\partial}{\partial C}\bar{f}(C(K,U),U)$ is the marginal utility of consumption of the manager. $1 + \Omega_I(I(K,U),K)$ is the cost of investment goods in terms of current period consumption numeraire. Therefore the left-hand side of equation (23) is the marginal cost of investment in utility terms. It is clear from equation (19) that $\frac{G(K,U)}{K}$ measures the increase in continuation utility for one additional unit of investment. Optimality (incentive compatibility) requires marginal cost of investment to equal its marginal benefit. Equation (23) can be interpreted as the Q-theory relation from the perspective of the manager. Note that $\frac{G(K,U)}{K}$ is the marginal benefit of an additional unit of installed capital measured in utility terms. Normalizing by marginal utility of consumption, $\frac{G(K,U)}{K \times \frac{\partial}{\partial C}\bar{f}(C(K,U),U)}$ is the shadow price of an additional unit of installed capital measured in current period consumption numeraire, that is marginal Q.

Thanks to the above lemma, feasibility of allocation rules can be summarized by the constraints in (12), (13), (15) and (23). The principal choose among all feasible plans to maximize profit. In our formulation, the utility function of the manager is homogeneous of degree $1-\gamma$ in consumption. Using Proposition 1, the operating profit function is linear in capital K. Together, this imply that the value function of the principal, $V(K,U)$ satisfies:

$$V(K,U) = v\left(\frac{U}{K^{1-\gamma}}\right)K \tag{24}$$

for some $v:\mathbb{R}\to\mathbb{R}$. We denote

$$u = \frac{U}{K^{1-\gamma}} \tag{25}$$

to be the normalized utility, and call $v(u)$ the normalized value function. Note also, $v(u)$ is the "average Q" of the firm, i.e. the ratio of the total value of the firm divided by the replacement cost of capital. We also normalize the policies and write:

$$c(u) = \frac{C(K,U)}{K};\ i(u) = \frac{I(K,U)}{K};\ d(u) = \frac{D(K,U)}{K};\ n(u) = \frac{N(K,U)}{K}$$

$$g(u) = \frac{G(K,U)}{(1-\gamma)UK};\ h(u) = \frac{H(K,U)}{(1-\gamma)UK}.$$

Using the above notation, condition (23) can be written as:

$$g = (\beta+\kappa)c^{-\frac{1}{\psi}}\left[(1-\gamma)u\right]^{\frac{1/\psi-1}{1-\gamma}}\left[1+\omega\left(i(u)-i^*\right)\right].$$

The normalized value function $v(u)$ and the policies must solve the following HJB equation:

$$0 = \max_{c,i,dg,h}\left\{\begin{array}{c} d+[i-r-\kappa-\delta]\,v(u) \\ +(1-\gamma)uv'(u)\left[\begin{array}{c} \frac{\beta+\kappa}{1-1/\psi}\left(1-c^{1-\frac{1}{\psi}}\left[(1-\gamma)u\right]^{\frac{1/\psi-1}{1-\gamma}}\right) \\ -i+\delta-\frac{1}{2}\gamma\left(\sigma_u^2+\sigma_o^2\right)+\gamma g\sigma_u^2+\gamma h\sigma_o^2 \end{array}\right] \\ +\frac{1}{2}(1-\gamma)^2u^2v''(u)\left[(g-1)^2\sigma_u^2+(h-1)^2\sigma_o^2\right] \end{array}\right\}$$

$$\text{subject to}\ :\ g = (\beta+\kappa)c^{-\frac{1}{\psi}}\left[(1-\gamma)u\right]^{\frac{1/\psi-1}{1-\gamma}}\left[1+\omega\left(i-i^*\right)\right]$$

$$0\le c+i+\frac{1}{2}\omega\left(i-i^*\right)^2+d\le A \tag{26}$$

Throughout this section, we will assume, without a proof that the value function $V(K,U)$ is strictly concave in (K,U). Intuitively, risk aversion of the agent induces concavity in the value function. A formal proof of this results, however, is quite involved and relies on technical conditions that are hard to establish in general.

B A Lower Bound on the Principal's Value Function

In this section, we start by computing a lower bound of the normalized value function. This will help us understand the basic trade off between incentive provision and risk sharing we focus on in this paper. It will also help us to motivate our assumptions on the parameters values of the model.

The lower bound on the principal's value function can be constructed by identifying a simple form incentive compatible plan. Consider the following class of compensation contract. The total payment from the principal to the agent is $\xi A K_t$, where $0 \leq \xi \leq 1$, and the rest of output is paid out to shareholders as dividend. Given ξ, the agent chooses optimally the division of $\xi A K_t$ between his own consumption and investment in the firm. We call a plan of this form a constant-share plan. As we vary ξ from 0 to 1, this procedure traces out the market value of a firm as a function of the utility delivered to the manager. We denote this value function as $v_{LB}(u)$, where u is the normalized utility as discussed in the above section. Note $v_{LB}(u)$ provides a lower bound to the value function of the principal, as a constant-share plans are incentive compatible by construction, but not necessarily optimal.

We denote the normalized consumption, investment of the agent under a constant-share plan ξ, by $c_{LB}(\xi)$, $i_{LB}(\xi)$, respectively. We also denote the normalized utility of the manager under the constant share plan ξ by $u_{LB}(\xi)$. It is convenient to introduce the notation

$$\hat{\beta} = \beta + \kappa + \left(1 - \frac{1}{\psi}\right)\delta + \frac{1}{2}\gamma\left(1 - \frac{1}{\psi}\right)\left[\sigma_u^2 + \sigma_o^2\right]. \tag{27}$$

We first make the following assumption on $\hat{\beta}$.

Assumption 1:

$$\hat{\beta} = \beta + \kappa + \left(1 - \frac{1}{\psi}\right)\delta + \frac{1}{2}\gamma\left(1 - \frac{1}{\psi}\right)\left[\sigma_u^2 + \sigma_o^2\right] > 0 \tag{28}$$

The follow lemma derives the formulas for $c_{LB}(\xi)$, $i_{LB}(\xi)$ and $u_{LB}(\xi)$ in the case with no adjustment cost. The case with adjustment cost is provided in the Appendix.

Lemma 2. (Manager's Choice Under Constant-Share Plans) *Assuming no adjustment cost, i.e. $\omega = 0$. Under Assumptions 1 and 2, given a constraint-share plan ξ, the manager's investment and consumption decision are given by:*

$$c_{LB}(\xi) = (1 - \psi)\xi A + \psi\hat{\beta};\ i_{LB}(\xi) = \psi\left(\xi A - \hat{\beta}\right). \tag{29}$$

The normalized utility of the manager is

$$u_{LB}(\xi) = \frac{1}{1-\gamma}\left\{\frac{\beta+\kappa}{\left[(1-\psi)\,\xi A+\psi\hat{\beta}\right]^{1/\psi}}\right\}^{\frac{1-\gamma}{1-1/\psi}}. \tag{30}$$

Proof. See Appendix. □

Note the marginal product of capital is constant, A. ξA is the marginal product of capital to the manager under constant-share plans. A higher ξ corresponds to a higher return on investment from the manager's perspective. Equation (29) implies that investment is always increasing in ξ. This is due the substitution effect: investment is increasing in the return on capital everything else being equal. The sensitivity of investment with respect to ξ is governed by the the manager's attitude toward intertemporal elasticity of substitution. Consumption of the manager maybe increasing or decreasing in ξ depends on the IES of the agent. For $\psi < 1$, income effect dominates, and consumption is increasing in ξ. For large IES, substitution effect dominates, and consumption decreases with ξ.

Note Assumption B guarantees that the life-time utility of the agent is well-defined for all $\xi \in [0,1]$. Intuitively, when $\xi = 0$, the only way that the agent can provide himself consumption is to set $i_{LB}(\xi)$ to be negative. This may imply a consumption process that converges to zero too fast, and render the utility of the agent $-\infty$.[9]. Assumption1 guarantees that this will never be the case. Under Assumptions 1, $u_{LB}(\xi)$ is an increasing function of ξ. As we vary ξ from 0 to 1, $u_{LB}(\xi)$ increases from $u_{LB}(0)$ to $u_{LB}(1)$. We define

$$u_L = u_{LB}(0) = \frac{1}{1-\gamma}\left\{\frac{\beta+\kappa}{\left[\psi\hat{\beta}\right]^{1/\psi}}\right\}^{\frac{1-\gamma}{1-1/\psi}}, \tag{31}$$

and

$$u_H = u_{LB}(1) = \frac{1}{1-\gamma}\left\{\frac{\beta+\kappa}{\left[(1-\psi)\,A+\psi\hat{\beta}\right]^{1/\psi}}\right\}^{\frac{1-\gamma}{1-1/\psi}}.$$

Note that u_L is the lowest utility of the agent under any feasible plan. Since consumption and investment are unobservable, the most the principal can do to lower the utility of the agent is to set

$$C_t + I_t = 0, \text{ for all } t.$$

[9]For example, Assumption 1 may fail if $\psi < 1$ and the capital depreciation rate δ is large.

