Dividend Policy, Production and Stock Returns: A Dividend-Based Explanation of Asset Pricing Anomalies

Hwagyun Kim*and Changwoo Nam†

May 8, 2011

Abstract

This paper proposes an asset pricing model in a production economy where cash flows are determined by firms' optimal dividend and investment decisions. Extensive and intensive decision margins in dividend payout are modeled with cash holding and investment adjustment costs. The model implies that delays in dividend distribution of young and growing firms play instrumental roles in explaining various asset pricing anomalies by changing the covariations of firms' cash flows and the stochastic discount factor. Quantitative results show that model-implied dividend policies and investments are consistent with data, and the cross sections of stock returns are well explained by the interactions between productivity shocks and the lumpy dividend policies. Additionally, the model produces countercyclical variations in the market risk premium.

JEL: G12, G35, E22

Keywords: Asset Pricing, Cross-Sectional Stock Returns, Payout Policy, Dividend, Firm Heterogeneity

*Email: hkim@mays.tamu.edu, Mailing Address: Department of Finance at Mays Business School, 360 Wehner, 4218 TAMU, Texas A&M University, College Station, TX 77843-4218

†Email: cnam@econmail.tamu.edu, Mailing Address: Department of Economics, Texas A&M University, 4228 TAMU, College Station, TX 77843-4228

1 Introduction

Do corporate decisions made by firm managers matter in explaining the cross-sectional properties of stock returns? Especially, if decisions on investment and dividend payout can affect cash flows of firms, these have direct implications on the equity prices of firms, since the cash flow processes of firms are one of the two main pillars determining stock prices. The other pillar, stochastic discount factor, is closely related to the changes of market investor's wealth given their attitudes toward risk and uncertainty. Since the distributed cash flows of firms will constitute the current and future wealths of market investors, corporate financial decisions related to investments and dividends can have important links to variations in asset prices via channels of both firms' cash flows and investors' discount factors. This paper studies the implications of corporate dividend policies on stock returns in a general equilibrium framework.

There are several studies that attempt to connect the investment and production behaviors of firms to stock returns (e.g., Cochrane (1996), Gomes, Kogan, and Zhang (2003), Zhang (2005), Li, Livdan, and Zhang (2009), and Livdan, Sapriza, and Zhang (2009)). However, the existing literature pays little attention to dividend policy, and the authors assume that dividends are given as residual cash flow. Simple as it may be, the assumption of this residual dividend policy is not consistent with empirical evidence. Figure 1 plots earnings per share and dividends per share of several firms. Dividends appear to be managed with the following patterns: First, there are periods of no dividend payments which make earnings much more volatile, and second, dividends and earnings share a common trend over time, though the latter more volatile, other than the period of zero dividends.[1] Thus, modeling whether and when to pay dividends seems relevant in describing dividend policy.

Empirical studies report that highly related are the extensive/intensive margins of dividend payout and key firm characteristics frequently used in empirical asset pricing studies. Fama and French (2001) find that four characteristics have effects on the decision to pay dividends: profitability, investment opportunities, market-to-book ratio, and size. Larger and more profitable firms are more likely to pay dividends, and paying dividends is less likely for firms with more investments. Grullon, Michealy, and Swaminathan (2002) report that the firm profitability declines after a dividend increase and rises after a dividend decrease. DeAngelo, DeAngelo, and Stulz (2006) present evidence that the probability that a firm pays dividends is significantly related

[1] It is well known that, at the aggregate level, dividend payments tend to be smooth relative to earnings, suggesting that corporate managers manage dividends (Lease, John, Kalay, Lowenstein, and Sarig (1999)).

to the mix of earned capital and contributed capital in its capital structure. Firms with a greater proportion of earned capital are more likely to be dividend payers. Bulan, Subramanian, and Tanlu (2007) use duration analysis to study the timing of dividend initiations in a firm's life cycle, and document that firms initiate dividends after reaching the maturity phase in their life cycles. Putting together, initiators are the firms that have grown larger and have fewer growth opportunities than do non-payers at the same stage in their life cycles.

Based on these observations, we develop a dynamic general equilibrium model of asset prices with many firms whose dividend processes are endogenously determined by firms' optimal dividend policies and investment behaviors. In so doing, both extensive and intensive decision margins in dividend payout are modeled. Corporate managers will maximize the present value of current and future cash flows net of dividends and the costs involved with investment and cash holding. Cash holding is assumed costly to reflect the agency cost or the conflict of interests between shareholders and firm managers. Easterbrook (1984) and Jensen (1986) are the classic studies suggesting that firm managers pursue self-interests, and increase cash holdings for various reasons. In this case, they argue that dividends can help shareholders to reduce the associated agency costs. Recently, Nikolov and Whited (2009) estimate a dynamic model of firm investment and cash accumulation to find that agency problems affect corporate cash holding decisions. They model three specific mechanisms that misalign managerial and shareholder incentives: managerial bonuses on current profits, limited managerial ownership of the firm, and a managerial preference for firm size. Our setup can be viewed as an attempt to incorporate these findings in the corporate finance into a general equilibrium framework.