In this case, the agent will chose optimally to provide himself with consumption by setting I_t negative, the resulting life-time utility of the agent is given by (31). Similarly, given the constraint (4), the highest utility that the principal can provide the agent is by setting

$$C_t + I_t = AK_t, \text{ for all } t.$$

This is equivalent to a constant-share plan with $\xi = 1$. The close interval $[u_L, u_H]$ is therefore the domain of the normalized value function $v(u)$. Under our assumptions, it is possible to provide the manager with higher utility than u_H or lower utility than u_L and all utility levels on the interval $[u_L, u_H]$ can be delivered by some feasible plan.

Note that ξ can be loosely interpreted as the equity share of the manger. Under constant share plans, ξ is the fraction of total output of the firm delivered to the manager. If $\xi = 1$, then the manager effectively owns the firms, as all output of the firm is consumed by the manager. As $\xi \to 0$, the share of output of the firm consumed by the manager gets smaller. If $\xi = 0$, the manager consume from the capital stock of the firm by choosing negative investment. Therefore the case $\xi = 0$ should be interpreted as the least equity share of the manger in the firm under all feasible plans, rather than zero equity share. Using this interpretation, equation (30) defines an monotonic mapping from the space of utilities to the space of manager's equity share, the $[0, 1]$ interval:

$$\xi(u) = \frac{1}{1-\psi}\left\{(\beta+\kappa)^{\psi}\left[(1-\gamma)u\right]^{\frac{1-\psi}{1-\gamma}} - \psi\hat{\beta}\right\}. \tag{32}$$

Because any monotonic transformation of utility function represent the same preference, equation (30) define a convenient normalization of utility that can be used to compare the value function of the firm for different preference parameters of the managers.[10]

Under a constant-share plan ξ, the investment policy of the firm is given by equation (29). The dividend paid by the firm at time t is $(1-\xi)AK_t$. Since there is no aggregate risk, the value of such a firm can be calculated using Gordon's formula:

$$\frac{(1-\xi)A}{r+\kappa+\delta-i(\xi)}K(t). \tag{33}$$

As we vary ξ from 0 to 1, Equations (30) and (33) together define the value of the firm as a function of the agent's utility u under plans with constant share. This is summarized by the lemma below. Again, the case with adjustment cost is discussed in the Appendix.

[10] Note as we change the preference parameters of the agent, the interval $[u_L, u_H]$ will also change. Mapping utilities to equity share allows us to "normalize" the domain of firms' value function to $[0, 1]$ across all specifications of the utility function.

Lemma 3. (Lower bound of the Value Function) *The value of the firm under constant-share plans is given by;*

$$v_{LB}(u) = \frac{A - \xi(u)}{r + \kappa + \delta + \psi\left[\hat{\beta} - \xi(u)\right]}, \;\; u \in [u_L, u_H]$$

where $\xi(u)$ *is given in (32).*

Proof. See Appendix. □

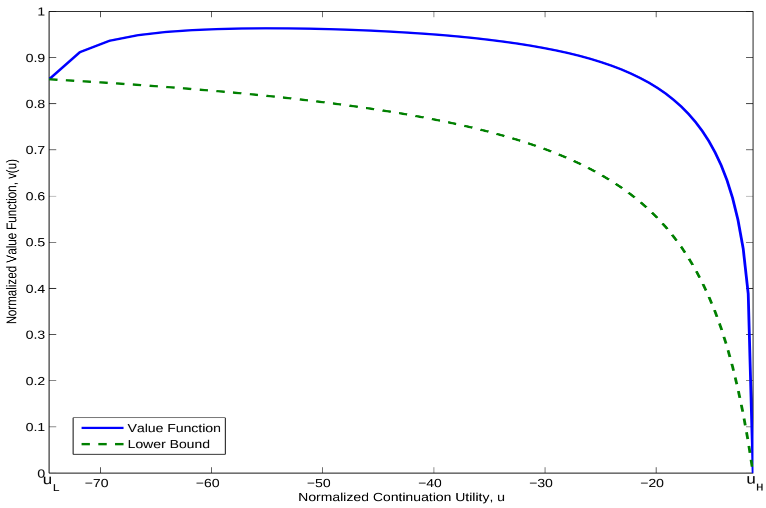

Figure 1: Normalized Value Function and the Lower Bound of the value function.

Note the $v_{LB}(u)$ provides a lower bound for the value function of the principal, because constant-share plans are incentive compatible by construction. Moreover, $v(u)$ must coincide with $v_{LB}(u)$ on the boundaries w_L, and w_H, because the plans that provide agents with utilities u_L and u_H are unique. Figure 1 depicts that true value function $v(u)$ and the lower bound $v_{LOB}(u)$ on the domain $[u_L, u_H]$ for our model under the calibrated parameter values we discuss in Section IV of the paper. Note that $v(u)$ is above the lower bound in the entire domain, and coincides only on the two boundaries. The value function $v(u)$ is the solution to the ODE in (26) with the boundary conditions

$$v(u) = v_{LB}(u) \;\text{ for } u = u_L, u_H.$$

We are now ready to present some basic properties of the optimal incentive compatible contract.

C Incentive Provision and Risk Sharing

The policies g and h are essential in understanding the dynamics of the contract because they measure the sensitivity of the continuation utility of the manager with respect to unobservable and the observable Brownian shocks under the optimal contract. Standard results from the static setting imply the use of relative performance evaluation in incentive compatible contracts (for example, Holmström [1982]), that is, the agents should not be rewarded for higher output due to observable exogenous shocks. The above is no longer true in our dynamic setting because observable shocks have persistent effect and then it is optimal to adjust incentive provision in response to observable shocks[11]. We provide a generalization of this result and show that under the optimal contract, the partial derivative of the value function with respect to continuation utility should not respond to observable shocks. We use $V_U(K,U)$ to denote the partial derivative of firm's value with respect to the promised utility, U, and the notation $Cov_t[dV_U(K_t,U_t), dB_{o,t}]$ to denote

$$\lim_{\Delta\to 0}\frac{1}{\Delta}Cov_t\left[V_U(K_{t+\Delta},U_{t+\Delta}) - V_U(K_t,U_t), B_{o,t+\Delta} - B_{o,t}\right].$$

Proposition 2. (Constancy of $V_U(K_t,U_t)$)

Under the optimal contract, $V_U(K_t,U_t)$ does not respond to observable shocks.[12] *Consequently,*

$$Cov_t\left[dV_U(K_t,U_t), dB_{o,t}\right] = 0 \tag{34}$$

To understand Proposition 2, note that the design of the continuation contract can be viewed as a process of assigning continuation utilities across time and states of nature. Optimality requires that the marginal cost of utility across states is equalized if continuation utility can be assigned in an unconstrained way. Note that the principal must respect incentive compatibility when allocating continuation utilities across unobservable states; however there is no constraint on doing such across observable states. In the absence of aggregate uncertainty, the relative price of consumption goods is constant across states. If utility is additively separable, as is true in the case of expected utility, the marginal cost of utility across observable states must be constant. In continuous time, this is equivalent to the diffusion coefficient of $\frac{\partial}{\partial U}V(K,U)$ on B_o being zero, as in (34). Recursive utility is in general not additively separable. However, the normalization we choose in the representation of the preference has the feature that aggregation of utility

[11] See Li [2011] and Hoffmann and Pfeil [2010] for this issue.

[12] A similar condition is provided in Piskorski and Tchistyi [2011] in the case of risk-neutral agent, and in Li [2011] in the case of risk averse agent with expected utility. Both papers consider the case of jump risk.

across states is additive.[13] In general, one can prove that in the absence of aggregate uncertainty, the above proposition holds as long as we choose the normalization with $\mathcal{A}(U) = 0$, where $\mathcal{A}$ is the "variance multiplier" of stochastic differential utility (in the language of Duffie and Epstein [1992]) in Equation (8).[14]

Optimality also requires the marginal cost of utility be equalized across time. Because expected utility is additively separable with respect to time, the same argument above implies the following Corollary of Proposition 2.

Corollary 1. *In the case of expected utility, $e^{(\beta+\kappa-r)t}V_U(K_t, U_t)$ is a martingale.*

Note that $V(K_t, U_t)$ is the time t value of the firm, and U_t is the time$t$ utility of the agent. The term, $e^{(\beta+\kappa-r)t}$, adjusts for discounting in firm value and that in utility. In settings where consumption and effort are separable, $V_U(K_t, U_t)$ equals the inverse of the instantaneous marginal utility of consumption, and Corollary 1 leads to the well-known "inverse Euler equation" of Rogerson [1985], Spear and Srivastava [1987]and Golosov et al. [2003]. Our Proposition 2 can therefore be viewed as a generalization of the "inverse Euler equation" to the case of recursive preference with the presence of observable shocks. It is worth noting that our Proposition 2 implies that, under the same separability conditions in Golosov et al. [2003], current consumption of the agent does not respond to contemporaneous observable shocks.

Because consumption and investment are non-separable in our context, the inverse marginal utility of the agent will no longer be a martingale. However, $e^{(\beta+\kappa-r)t}V_U(K_t, U_t)$ will be a martingale if utility is separable with respect to time (for example in the special case of CRRA), and will not respond to observable shocks in $dB_{o,t}$ as long as the appropriate normalization of utility is chosen so that aggregation of utility is additively separable with respect to states.

Since h is the measure of the sensitivity of continuation utility with respect to observable shocks, a direct application of Proposition 2 allows us to provide a characterization of the optimal policy h, which is summarized as follows.

[13] The stochastic differential utility representation of the preference we choose can be viewed as the limit of the following discrete time representation of the preference:

$$U_t = \left\{ \left(1 - e^{-\beta\Delta}\right) C_t^{1-1/\psi} + e^{-\beta\Delta} \left(E_t \left[U_{t+\Delta}\right]\right)^{\frac{1-1/\psi}{1-\gamma}} \right\}^{\frac{1-\gamma}{1-1/\psi}}.$$

Note that, in this case, the certainty equivalence functional (Epstein and Zin [1989]) is the expectation operator, and therefore additively separable.

[14] Ai and Li (2001) provide a generalization of Proposition 2 in an environment where the unobservable shock requires a risk premium.

Proposition 3. (Optimal Response to Observable Shocks) *The optimal choice of sensitivity of continuation utility with respect to observable shocks, h is given by:*

$$h = 1 - \frac{\gamma v'(u)}{(1-\gamma) uv''(u)}.$$

If the (un-normalized) value function $V(K, U)$ is strictly concave in K, then $h > 0$ for all $u \in [u_L, u_H]$.

Assume $V(K, U)$ is strictly concave in U, and u_{MAX} is an interior maximizer of $v(u)$ on $[u_L, u_H]$, then $h > 1$ for all $u < u_{MAX}$ and $h < 1$ for all $u > u_{MAX}$.

Proof. See Appendix. □

There are two important implications of Proposition 3. First, continuation utility does respond positively to observable shocks. This is in contrast with the results obtained in the static setting, for example, Holmström [1982]. Note an observable shock is $dB_{o,t}$ increases capital stock, K. By Proposition 2, $V_U(K, U)$ cannot respond to this shock; therefore, the static result $h = 0$ holds only if the cross partial derivative $\frac{\partial^2}{\partial K \partial U} V(K, U) = 0$. In the case $\frac{\partial^2}{\partial K \partial U} V(K, U) > 0$, an increase in K raises $V_U(K, U)$. It is therefore optimal for the principal to increase U to equalize $V_U(K, U)$ across states if $V(K, U)$ is concave in U. Consequently, $h > 0$, that is, continuation utility responds positively to observable shocks. We show in the appendix that homogeneity of the value function (24) implies that concavity of the value function with respect to K is equivalent to $\frac{\partial^2}{\partial K \partial U} V(K, U) > 0$. The intuition is that concavity and homogeneity of the value function implies complementarity of capital and promised utility: the cost of providing promised utility is lower when the size of the firm is large. In this case, the principal chooses optimally to allocate higher continuation utility in the states where the size of the firm is large.

The second implication of the above proposition is responding to an observable shock, $h < 1$ if and only $u > u_{MAX}$. To understand the implication of this result, it is useful to write down the law of motion of the normalized continuation utility, u:

$$du_t = (1-\gamma) u_t [\mu_u(u_t) dt + (g-1) \sigma_u dB_{u,t} + (h-1) \sigma_o dB_{o,t}],$$

Note that $(1-\gamma) u > 0$ under our formulation of the preference. Therefore $h - 1$ determines that sign of loading of the normalized utility u with respect to the observable shock. $0 < h < 1$ implies that with an increase in K_t due to the observable shock $dB_{o,t}$, the change in promised utility U is positive, but less in proportion than that in K_t. Consequently, promised utility U increase with $dB_{o,t}$ in level, but decreases with $dB_{o,t}$ after normalized by K_t. This implies that responding to an observable shock, $dB_{o,t}$, u_t moves to the left in the u-space. Similarly, in the region $u < u_{MAX}$, $h > 1$, and the normalized continuation utility moves to the right after an observable shock.

To explain the intuition behind this result, note that the firm value per-capital is low on the boundaries of the continuation utility region and high in the interior around u_{MAX}. Therefore, if the continuation utility level is lower than u_{MAX}, the per-capital firm value is increasing in the continuation utility. By tying the motion of the continuation utility to the observable noise positively, the shareholder obtains higher per-capital value if the firm size becomes bigger and lower per-capital value when the firm size becomes smaller due to observable shocks. Since the firm becomes bigger and smaller with equal chances, the shareholders are better off by this arrangement. Symmetrically, if u_t is greater than u_{MAX} and the per-capital firm value is decreasing in the continuation utility, by tying the continuation utility negatively to the observable noise, the shareholders, again, obtain higher per-capital firm value when the capital level becomes higher and obtain lower per-capital firm value when the capital level becomes lower due to observable shocks.

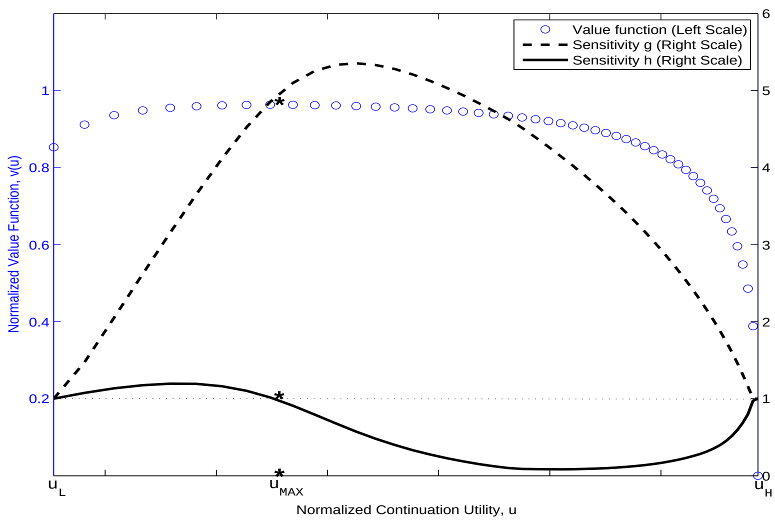

Figure 2: Value Function and Optimal Sensitivity of Continuation Utility w.r.t. Shocks.

Figure 2 depicts the optimal sensitivity of continuation utility with respect to the observable shocks, h (right scale, solid line), that with respect to the unobservable shock, g (right scale, dashed line), and the normalized value function $v(u)$ (left scale, circled line) as functions of the normalized promised utility, u. All quantities are calculated under the parameter values we discuss in the calibration section of the paper. The maximizer of $v(u)$, u_{MAX} is represented by a star on the horizontal axis. Note h is less than 1 to the left of u_{MAX}, and higher than 1 to the right. To understand the condition on $h > 1$, consider first the case in which $u > u_{MAX}$. Here, $V_U(K, U) = v'(u) K^{\gamma} < 0$. In this case, the marginal cost of providing an additional until to the agent is positive. An increase in K increases the marginal cost of continuation utility if U increases proportionally, so that u is kept constant. To equalize $V_U(K, U)$ across states, u should move to the left due to the concavity of $v(u)$, that is $h < 1$.