The model features rich firm dynamics and generates several cross-sectional implications of asset returns and firm characteristics. Regarding the firm dynamics, both analytic and quantitative results show that younger firms with small capital tend to invest more in capital, and they withhold paying dividends. These firms initiate dividend payouts mainly to reduce the increasing cash holding and investment costs, as capital is accumulated. It turns out that this extensive decision margin on firm cash flow depends on firm characteristics associated with productivity shocks.[2] Thus, this model endogenizes not only the amount but also the timing of cash flows distributed from firms to shareholders. Note that the latter is reminiscent of duration-based ex-

[2]These theoretical findings are consistent with the firm life cycle theory of dividends. For more details, see Mueller (1972), Grullon, Michealy, and Swaminathan (2002), DeAngelo, DeAngelo, and Skinner (2004), DeAngelo, DeAngelo, and Stulz (2006). On the firm characteristics for the propensity to pay dividends, see Fama and French (2001) and Denis and Osobov (2008).

planations of the value premium examined by Lettau and Wachter (2007) and Santos and Veronesi (2009). These papers view that growth stocks pay later, while value firms pay now. Alas, if there exist long-run risks (persistent shocks in economic growth, for instance) in an economy and the discount factor co-varies with stock prices, counterfactual growth premiums can occur. To overcome this, Lettau and Wachter (2007) assume that shocks to the discount factor do not co-vary with cash flows shocks, and Santos and Veronesi (2009) assume that cash flows are highly volatile. The former is difficult to justify in an equilibrium, while the latter needs an abnormal amount of cash flow fluctuations. In our model, young and growing firms typically do not pay dividends, and therefore covariations with the discount factor are close to zero. Meanwhile mature firms with larger amounts of capital pay dividends, and tend to have less room for growth. Thus, their short-run cash flows become risky and accordingly priced. In this vein, our paper provides a novel explanation of the value premium which extends and generalizes the insight from the durations of assets based on firms' cash flows.[3]

With this model in hand, we analyze the behaviors of stock returns generated from the simulated cash flows of firms. According to our quantitative results, the expected returns on value and size sorted portfolios are consistent with the historical data (e.g., Fama and French (2007) and Chen, Petkova, and Zhang (2008)). In addition, the simulated expected returns on decile portfolios sorted by book-to-market equity ratio and size present features of value and size premia in the empirical facts (Fama and French (1992)). Interestingly, pooling new decile portfolios sorted by fitted propensity to pay dividends, the simulated expected returns on them consistently increase in the likelihood of dividend payout. This implies that the factors affecting dividend policy largely overlap with the characteristics used to form pricing factors. Finally, the simulated market risk premiums and volatilities of aggregate dividend and asset returns reveal endogenously countercyclical variations, again in line with empirical evidence.

This paper is related to growing literature that links the cross-sectional properties of stock returns to firm characteristics. Gomes, Kogan, and Zhang (2003) construct a dynamic general equilibrium production economy to explicitly link expected stock returns to firm characteristics such as the firm size and the book-to-market equity ratio. Zhang (2005) shows that the value premium occurs from the asymmetric cost of reversibility and the countercyclical price of risk, and assets in place are riskier than growth options especially in bad times when the price of risk is high. Li, Livdan, and Zhang (2009) use a simple q-theory model, and ask if it can explain external financing

[3]The size premium can also arise due to the delay in dividend payout and the profitability of firms in this model.

anomalies, both qualitatively and quantitatively. Livdan, Sapriza, and Zhang (2009) analyze the effect of financial constraints on risk and expected returns by extending the investment-based asset pricing framework to incorporate retained earnings, debt, costly equity, and collateral constraints on debt capacity. However, this paper differs from all these studies in that our model explains the cross section of stock returns using endogenous dividend policies as well as production and investment. As mentioned earlier, models missing this feature can have counterfactual implications on the relationships between stock returns and firm characteristics.

Our work also belongs to another line of literature on the asset pricing model, which extends Lucas (1978) to have many trees (dividends). Cochrane, Longstaff, and Santa-Clara (2008) solve a model with two Lucas trees. They show that expected returns, excess returns, and return volatility as functions of dividend share vary through time, and returns are predictable from price-dividend ratios. Martin (2009) further extends Cochrane, Longstaff, and Santa-Clara (2008) with many trees. Menzly, Santos, and Veronesi (2004) propose a general equilibrium model with multiple securities in which investors' risk preferences, and expectations of dividend growth are time-varying with external habit formation. Santos and Veronesi (2009) show that substantial heterogeneity in firm's cash-flow risk yields both a value premium as well as most of the stylized facts about the cross-section of stock returns, but it generates a "cash-flow risk puzzle". All these papers assume exogenous processes for dividend processes. Therefore, these papers are restrictive in linking firm characteristics to asset prices. In this context, our paper can be viewed as a general equilibrium justification of Lucas models with multiple trees.