In the case $u < u_{MAX}$, $V_U(K, U) = v'(u) K^{\gamma} > 0$. This is the region where to punish the agent, the principal has to suffer; therefore it benefit both parties to increase the promised utility of the agent. An increase in K raises the marginal benefit of promised utility if U increases proportionally so that u is kept constant. Concavity of the value function $v(u)$ implies that it is optimal to further increase u to counteract the effect of K on $V_U(K, U)$. This requires u move to the right, and $h > 1$.

The following proposition states that $g \geq h$, that is, continuation utility is more sensitive to unobservable shocks, $dB_{u,t}$ than to observable shocks, $dB_{o,t}$. The intuition is clear, to provide incentives for the managers to invest, continuation utility must respond strongly to unobservable shocks. Continuation utility do not respond as strongly to observable shocks, because it has nothing to do with incentive provision.

Proposition 4. *Optimal Response to Unobservable Shocks*
$g \geq h$ *for all* $u \in [u_L, u_H]$.

The sensitivity of continuation utility with respect to unobservable shocks, g is the dashed line Figure 2. Just as in a standard principal-agent problems, there is a trade-off between risk sharing and incentive provision in our model. To provide incentive for the agent to invest, continuation utility needs to increase after an unobservable shock, $dB_{u,t}$. More sensitivity implies higher incentives for the agent to invest; however, at the same time, it is also associated higher welfare loss due to risk aversion and variations in the continuation utility. The optimal choice of g must trade off incentive provision against risk sharing.

Finally, we note that on the interval (u_{MAX}, u_H), an positive observable shock $dB_{o,t}$ moves firms to the left and a positive unobservable shock $dB_{u,t}$ moves firms to the right in the space of normalized utilities. If observable shocks are more important in accounting for the total volatility in capital accumulation, then we expect firm size as measure in K to be negatively correlated with promised utility, u. In this risk sharing is quantitatively more important than incentive provision, and good shocks is typically (when they are observable) associated a lower equity share of the manager in the firm. In the calibration of our model, we found this to an empirically plausible specification of the parameters. In this case, the predictions of our model is consistent with several empirical regularities on the relationship between firm decisions and firm size. For example small firms tend to invest more, pay less dividend, grow faster, and have higher pay-performance sensitivity than larger firms.

D Relationship with Q-Theory of Investment

Our next result provides a characterization of firms' investment policies as a function of the normalized utility, u.

Proposition 5. (Dividend and Investment Policy) *Suppose $V(K,U)$ is strictly concave in K. Also assume that*

$$\frac{V(U,K)}{K} \leq 1 \text{ for all } (K,U). \tag{35}$$

Then there exist $u_{SWTICH} \in (u_{MAX}, u_H)$ such that

It satisfies $v(u_{SWITCH}) - (1-\gamma)\, v'(u_{SWITCH}) = 1$.

On $[u_L, u_H]$, $\forall u > u_{SWITCH}$, $c(u) + i(u) = A$ and $\forall u < u_{SWITCH}$, $c(u) + i(u) = 0$. That is, all output is paid out as dividend to the left of u_{SWITCH} and no dividend is paid to the right of u_{SWITCH}.

Proof. See Appendix. □

The assumption $\frac{V(U,K)}{K} \leq 1$ can be motivated by the fact that in the case of first best, the value of the firm normalized by its capital stock, or average Q is always 1. Note in economies without moral hazard, standard Q-theory implies that the value of a unit of capital is always equals its replacement cost, 1, in the absence of adjustment cost. Intuition suggests that the value of firms in our economy should always be lower than the case of first best. We suspect condition (35) holds generally, but are not able to provide a formal proof. We observe that condition (35) is never violated in all numerical examples we compute.[15]

The above proposition implies that firms' investment is negative to the left of u_{SWITCH}, and positive to the right. If firm size is negatively correlated with normalized utility, u,

[15] A formal proof is complicated by two issues. First, the value of firms is lower because the manager earns "information rent" and lower that value of the firm. The second is a general equilibrium effect, moral hazard lowers investment and therefore the steady-state level of the marginal product of capital. We suspect that the first effect always dominates, but are not able to establish this result under general conditions.

as we explained in the end of the last subsection, then small firms (firms with higher promised utility u) invest more than large firms. This pattern is shown in Figure 3, where we plot managerial consumption and investment policies. The solid line is managerial consumption normalized by K (right scale) and the dashed line is firm's investment policy normalized by K (right scale). There is a discontinuous point in investment policy, u_{SWITCH}. By the proposition, to the left of u_{SWITCH}, which is represented by the circle on the $u-$axis, $c(u) + i(u) = 0$. Because of the utility function of the manager satisfies the Inada condition at 0, consumption is always strictly positive. This implies investment must be negative in this region, as shown in the figure. That is, in this region, all output are paid as dividend to the shareholders, and manager obtains consumption by setting investment to be negative. On the other hand, for $u > u_{SWITCH}$, $c(u) + i(u) = A$. Here firms do not pay dividend, and consumption and investment are both positive. To understand this pattern of the optimal contract, note by homogeneity of the value function,

$$\frac{\partial}{\partial K} V(K, U) = v(u) - (1-\gamma) u v'(u).$$

The discontinuous point of investment, u_{SWITCH}, is in fact the point where $\frac{\partial}{\partial K} V(K, U) = 1$. We show in the appendix that the concavity of value function V with respect to K is equivalent to

$$\frac{\partial}{\partial u}\left[v(u) - (1-\gamma) u v'(u)\right] > 0.$$

In this case, as shown in Figure 3, $\frac{\partial}{\partial K} V(K, U)$, represented by the squares, is increasing in u. Consequently, $\frac{\partial}{\partial K} V(K, U) > 1$ for firms with $u > u_{SWITCH}$ and $\frac{\partial}{\partial K} V(K, U) < 1$ for firms with $u < u_{SWITCH}$. In the absence of adjustment cost, the marginal cost of investment measured in current period consumption numeraire is 1. The pattern of invest policy in our model thus resembles that in neoclassical models. Firms choose the maximum level of investment whenever the marginal benefit of investment, $\frac{\partial}{\partial K} V(K, U)$ exceed 1, and choose the maximum level of disinvestment whenever $\frac{\partial}{\partial K} V(K, U) < 1$. In this sense, optimal investment has to satisfy the Q-theory relationship from the perspective of shareholders.

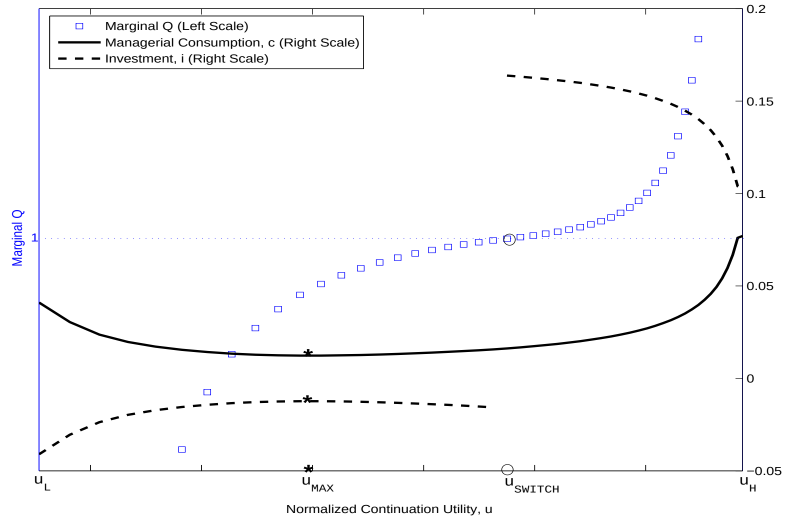

Figure 3: Marginal Q, Managerial Consumption and Investment.

Just like in neoclassical models, concavity of firms' value function and absence of adjustment implies that firms tend to choose investment policies to reach the point where marginal benefit and marginal cost of investment equalizes, i.e., $\frac{\partial}{\partial K} V(K, U) = 1$ as soon as possible. The difference is that concavity of firms' value function arises in neoclassical model because of the concavity of the production function. Here the operating profit function of firms are linear, and the concavity of the value function is induced by risk aversion of the agent. Note that investment policy is discontinuous in u in our model due to the absence of adjustment cost. In the more general version of the model with quadratic adjustment cost, investment will be continuous, and marginal Q will be equalized to the marginal cost of investment at all times, just as in neoclassical models with adjustment cost.

As we remarked before, the incentive compatibility constraint can be interpreted as a Q-theory relation from the perspective of the manager: incentive compatibility requires that managers equalize the marginal benefit of investment and marginal cost of investment. In this sense, Q-theory has to hold from both from the perspective of shareholders and managers. Optimal investment policy requires perfect alignment of managers' and shareholder's interest.