The outline for the rest of the paper is as follows: The recursive competitive equilibrium with dynamic problems is described in Section 2, and our theoretical findings about the firm characteristics and the cross-section of stock returns are explored in Section 3. Section 4 outlines the quantitative analysis of cross-sectional stock returns including the baseline calibration, and discusses the consistency of our findings related to similar literature. Finally, Section 5 concludes.

2 Model

2.1 Individual Firm Dynamics

This section sets up the dynamic and stochastic problem of firms. The economy is composed of a continuum of competitive firms that produce a homogeneous product. Firms are subject to an aggregate productivity shock (x_t) and a firm-specific productivity shock (z_{it})[4]. The aggregate shock, x_t, develops according to a first-order autoregressive stationary and monotone Markov transition function, denoted by $dQ_x(x_{t+1}|x_t)$,

$$x_{t+1} = \mu_x(1-\rho_x) + \rho_x x_t + \epsilon_{xt+1},$$

in which ϵ_x follows truncated $\mathcal{N}(0, \sigma_x^2)$, and x_t serves as the driving force of economic fluctuations and systematic risks. x_t has the finite support as $[\underline{x}, \bar{x}]$. Similarly, the firm specific productivity shocks, $\{z_{it}\}_{i\in[0,1]}$ with $dQ_z(z_{it+1}|z_{it})$ are uncorrelated across firms, and follow

$$z_{it+1} = \rho_z z_{it} + \epsilon_{it+1},$$

in which ϵ_i follows truncated $\mathcal{N}(0, \sigma_z^2)$, and ϵ_{xt+1} is independent of ϵ_{it+1} for all i. z_{it} has the finite support as $[\underline{z}, \bar{z}]$.

The production function is given by

$$y_{it} = e^{z_{it}+x_t} K_{it}^{\theta},$$

in which K_{it} is the capital of firm i at time t, and has the compact support of $[\underline{K}, \bar{K}]$. The assumption on capital implies that firms have limited investment opportunities as firms accumulate capital close to the upper bound. The production function exhibits decreasing returns to scale: the curvature parameter satisfies $0 < \theta < 1$. The operating profit is, then, defined as,

$$\Pi(K_{it}, x_t, z_{it}) = e^{z_{it}+x_t} K_{it}^{\theta} - f,$$

in which f is the nonnegative fixed production cost, which is paid in every period[5].

With this production technology in hand, we now describe the firm's decision process below. In the beginning of period t, firm i observes current aggregate and firm-specific shocks. Then, it decides whether to pay dividends to shareholders ($a_{it} = 1$: paying dividends, $a_{it} = 0$: not-paying dividends). In particular, the dividend payment (D_{it}) is determined by a fixed payout ratio $0 < \delta < 1$ as

$$D_{it} = a_{it}\delta\Pi(K_{it}, x_t, z_{it}) = a_{it}\delta e^{z_{it}+x_t} K_{it}^{\theta}. \tag{1}$$

[4] This convention is based on recent macro-finance literature: Zhang (2005), Li, Livdan, and Zhang (2009), and Livdan, Sapriza, and Zhang (2009).

[5] For mathematical simplicity, we assume that f is zero in driving our theoretical results.

To motivate our choice of dividend policy dictated by (1), we plot Figures 1-3. As mentioned earlier, Figure 1 illustrates the existence of zero dividend periods for individual firms. Although we do not tabulate here, a significant number of firms do not pay dividends at any given point of time. Thus, we believe that it is pertinent to include the discrete decision of dividend distribution. In addition, this figure suggests that dividend and corporate earnings follow a common trend, excluding the period of no dividend, which makes (1) a reasonable approximation of reality. We further investigate this issue in Figures 2 and 3. Panel (A) in Figure 2 displays the aggregate payout ratios in the U.S. stock market during 1950 and 2009. Although there exist some variations, the payout ratio is stable around 0.5. To see if the variations result from the usual suspects such as profitability and investments, we run rolling regressions of the aggregate payout ratio onto those variables. We find that both variables are mostly not different from zero, and the negative sign of the profitability is inconsistent with the related theory. At an individual firm level, we project dividends onto earnings and investment, and display in Figure 3. It shows that dividend-earnings ratio is highly concentrated and significant, hinting that the common payout ratio is a reasonable assumption. Meanwhile, the coefficient for investment is concentrated around a small positive number, and often insignificant. Note that the sign should be negative if residual dividend policy is true. That is, the results suggest that the residual policy does not capture the data. All the evidence supports the selection of dividend function.

After the decision of paying dividends, the firm manager of firm i chooses the optimal investment I_{it}^a to maximize the present value of future cash streams under the operating profit less dividends. This cash accumulation decision is motivated from the empirical findings in the corporate finance literature[6]. Related, firms often need to hold a significant amount of funds (cash and marketable securities) to allow for future acquisitions and to cover its considerable legal and business risks (Harford (1999)).[7] It is also well known that there exists an agency conflict between managers and equity providers, and increasing cash holdings further intensify this type of agency problem.[8]

[6] For instance, Nikolov and Whited (2009) argues that although numerous empirical researchers have studied the effects of agency conflicts on cash holding, this topic remains of interest because no single prominent conclusion has emerged from these exercises.

[7] From the perspective of corporate governance, Harford (1999) states that cash is an important tool for firms operating in imperfect capital markets. However, firms often build up much more cash than they need to meet expected financing requirement, which also provides another rational on why managers accumulate cash reserves, and it is subject to a source of agency cost. Harford, Mansi, and Maxwell (2008) find that firms with weaker corporate governance structures actually have smaller cash reserves, and weakly controlled managers tend to spend cash quickly on acquisitions and capital expenditures.

[8] Readers are referred to Jensen and Meckling (1976), Jensen (1986), Easterbrook (1984) for more

That is, increasing the cash holdings of firms is a costly business. Incorporating this stylized fact, we set up a firm's dynamic decision problem with a non-convex agency cost function in the following way.

$$V_i(S_{it}) = \max_{\{a_{it}=0, a_{it}=1\}} V_i^{a_{it}}(S_{it}) \text{ for all } (K_{it}, x_t, z_{it}) \equiv S_{it} \in \mathcal{S}, \tag{2}$$

$$V_i^{a_{it}}(S_{it}) = \max_{I_{it}^a} \{v_{it}^a - \Xi(v_{it}^a, K_{it}, x_t, z_{it}) + \beta \mathbb{E}_t [V_i(S_{it+1}) | S_{it}]\},$$

in which $V^{a_{it}=0}$ is the present value of future cash streams reserved when it does not pay dividends, $V^{a_{it}=1}$ is the present value of future cash flows held by a firm paying dividends, v_{it}^a is the current cash held and used by the firm manager before the cost of reserving the cash, $\Xi(v_{it}^a, K_{it}, x_t, z_{it})$ is the cash holding cost, $\mathcal{S}$ is a compact product space of state variables, and $\mathbb{E}_t$ is the expectation operator at time t.[9] Thus, the firm i maximizes the present value of the net cash flows into the corporate cash holdings, defined as $v_{it}^{a_{it}} - \Xi(v_{it}^{a_{it}}, K_{it}, x_t, z_{it})$ with the decision of paying or not paying dividends each period as well as an investment decision.[10] Furthermore, we define v_{it}^a as

$$v_{it}^a \equiv \Pi(K_{it}, z_{it}, x_t) - D_{it} - \Phi(I_{it}^a, K_{it}),$$

and

$$\Phi(I, K) \equiv \frac{\phi}{2}\left(\frac{I}{K}\right)^2 K + I,$$

in which Φ is a quadratic investment adjustment cost with a constant $\phi > 0$.

Regarding the cash holding cost, $\Xi(v_{it}^{a_{it}}, K_{it}, x_t, z_{it})$, we assume the following quadratic form:

$$\Xi(v, K, x, z) \equiv \frac{\xi_1 e^{\xi_2(x+z)}}{2}\left(\frac{v}{K}\right)^2 K,$$

in which $\xi_1, \xi_2 > 0$ are constant parameters. This reflects the idea that shareholders would expect more dividends to be paid in good times, and increasing the cash holdings of firms becomes more costly. The law of motion for capital is expressed as

$$K_{it+1} = I_{it}^a + (1 - \bar{\delta}) K_{it},$$

in which $\bar{\delta}$ is a constant depreciation rate of capital stocks. This setup allows that

on this issue.

[9] The superscript a is sometimes suppressed, when it is not confusing, for the purpose of expositional ease. For instance, when we integrate over i, we will simply use v_{it} instead of v_{it}^a.

[10] We can define a felicity function for firm manager, F such that her periodic utility is $F(v_{it}^a - \Xi(v_{it}^a, K_{it}, x_t, z_{it}))$, where $F' > 0$, and $F'' \leq 0$. In this context, our setup assumes that firm managers are risk-neutral.

capital in the next period may depend upon dividend policy, especially around the time of initiating dividend payment, since whether a firm pays a proportion of cash flow from the profit as dividend to a shareholder determines the amount of investment given the investment opportunity set.