A direct implication of Proposition 3 is that firms with $u > u_{SWITCH}$ invest more, grows faster, and pays more dividend compared to firms with $u < u_{SWITCH}$. As we remarked before if observable shocks are quantitatively large compared to unobservable shocks, in the region $[u_{MAX}, u_H]$, firms to the right of u_{SWITCH} is smaller in size (as measured by total stock of capital, and total number of employees) than firms to the left of u_{SWITCH}. In Section III and IV, we show this is indeed the case under our calibrated parameter values.

Finally, we note that an econometrician who measures marginal Q by average Q is likely to rejected the Q-theory in our economy as the relationship between investment and average Q is not monotone. Despite the homogeneity of technology, average Q, the ratio of the total value of the firm and the total capital stock of the firm does not equal marginal Q, $\frac{\partial}{\partial K}V(K,U)$. As shown in Figure 3, investment is increasing in average Q in the region $u > u_{SWITCH}$. However at u_{SWITCH}, there is a discrete jump in investment and further increasing in average Q is associated with negative levels of investment. In the more general model with quadratic adjustment cost, the change in investment will be continuous, but the non-monotone pattern will remain. As we show in the calibration section of the paper, due to the non-monotonicity of investment with respect to $v(u)$, average Q has little predictive power of investment, and firms' cash flow strongly predict investment. This is mechanism of our model thus provides a potential resolution of the empirical failure of Q-theory regressions.

E The Role of Risk Aversion and IES

The preferences of managers are represented by recursive utility and allows a separation of risk aversion and IES. In this section we show risk aversion and IES play very different roles in terms of the design of the optimal contract.

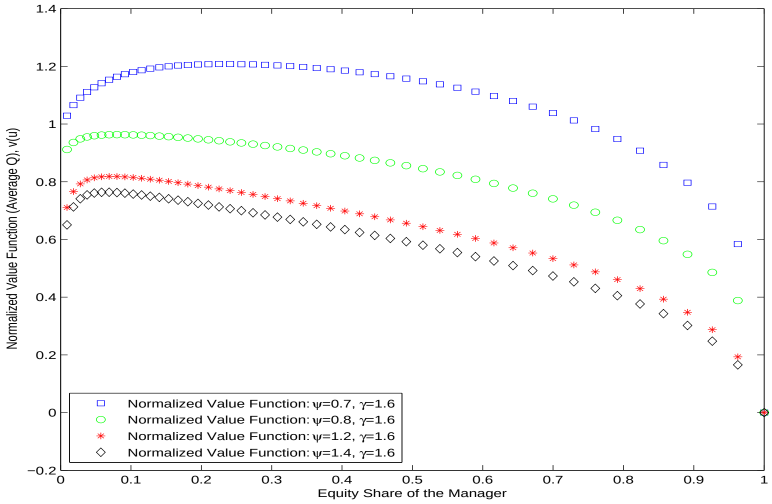

Figure 4: Normalized Value Function and IES.

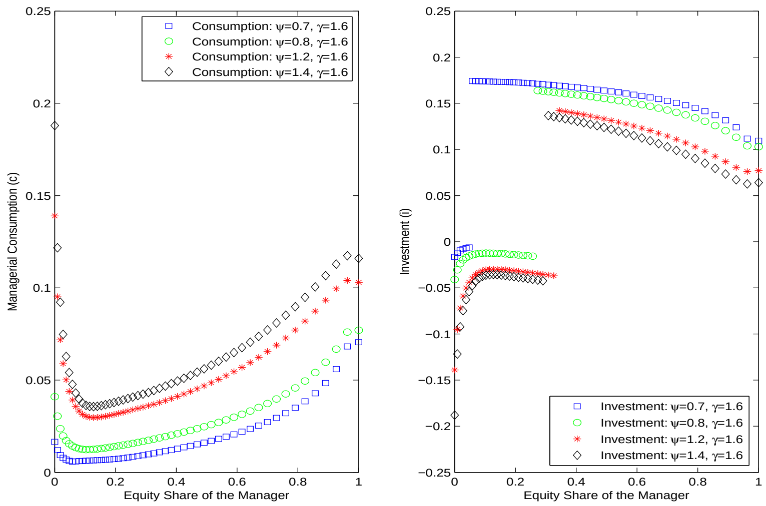

Figure 5: Managerial Consumption, Investment and IES.

We first show that firm value is monotonically decreasing in the IES parameter of managers. Figure 4 plots the value function of the firm for fixed A, fixed level of risk aversion, but different levels of IES of the manager.[16] It is clear that firms' value is

[16]Note that A is endogenous and depends on the steady-state level of capital. Therefore while the

a decreasing function of IES. Figure 5 plots the consumption and investment policies for different parameter values of IES. Figure 5 make it clear that low IES is associated with higher value of the firm because managerial consumption is low and investment is high compared to models with higher levels of IES. The intuition of this result is best explained by an observation made by Tallarini [2000]. In an economy with representative firm and without moral hazard in investment, Tallarini [2000] shows that the quantity dynamics in neoclassical production economies are mostly affected by the IES parameters of the Representative agent. The intuition for his result is that the neoclassical capital accumulation technology is a device to trade off consumption over time and does not allow agents to trade off consumption across states of the world. As perhaps best summarized by Cochrane [1993], the neoclassical capital accumulation technology is Leontiff across states of nature. Consequently, the decision for investment depends mainly on the agent's attitude toward intertemporal substitution, rather than risk aversion. In our economy, firms' value is affected by the conflict of interest between shareholders and managers. Managers, who make the actual decision of investment do not own the entire firm, and therefore have incentives to sacrifice investment for his own private benefit. In the absence of incentive provision schemes, the first best investment policy is not optimal from the perspective of the managers because the marginal cost of investment is higher than the marginal benefit to a benevolent social planner. The incentive to deviate from this first best allocation is stronger for high levels of IES because by definition, high IES means quantities are more sensitive to changes in shadow prices, or marginal utilities. Consequently, high IES is associated with larger deviations from first best investment levels, and therefore the value of firms.

marginal product of capital A are common across all examples, the productivity parameters, z are different across these examples because the steady-state level of capital are different. Therefore the comparison here is a partial equilibrium exercise assuming fixed marginal product of capital. This exercise nevertheless highlights the basic intuitions associated with preference parameters. Comparisons across models with fixed z are possible after we solve for the aggregation problem. We obtain similar qualitative results in the general equilibrium analysis.

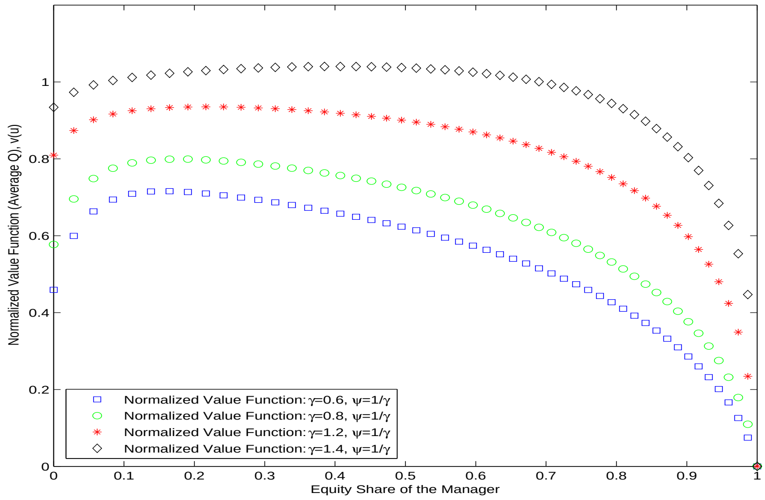

Figure 6: Normalized Value Function with Expected Utility.