Recall that firm managers are subject to costs due to the agency problem mentioned above. This not only has a direct implication on the payout policy of firms, but also affects firms' capital accumulation paths over time. Furthermore, note that the total cost defined as $\Phi + \Xi$ is state-dependent, depending on the firms' choices of extensive margins on the dividend payout. Specifically, even if each of Φ and Ξ is a convex function, total cost function is non-convex in investment. Figure 4 displays this feature. Notice that the level of investment corresponding to the minimum cost of non dividend payment, $(\Phi^0 + \Xi^0)$ is bigger than that of the minimum cost when paying dividend, $(\Phi^1 + \Xi^1)$. This property results from the fact that dividend payment effectively reduces the range of possible investments by lowering down cash flow in hand. In addition, the total cost function shifts to the right as capital increases. Once a firm has enough cash in hand and fewer marginal profitable investment options, it has the incentive to pay dividends to reduce the cash holding cost as well as the investment adjustment cost.

The recursive problem of firm (2) is well defined, and the existence of a solution can be easily established, summarized in the following proposition.

Proposition 2.1.1. *There exists a solution to the functional equation* (2).

Proof. We refer to Theorem 9.6 in Stokey and Lucas (1989).

Now we characterize the firm dynamics implied by our model. In light of optimal dividend policy, the probability of paying dividends is described as

$$\mathbb{P}\{a = 1|S\} = 1$$

if and only if $V^1 > V^0$ given the state vector $S \in \mathcal{S}$. In this context, the propensity to pay dividends (PPD, hereinafter) following Fama and French (2001) is defined as,

$$\mathbb{P}\{a_{t+u} = 1|S_t\} \equiv \int_{\mathcal{S}} a_{t+u} dQ\left(S_{t+u}|S_t\right), \tag{3}$$

in which dQ is the transition density of S. This PPD measure is reminiscent of a hazard function in that a firm usually does not pay dividends at its initiation, yet it tends to pay dividends as it matures. In addition, (3) states that capital, aggregate shock, and firm-specific shock (i.e., the elements of S_t) determine optimal dividend policy, and can be related to the size and the profitability of firms. Although we will investigate this issue in detail through numerical analysis, we show some theoretical predictions consistent with the empirical findings by Fama and French (2001) below. For the purpose of illustration, we make a simplifying assumption that there are some dividend payers and non-payers who will not change their types for this and next sections.[11]

Assumption 2.1.2. There exist compact sets $\mathcal{S}^0$ and $\mathcal{S}^1$ such that for bounded positive integer sets $\{u^0\}$ and $\{u^1\}$,

$$\mathcal{S}\left(u^0\right) \equiv \left\{S_t \in \mathcal{S} \middle| \; \mathbb{P}\left\{a_{t+s} = 0 | S_t\right\} = 1, \; \forall 0 \leq s \leq u^0\right\}, \text{ and } \bigcup_{\{u^0\}} \mathcal{S}\left(u^0\right) = \mathcal{S}^0,$$

and

$$\mathcal{S}\left(u^1\right) \equiv \left\{S_t \in \mathcal{S} \middle| \; \mathbb{P}\left\{a_{t+s} = 1 | S_t\right\} = 1, \; \forall 0 \leq s \leq u^1\right\}, \text{ and } \bigcup_{\{u^1\}} \mathcal{S}\left(u^1\right) = \mathcal{S}^1.$$

Assumption 2.1.2 describes that if a firm's state vector is located in $\mathcal{S}^0$, then it currently does not pay dividends and it has no propensity to pay dividends to shareholders in a near future. But, in $\mathcal{S}^1$, a firm is paying a proportion of operating profit as dividends for some periods. Then, we show the following.

Proposition 2.1.3. *Suppose that $\phi = 1$ in the investment adjustment cost function for simplicity, for $S \in \mathcal{S}^0$ and $\frac{\Pi}{3} < I^0 \leq \Pi$. Under the technical Assumption A.1.1 in Appendix A,*

$$\frac{\partial I^0(S)}{\partial K} > 0, \; \frac{\partial I^0(S)}{\partial x} > 0, \; \textit{and} \; \frac{\partial I^0(S)}{\partial z} > 0,$$

in which I^0 is the optimal investment when a firm does not pay dividends.

Proof. See Appendix A.

According to Proposition 2.1.3, the optimal investment increases in capital, aggregate shock, and firm-specific shock, when a firm does not pay dividends. If a firm starts to pay out, we find that this relationship becomes non-monotonic and is unable to show. The complexity comes from the shift in conditional likelihood of initiating dividend payment as either productivity shock or level of capital varies. For instance, if the aggregate productivity shock x_t increases, then investment increases when there

[11] This exogeneity is not used in our main quantitative section.

is no dividend payment. But, the increases in firm profitability can also increase the value of firm with or without dividend payment. Note that increases in productivity can give more pressure to the firm manager toward paying dividends via the agency cost channel. Then, the conditional probability of dividend payment can increase as x or z goes up. This is illustrated in Figure 5. As x increases, the level of capital that triggers dividend payment shifts leftward, implying that dividend is going to be distributed at an earlier stage of firm maturity.

We expect that the optimal investment experiences a one-time reduction as dividend starts getting paid, followed by a gradual increase of investment in capital. This conjecture is confirmed in our numerical and empirical analysis (Section 4.2).

2.2 Preferences and Asset Prices

We assume that there is a representative household in this economy. The household holds a continuum of stocks from a set of firms. This setting is borrowed from Lucas tree model (Lucas (1978), Cochrane, Longstaff, and Santa-Clara (2008), and Martin (2009)). The preference of this household is given by

$$\mathbb{E}_t \left[\sum_{u=0}^{\infty} \beta^u \frac{C_{t+u}^{1-\gamma}}{1-\gamma} \right],$$

in which $\beta \in (0,1)$, and $\gamma > 1$. The budget constraint of the representative household is written as

$$C_t + \int P_{it}\varphi_{it+1} di = \int \varphi_{it} \left(P_{it} + D_{it} \right) di, \tag{4}$$

where φ_{it} is an outstanding share of stock i publicly traded from firm i in the stock market, P_{it} is the price of asset of firm i at time t, and D_{it} is the dividend from firm i at time t. We assume that φ_{it} for all i and t is equal to one, hence

$$C_t = \int D_{it} di. \tag{5}$$

This is a Lucas tree model with multiple assets. But, it is important to note that the dividend processes of firms, $\{(D_i)_t\}_{i\in[0,1]}$ are endogenously determined by the firm problem we developed previously, in lieu of using exogenous processes for endowments. In addition, our setup incorporates financial frictions as well as real frictions, allowing interactions between the two main pillars in corporate decision making.

Given the simplicity of the household setup, we can readily compute the fundamental asset pricing equation through the Euler equation of the representative household as

$$P_{it} = \mathbb{E}_t \left[\sum_{u=1}^{\infty} M_{t,t+u} D_{it+u} \right], \tag{6}$$

in which $M_{t,t+u}$ is the stochastic pricing kernel such that,

$$M_{t,t+u} = \beta^u \left(\frac{C_{t+u}}{C_t} \right)^{-\gamma} = \beta^u \left(\frac{\int D_{it+u} di}{\int D_{it} di} \right)^{-\gamma}.$$

Cochrane, Longstaff, and Santa-Clara (2008) show that the volatility of consumption growth is endogenously related to a non-linear function of dividend shares (D_{it}/C_t), which leads to time-variations in risk premium. Our model shares this feature, and endogenizes the dividend processes of individual firms, and aggregate dividend to explicitly link economic factors to asset returns, without relying on some ad-hoc statistical processes.

Recently, there are several papers studying asset returns via variables related to the investment and production sides of firms (See Zhang (2005), Li, Livdan, and Zhang (2009), and Livdan, Sapriza, and Zhang (2009)). However, we depart from this literature by emphasizing the role of endogenous dividend policy to explain the cross-sectional behaviors of asset returns. In addition, they assume a version of stochastic risk aversion to generate time-varying risk premium unlike this model in which the risk premium is endogenously countercyclical.

Expected returns, market sizes, book-to-market equity ratios and some sort of risks are functions of three state variables $K_{it,}$ z_{it} and x_t. The risk and expected return of firm i satisfy

$$\mathbb{E}_t \left[R_{it+1} \right] = R_{ft} + \beta_{it} \lambda_{Mt} \text{ or } \mathbb{E}_t \left[r_{it+1} \right] - r_{ft} = \beta_{it} \lambda_{Mt}.$$

The *quantity of risk* is given by

$$\beta_{it} \equiv - \frac{\mathbb{C}ov_t \left[M_{t,t+1}, R_{it+1} \right]}{\mathbb{V}ar_t \left[M_{t,t+1} \right]},$$

the *price of risk* is given by

$$\lambda_{Mt} \equiv \frac{\mathbb{V}ar_t \left[M_{t,t+1} \right]}{\mathbb{E}_t \left[M_{t,t+1} \right]}.$$

and the maximum conditional *Sharpe* ratio is given by

$$\left|\frac{\mathbb{E}_t\left[R_{t+1}\right]-R_{ft}}{\sigma_t\left[R_{t+1}\right]}\right| \leq \frac{\sigma_t\left[M_{t,t+1}\right]}{\mathbb{E}_t\left[M_{t,t+1}\right]},$$

in which R is assumed to be located on the mean-standard deviation frontier (Cochrane (2005)). The market size is P_{it} since we already assume that a supplied outstanding share of stock i, φ_{it} is one. Then, *book-to-market equity ratio* is defined as

$$\frac{BE_{it}}{ME_{it}} \equiv \frac{K_{it}}{P_{it}}.$$

2.3 Equilibrium

By developing a recursive competitive equilibrium model, we characterize the aggregate behaviors of the economy. We assume that aggregate demand is automatically cleared at the aggregate output. Asset prices (P_{it}) are real prices normalized by the output price, and we can suppress the output price. Let μ_t denote the measure over the capital stocks and idiosyncratic shocks for all the firms at time t and let $\Psi\left(\mu_t, x_t, x_{t+1}\right)$ be the law of motion for the firm distribution μ_t. Then $\Psi\left(\mu_t, x_t, x_{t+1}\right)$ can be stated formally as