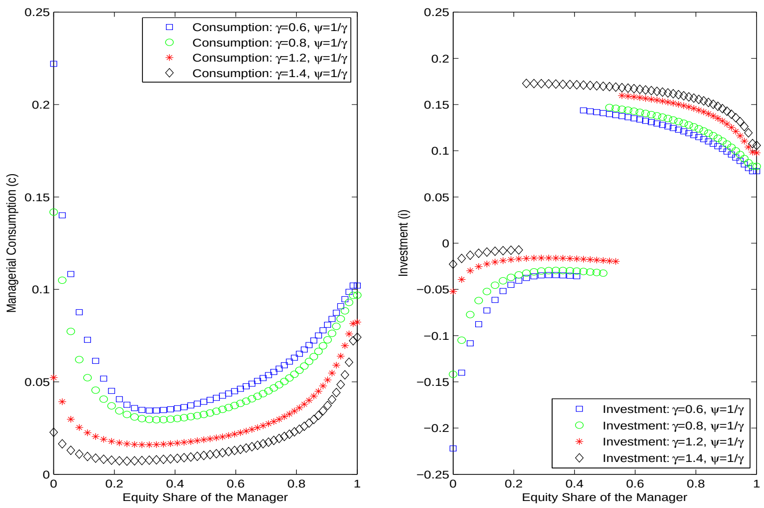

Figure 7: Managerial Consumption, Investment with Expected Utility.

To made this point clear, we plot in Figure 6 the value functions of the firm for fixed A, but different levels of IES (or risk aversion) assuming expected utility. Figure 7 depicts the managerial consumption and investment policies for the same parameter configurations. Note that in the case of expected utility, firms' value and investment are increasing in risk aversion, i.e. decreasing in IES. Managerial consumption is decreasing in risk aversion and increasing in IES. These figures further illustrate that in the case of expected utility, firms values are affected by risk aversion only because risk aversion are tied to IES under expected utility.

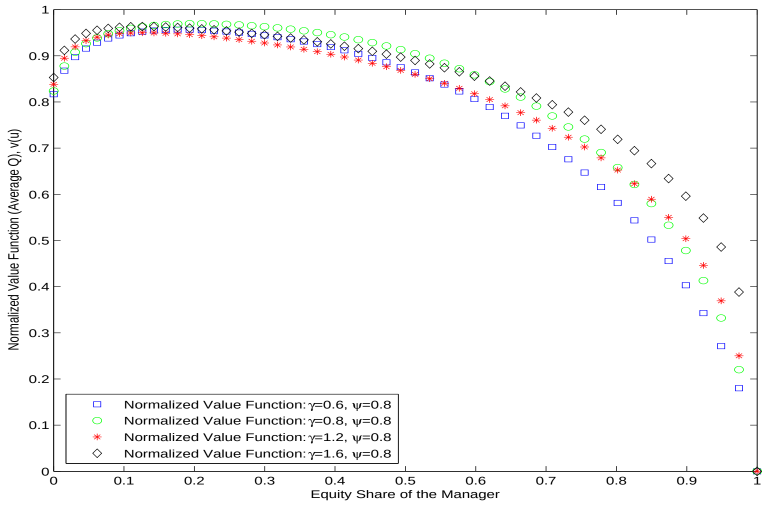

Figure 8: Managerial Consumption, Investment and Risk Aversion.

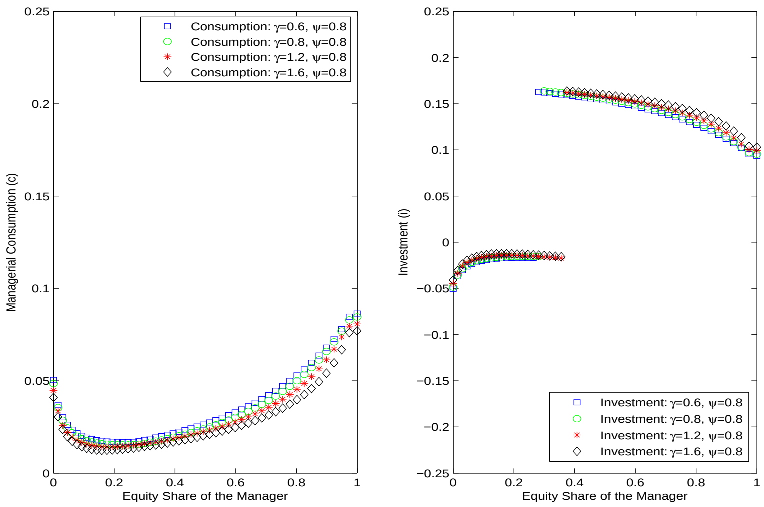

Figure 9: Managerial Consumption, Investment and Risk Aversion.

In Figure 8, we plot the value function of firms for different levels of risk aversion. We also plot managerial consumption and investment decisions in Figure 9. We note that firms' value show much less variation when we change the risk aversion parameter holding IES fixed. Furthermore, they are not monotone in risk aversion. An increase in risk aversion has two effects. First it increases the precautionary saving motive of the manager, and therefore raises the incentives of the managers to invest, all else being equal. This effect tend to increase firm investment and therefore firm value. Next, an increase in risk aversion raise the welfare cost of incentive provision, which requires the continuation utility to vary with realizations of idiosyncratic shocks. The two effects offsets with each other and results in the non-monotone pattern in firms' value.

III Aggregation and the Distribution of Firms

A Aggregation and Numerical Procedure

In this section, we close the model by imposing the market clearing condition and verify our earlier conjecture on the existence of steady state where aggregate consumption of shareholders is constant. In our economy, firms' characteristics can be summarized by state variables (K, U), or equivalently, (K, u), where u is the normalized continuation utility of managers as defined in Equation (25). We use $\phi(t, K, u)$ to denote the density of the joint distribution firms at time t and use $\phi(t, u)$ to denote the density of the marginal distribution of u at time t. That is, $\phi(t, K, u)$ is the measure of firms with characteristic (K, u) at time t and $\phi(t, u)$ is the total measure of firms with normalized promised utility u at time t. Law of large numbers implies that these densities should integrate to $\frac{1}{\kappa}$ as firms enter into the economy at rate 1 and dies at Poisson rate κ.

We first result is on the initial condition of the size of new entrant firms.

Proposition 6. *The normalized initial promised utility to any manager is u and the total capital stock of all initial entrant at any time t is given by*

$$\int_{[U_L,U_H]} \left(\frac{U}{u}\right)^{\frac{1}{1-\gamma}} \Lambda(U)\, dU. \tag{36}$$

Note for each manager with reservation utility U, the manager choose the initial size of the tree. Optimality requires that the marginal benefit of investing in the tree to equal its marginal cost, which is 1 measured in current period consumption numeraire. That is, it is optimal to investment until $\frac{\partial}{\partial K}V(K, U) = 1$. This implies that the normalized promised utility of any manager upon entrance is u_{SWTICH} by Proposition 4. Consequently, the initial capital stock given to a manager with reservation utility U is

$$K_0(U) = \left(\frac{U}{u}\right)^{\frac{1}{1-\gamma}}.$$

The total initial capital stock of all new entrant firms is therefore given by Equation (36) above.

We first show that the density $\phi(t,u)$ satisfies a version of the Kolmogorov forward equation.

Proposition 7. (Marginal Distribution of u) *On $(u_L, u_{SWITCH}) \cup (u_{SWITCH}, u_H)$, $\phi(t,u)$ satisfies*

$$\begin{aligned}\frac{\partial}{\partial t}\phi(t,u) &= -\kappa\phi(t,u) - (1-\gamma)\frac{\partial}{\partial u}[\phi(t,u)\, u\mu_u(u)\, dt] \\ &\quad + \frac{1}{2}(1-\gamma)^2 \frac{\partial^2}{\partial u^2}\left[\phi(t,u)\, u^2\left[(g(u)-1)^2\sigma_u^2 + (h(u)-1)^2\sigma_o^2\right]\right].\end{aligned}$$

Furthermore, $\phi(t,u) \to \phi(u)$, where $\phi(u)$ is the stationary distribution of u and satisfies

$$\begin{aligned}0 = -\kappa\phi(u) - (1-\gamma)\frac{\partial}{\partial u}[\phi(u)\, u\mu_u(u)\, dt] \\ + \frac{1}{2}(1-\gamma)^2 \frac{\partial^2}{\partial u^2}\left[\phi(u)\, u^2\left[(g(u)-1)^2\sigma_u^2 + (h(u)-1)^2\sigma_o^2\right]\right]\end{aligned}$$

on $(u_L, u_{SWITCH}) \cup (u_{SWITCH}, u_H)$.

The forward equation that describe the two-dimensional distribution $\phi(t,K,u)$ can be derived in a similar fashion. However, as noted by Ai [2011], in economies with homogeneous decision rules, the two dimensional distribution $\phi(t,K,u)$ can be summarized by a one dimensional measure $m(t,u)$ defined as follows:

$$m(t,u) = \int \phi(t,u,K)\, K dK. \tag{37}$$

The following proposition shows that $m(t,u)$ also satisfies a version of the Kolmogorov forward equation.