$$\mu_{t+1}\left(\mathcal{S}; x_{t+1}\right) = T\left(\mathcal{S}, \left(K_t, z_t\right); x_t\right) \mu_t\left(K_t, z_t; x_t\right), \tag{7}$$

in which the operator T is defined as

$$T\left(\mathcal{S}, K_t, z_t; x_t\right) \equiv \int\int 1_{\left\{I_t+\left(1-\bar{\delta}\right)K_t, z_{t+1}\in\mathcal{S}\right\}} dQ_z\left(z_{t+1}|z_t\right) dQ_x\left(x_{t+1}|x_t\right), \tag{8}$$

in which $1_{\{.\}}$ is the indicator function. The operator T determines the law of motion of the firm distribution μ_t.

The total economic output can be written as

$$Y_t \equiv \int y\left(K_t, z_t; x_t\right) d\mu_t\left(K_t, z_t\right),$$

in equation (9). The resource constraint for this economy is given by

$$Y_t = \int \left(D_{it} + v_{it} + \Phi_{it}\right) di. \tag{9}$$

Definition 2.3.1. A recursive competitive equilibrium is characterized by (a) an optimal investment rule $I^*\left(S_t\right)$, an optimal dividend policy $D^*\left(S_t\right)$ as well as a value function $V^*(S_t)$ for each firm, (b) an optimal consumption $C^*(\mu_t, x_t)$ for a representative household given asset prices $P(S_t)$, and (c) a law of motion of firm distribution Ψ^* such that:

1. $I^*(S_t)$ and $D^*(S_t)$, hence $v^*(S_t)$ solve the value-maximization problem (2) for each firm,

2. $C^*(\mu_t, x_t)$ solves the household utility-maximization problem,

3. $P(S_t)$ is determined in (6),

4. Consistency: (9) holds for the consistency of the production of all firms in the industry with the aggregate output Y_t. (7) and (8) hold for the consistency of the law of motion of firm distribution Ψ^* with firms' optimal decisions.

5. Market clearing condition: from (5) and (9),

$$Y_t = C_t + \int (v_{it} + \Phi_{it})\, di,$$

   and $\varphi_{it} = 1$ for all i and t.

Proposition 2.3.2. *There exists a unique recursive general equilibrium.*

Proof. See Appendix A.

3 Cross Sections of Stock Returns

Since our model incorporates key aspects of corporate decision making on capital accumulation and dividend payout which, in turn, determine the wealth of investors in equilibrium, the model offers a theoretical laboratory to analyze the stylized facts from the empirical asset pricing studies, such as (Fama and French (1992)). In particular, we focus on the value and size premia. The first step is to illustrate the theoretical possibility via simple examples, paving the way to the quantitative analysis.

3.1 The Value Premium

The expected excess return is decomposed into a term indicating the discount risk effect and the other term representing the cash flow risk effect following (Santos and Veronesi (2009)) as

$$\mathbb{E}_t[R_{it+1}] - R_{ft} = \left(\beta_{it}^{\text{cf}} + \beta_{it}^{\text{disc}}\right)\lambda_{Mt}, \tag{10}$$

in which

$$\beta_{it}^{\text{cf}} = -\frac{\mathbb{C}ov_t\left[M_{t,t+1}, \frac{D_{it+1}}{P_{it}}\right]}{\mathbb{V}ar_t\left[M_{t,t+1}\right]},$$
$$\beta_{it}^{\text{disc}} = -\frac{\mathbb{C}ov_t\left[M_{t,t+1}, \frac{P_{it+1}}{P_{it}}\right]}{\mathbb{V}ar_t\left[M_{t,t+1}\right]}.$$

Alternatively, the expected stock return can be explained in two parts, the expected rate of capital gain (expected long-term dividend growth rate) and the expected dividend price ratio (Fama and French (2002), and Chen, Petkova, and Zhang (2008)) such that,

$$\mathbb{E}_t\left[R_{it+1}\right] = \frac{\mathbb{E}_t\left[P_{it+1}\right]}{P_{it}} + \frac{\mathbb{E}_t\left[D_{it+1}\right]}{P_{it}}. \tag{11}$$

Now we show that the value premium can result from both the discount risk and cash flow risk under some conditions.

Proposition 3.1.1. *Suppose that there are only two assets in the economy with $K_{1t} < K_{2t}$, $P_{1t} = P_{2t}$ with the Assumptions 2.1.2 and A.1.2. Further, we assume that*

$$(K_{1t}, x_t, z_{1t}) \in \mathcal{S}^0 \text{ and } (K_{2t}, x_t, z_{2t}) \in \mathcal{S}^1.$$

Then, the following relationships hold.

$$\beta_{1t}^{cf} < \beta_{2t}^{cf},\ \beta_{1t}^{disc} < \beta_{2t}^{disc},$$

hence,

$$\mathbb{E}_t\left[R_{1t+1}\right] < \mathbb{E}_t\left[R_{2t+1}\right].$$

Proof. See Appendix A.