Proposition 8. (The Measure m) *On $(u_L, u_{SWITCH}) \cup (u_{SWITCH}, u_H)$, $m(t,u)$ satisfies*

$$\begin{aligned}\frac{\partial}{\partial t}m(t,u) &= -(\kappa + \delta - i(u))\, m(t,u) \\ &\quad -(1-\gamma)\frac{\partial}{\partial u}\left\{m(t,u)\, u\left[\mu_u(u) + (g(u)-1)\sigma_u^2 + (h(u)-1)\sigma_o^2\right] dt\right\} \\ &\quad + \frac{1}{2}(1-\gamma)^2\frac{\partial^2}{\partial u^2}\left[m(t,u)\, u^2\left[(g(u)-1)^2\sigma_u^2 + (h(u)-1)^2\sigma_o^2\right]\right].\end{aligned}$$

Assume $\kappa+\delta-i(u)>0$ *for all* $u\in[u_L,u_H]$, $m(t,u)\to m(u)$, *where* $m(u)$ *satisfies*

$$\begin{aligned} 0 &= -(\kappa+\delta-i(u))\,m(u) \\ &\quad -(1-\gamma)\frac{\partial}{\partial u}\left\{m(u)\,u\left[\mu_u(u)+(g(u)-1)\,\sigma_u^2+(h(u)-1)\,\sigma_o^2\right]dt\right\} \\ &\quad +\frac{1}{2}(1-\gamma)^2\frac{\partial^2}{\partial u^2}\left[m(u)\,u^2\left[(g(u)-1)^2\sigma_u^2+(h(u)-1)^2\sigma_o^2\right]\right]. \end{aligned}$$

on $(u_L,u_{SWITCH})\cup(u_{SWITCH},u_H)$.

Proofs of Proposition 7 and 8 can be found in Ai [2011]. Using the definition of m, the total capital stock in the economy in steady-state is simply the integral of m:

$$\mathbf{K}=\int_{u_L}^{u_H}m(u)\,du. \tag{38}$$

Because allocation rules are homogeneous in our economy, aggregate output, aggregate consumption of managers, and aggregate investment in existing firms can all be written as integrals against the measure m:

$$\int zK^{\alpha}N(K,U)^{1-\alpha}\,\Phi(K,U)\,dKdU=\int_{u_L}^{u_H}zN(u)^{1-\alpha}\,m(u)\,du,$$

$$\int C(K,U)\,\Phi(K,U)\,dKdU=\int_{u_L}^{u_H}c(u)\,m(u)\,du,$$

and

$$\int I(K,U)\,\Phi(K,U)\,dKdU=\int_{u_L}^{u_H}i(u)\,m(u)\,du,$$

respectively. Using the above, the market clearing conditions (21) and (22) can be written as:

$$\int_{u_L}^{u_H}n(u)\,m(u)\,du=1,$$

and

$$\mathbf{C}^P+\int_{[U_L,U_H]}\left(\frac{U}{u}\right)^{\frac{1}{1-\gamma}}\Lambda(U)\,dU+\int_{u_L}^{u_H}[c(u)+i(u)]\,m(u)\,du=\int_{u_L}^{u_H}zN(u)^{1-\alpha}\,m(u)\,du.$$

Note the construction of measure m reduces the dimensionality of the cross-section distribution of firms and greatly simplifies the computation of equilibrium.

Using the above result, the competitive equilibrium can be calculated as follows.

1. Step 1: Starting from an initial guess of the marginal production of capital A, we solve the optimal contract and allocation rules by solving the ODE (26). Numerically, we use the Markov chain approximation method (Kushner and Dupuis) described in the appendix to solve the optimal control problem.

2. Step 2: After obtaining the policy function $c(u), i(u), g(u), h(u)$, we use Proposition 6 and 7 to construct the density $\phi(u)$ and the measure $m(u)$.

3. Step 3: We use measure m and Equation (38) to calculate the total capital stock in the steady state for the given A.

4. Step 4: We verify that A is the marginal product of capital using Equation (6). If $A > (<) \alpha z \mathbf{K}^{\alpha-1}$, we choose a smaller (larger) A and resolve the contracts by repeating the above steps. We iterate this procedure until convergence, that is, until $A = \alpha z \mathbf{K}^{\alpha-1}$.

B Distribution of Firms

The density $\phi(u)$ and measure $m(u)$ provide a convenient way to visualize the two-dimensional distribution of firms. Figure 11 plot the density of the marginal distribution of u, $\phi(u)$ (solid line, left scale) and the measure $m(u)$ (dashed line, right scale).

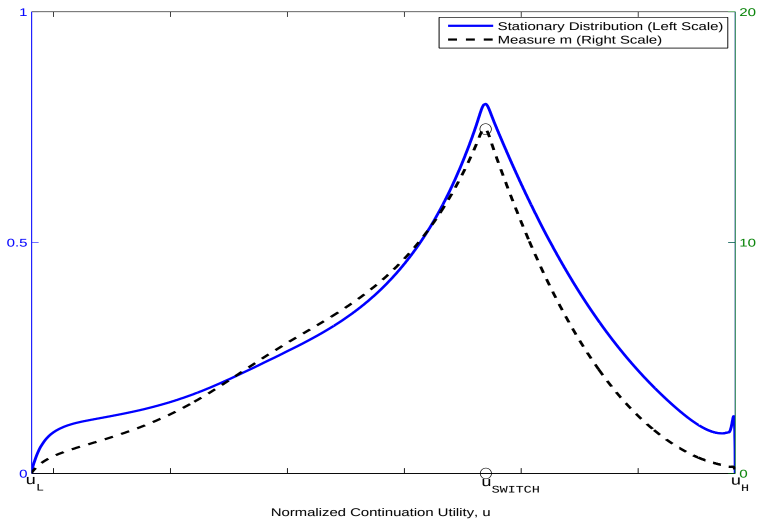

Figure 10: Density of Marginal Distribution of Normalized Utility and the Measure m.

Note both of these measures have a mode at u_{SWITCH} (marked as a circle). As we remarked before, this is the point where the marginal product and the marginal cost of capital equalizes, and therefore firms have a tendency to converge to this point over time. In addition, optimality of entrance implies new firms enter into this economy at u_{SWITCH} as well. Note also, the majority of firms in this economy concentrated on the left of u_{MAX}. As explained earlier, points to the left of u_{MAX} is a "bad region" of the contract: an increase in u will simultaneously increase the the utility of the manager, and the value of the firm. So firms tend to leave this region as soon as possible by choosing negative levels of investment.

By definition of $m(u)$ in Equation (36), the ratio $\frac{m(u)}{\phi(u)}$ is the average firm size at location u in steady state. In figure 11, we plot the average size of firms (solid line, left scale) against normalized utility. We also plot the investment policy (dashed line, right scale) on the same graph.

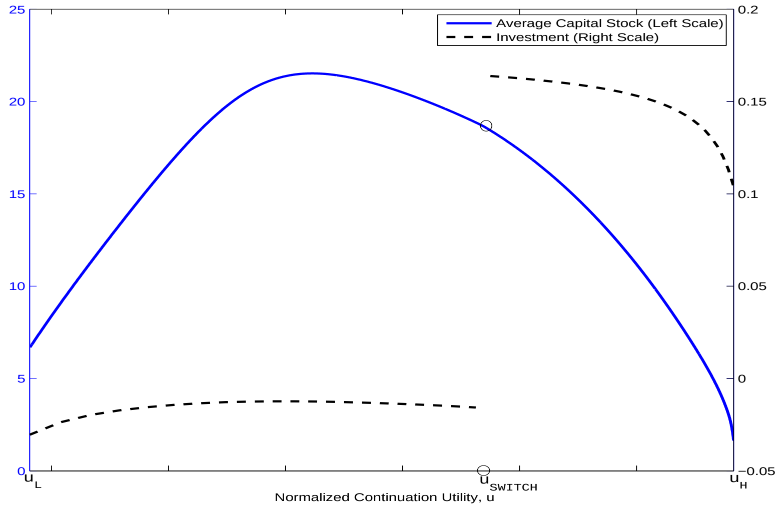

Figure 11: Average Capital Stock and Investment Policy.

To the right of u_{MAX}, where most of the firms in the economy reside, average firm size decreases with investment. This pattern shows clearly our earlier claim that small firms invest more, grow faster, and pay less dividend.

IV Quantitative Implications of the Model

In this section, we calibrate our model to evaluate the quantitative implications of the model on the size-investment and growth relationship and the predictability of investment by firms' average Q and cash flow. We choose the discount rate of the principal so that the interest rate in the economy is $r = 4\%$ in steady state. We calibrate the death rate of firms to be $\kappa = 5\%$ per year, and depreciate rate of firm specific capital $\delta = 8\%$ per year. Together, this imply a depreciation rate of capital of 13% per year, similar to those used in standard RBC models. We choose capital share $\alpha = 0.33$, which roughly matches the income share of capital and labor in US post war data. We choose the IES parameter of managers to be $\psi = 1.2$, so that the steady-state investment-to-output ratio is 25%, which matches its empirical counterpart in the US post war data. We choose the risk aversion parameter of the manager to be $\gamma = 1.6$. Model implications are generally robust to the choice of the risk aversion parameter.

We choose the total volatility of firm-specific capital to be 31% per year, so that the volatility of output at the firm level roughly matches its empirical counterpart in the COMPUSTAT data. We choose $\sigma_u = 6\%$ and $\sigma_o = 30\%$. This specification allows our model to roughly match the tail parameter of the size distribution of firms in the US post war data.

FURTHER RESULTS TO BE ADDED.

V Conclusion

We embedded moral hazard in firms' investment decisions in a general equilibrium model with heterogeneous firms. Using continuous time methods, we solve for the optimal incentive compatible contracts and the implied dynamics of firm investment and firm growth. Theoretically, we provide a characterization of the optimal incentive compatible contract when managers have recursive utility that separates risk aversion and IES. We show that the marginal cost of utility provision must not respond to observable shocks under appropriate normalization of managers' utility function for a large class of preferences. Quantitatively, our model implies that moral hazard in investment decisions is important in accounting the empirical relationship between firm size and firms' investment and dividend payout policies. Our model also provide a potential resolution of the failure of average Q in predicting firm investment when cash flow and other firm characteristics are included as explanatory variables.

VI Appendix

A Appendix 1

A.1 Operating Profit of the Firm

Solving the profit maximization problem in (3), the optimal choice of labor is

$$N_{j,t} = \left[\frac{(1-\alpha)\, z}{W_t}\right]^{\frac{1}{\alpha}} K_{j,t},$$

and the total operating profit is given by:

$$P\left(K_{j,t}\right) = \frac{\alpha}{1-\alpha}\left[(1-\alpha)\, z\right]^{\frac{1}{\alpha}} W_t^{1-\frac{1}{\alpha}} K_{j,t}. \tag{39}$$

Using the market clearing condition (21), we have

$$\int N_{j,t} dj = \left[\frac{(1-\alpha)\, z}{W_t}\right]^{\frac{1}{\alpha}} \int K_{j,t} dj = 1. \tag{40}$$

In steady-state, $\int K_{j,t} dj = \mathbf{K}$; therefore W_t is constant over time. Combing equation (39) and (40), the operating profit function can be written as:

$$P\left(K_{j,t}\right) = \alpha z \left[\frac{1}{\mathbf{K}}\right]^{1-\alpha} K_{j,t} = A K_{j,t},$$

where we denote

$$A = \alpha z \mathbf{K}^{\alpha-1} \tag{41}$$

as the marginal product of capital.

References

Andrew Abel and Janice Eberly. A unified model of investment under uncertainty. *American Economic Review*, 84:1369–1384, 1994.

Dilip Abreu, David Pearce, and Ennio Stacchetti. Toward a theory of discounted repeated games with imperfect monitoring. *Econometrica*, 58:1041–1063, 1990.

Hengjie Ai and Rui Li. Delegated investment and the cross-section of equity returns. *Working paper, Duke University.*, 2011.

Henhjie Ai. A note on aggregation in economies with homogenous decision rules. *Working paper, Duke University.*, 2011.

Rui Albuquerque and Hugo Hopenhayn. Optimal lending contracts and firm dynamics. *Review of Economic Studies*, 71:285–315, 2004.

Costas Arkolakis. A unified theory of firm selection and growth. *Working paper, Yale University*, 2011.

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

Ravi Bansal and Amir Yaron. Risks for the long run: A potential resolution of asset pricing puzzles. *Journal of Finance*, 59:1481–1509, 2004.

Ben Bernanke and Mark Gertler. Agency costs, net worth, and business fluctuations. *American Economic Review*, 79:14–31, 1989.

Ben Bernanke, Mark Gertler, and Simon Gilchrist. The financial accelerator and the flight to quality. *Review of Economics and Statistics*, 78:1–15, 1999.

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

Olivier Blanchard. Debt, deficits, and finite horizons. *Journal of Political Economy*, 93: 223–247, 1985.

Patrick Bolton, Hui Chen, and Neng Wang. A unified theory of Tobin' s q, corporate investment, financing, and risk management. *Journal of Finance, Forthcoming*, 2010.

Charles Carlstrom and Timothy Fuerst. Agency costs, net worth, and business fluctuations: A computable general equilibrium analysis. *American Economic Review*, 87: 893–910, 1997.

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

Gian Clementi, Thomas Cooley, and Sonia Giannatale. A theory of firm decline. *Working paper, MIT*, 2008.

John Cochrane. Rethinking production under uncertainty. *Manuscript, University of Chicago*, 1993.

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

Peter DeMarzo, Michael Fishman, Zhiguo He, and Neng Wang. Dynamic agency and q theory of investment. *Journal of Finance, Forthcoming*, 2009.

Avinash Dixit and Joseph Stiglitz. Monopolistic competition and optimum product diversity. *American Economic Review*, 67:297–308, 1977.

Darrell Duffie and Larry Epstein. Stochastic differential utility. *Econometrica*, 60:353–394, 1992.

Larry Epstein and Stanley Zin. Substitution, risk aversion, and the temporal behavior of consumption and asset returns: A theoretical framework. *Econometrica*, 57:937–969, 1989.

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

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

Eugene Fama and Kenneth French. Disappearing dividends: changing firm characteristics or lower propensity to pay? *Journal of Financial Economics*, 60:3–43, 2001.

Steven Fazzari, Glenn Hubbard, and Bruce Petersen. Financing constraints and corporate investment. *Brookings Papers on Economic Activity*, 1988:141–206, 1988.

Mikhail Golosov, Narayana Kocherlakota, and Aleh Tsyvinski. Optimal indirect and capital taxation. *Review of Economic Studies*, 70:569–587, 2003.

Fumio Hayashi. Tobin's marginal q and average q: A neoclassical interpretation. *Econometrica*, 50:213–224, 1982.

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

Florian Hoffmann and Sebastian Pfeil. Reward for luck in a dynamic agency model. *Review of Financial Studies*, 23:3329–3345, 2010.

Bengt Holmström. Moral hazard in teams. *Bell Journal of Economics*, 13:324–340, 1982.

Hugo Hopenhayn. Entry, exit, and firm dynamics in long run equilibrium. *Econometrica*, 60:1127–1150, 1992.

Boyan Jovanovic. Selection and the evolution of industry. *Econometrica*, 50:649–670, 1982.

Jakob Klette and Samuel Kortum. Innovating firms and aggregate innovation. *Journal of Political Economy*, 112:986–1018, 2004.

David Kreps and Evan Porteus. Temporal resolution of uncertainty and dynamic choice theory. *Econometrica*, 46:185–200, 1978.

Rui Li. Dynamic agency with persistent exogenous shocks. *Working paper, University of Wisconsin-Madison*, 2011.

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

Erzo Luttmer. Technology diffusion and growth. *Journal of Economic Theory, Forthcoming*, 2011.

Edwin Mansfield. Entry, gibrat's law, innovation, and the growth of firms. *American Economic Review*, 52:1023–1051, 1962.

Thomas Philippon and Yuliy Sannikov. Real options in a dynamic agency model with applications to financial development, ipos, and business risk. *Working paper, NYU Stern School.*, 2007.

Tomasz Piskorski and Alexei Tchistyi. Stochastic house appreciation and optimal mortgage lending. *Review of Financial Studies*, 24:1407–1446, 2011.

Vincenzo Quadrini. Investment and liquidation in renegotiation-proof contracts with moral hazard. *Journal of Monetary Economics*, 51:713–751, 2004.

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

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

Stephen Spear and Sanjay Srivastava. On repeated moral hazard with discounting. *Review of Economic Studies*, 54:599–617, 1987.

Jeremy Stein. Agency, information and corporate investment. In *Handbook of the Economics of Finance*. North Holland, 2003.

Thomas Tallarini. Risk-sensitive real business cycles. *Journal of Monetary Economics*, 45:507–532, 2000.