The firm 1 with small capital is a growth stock compared to the firm 2 with larger capital, a value firm by construction. Then the proposition states that the value premium prevails, as long as the growth firm has not been paying the dividend ($\mathcal{S}^0$), while the value firm pays ($\mathcal{S}^1$). It is worth mentioning that both the premiums from the cash flow and discount risks are higher for the value firm. Admittedly the assumptions used in this proposition are rather strong, which we do not impose in our quantitative analysis. However, this permits us to gain insight on the roles of lumpy dividend policy and the life cycle of firms in producing the value premium.

In this example, the cash flow beta (β_t^{cf}) of the value firm is higher than that of the growth firm, because the latter will not distribute its cash to a shareholder for a while, and the resulting conditional covariations will be zero. Thus, as long as it is a reasonable assumption that non-dividend paying firms are the growth firms, which will be shown in the next section, equity holders of those firms are exposed to lower cash flow risks, since they are unlikely to pay dividends in a near future.

Interestingly, this intuition carries over in evaluating the discount effect as well. According to the proposition, the equity of the firm 2 (value firm) will involve the higher risk than that of the firm 1 (growth firm) for betting on long-term dividend growths ($\beta_{1t}^{\text{disc}} < \beta_{2t}^{\text{disc}}$). At first glance, this result seems counterintuitive, because growth firms are assets with high durations, implying that the more sensitive are their prices to changes in the discount factor. It is well known that this leads to counterfactual growth premiums. Lettau and Wachter (2007) assume that the discount rate shock is uncorrelated with the aggregate dividend growth to turn this channel off, while Santos and Veronesi (2009) increase idiosyncratic cash flow risks to counter the effect. Although the same channel still exists in the model, there is another discount effect from the short-run fluctuations in our case: Stocks currently paying dividends (value firms) can be exposed to more discount risk than those not paying dividends (growth firms) which has zero covariations with changes in the discount rate until they start paying. If this dominates the effect of high duration, the value firm involves higher risks from the shocks to discount rate.

More concretely, under the assumptions in Proposition 3.1.1, we have the following relation

$$\beta_2^{\text{disc}} - \beta_1^{\text{disc}} \propto -\sum_{u=1}^{u^0-1} \mathbb{C}ov_t \left[M_{t,t+1}, (\mathbb{E}_{t+1} - \mathbb{E}_t) \left(\left(\frac{D_{2t+u+1}}{D_{2t+1}} \right)^{-\gamma} D_{2t+u+1} \right) \right], \tag{12}$$

where

$$\mathbb{C}ov_t \left[M_{t,t+1}, (\mathbb{E}_{t+1} - \mathbb{E}_t) \left[\left(\frac{D_{2t+u+1}}{D_{2t+1}} \right)^{-\gamma} D_{2t+u+1} \right] \right] < 0 \tag{13}$$

for $u = 0, 1, \cdots u^0$. The term (13) exhibits the discount risk effect of the value firm (firm 2) during a period in which the growth firm does not pay dividends. If the differences in future dividends are relatively small by the time both firms pay dividends (Assumption A.1.2), value premium can prevail because of the differences in dividend policy in the next upcoming periods. Assumption A.1.2 is a quantitative concern, which we closely examine in the next section. Notice from (12) that the value premium increases in u^0 that measures how long the firm 1 will delay dividend payment, i.e., persistence in dividend policy. As the growth firm delays paying dividends for a longer period, the value premium is likely to be higher. In sum, a firm's decisions on whether to pay dividend and its duration matter to generate the value premium in our model. Of course, the assumption on the equal market value of the two firms ($P_1 = P_2$) completely ties the firm's book value (K) with the book-to-market equity ratio (K/P). However, the following proposition shows that if firms with larger (smaller) book values (K) are those who (do not) pay dividends, their book-to-market equity ratios are consistent with the rank order of book values.

Proposition 3.1.2. *Suppose that there are only two assets without firm-specific shocks in the economy with $K_{1t} < K_{2t}$, the Assumption 2.1.2, and some parametric restrictions. If $(K_{1t}, x_t) \in \mathcal{S}^0$ and $(K_{2t}, x_t) \in \mathcal{S}^1$ additionally hold, then*

$$\frac{K_{1t}}{P_{1t}} < \frac{K_{2t}}{P_{2t}}.$$

Proof. See Appendix A.

The existing models such as Gomes, Kogan, and Zhang (2003) and Zhang (2005) have the property that a growth firm pays more dividend than a value firm, which is counterfactual. On the contrary, the above proposition suggests that, even without firm-specific shocks, the firm paying dividends (firm 2) is more likely to be a value stock compared to the firm 1 who is not paying dividends. Because the firm 1 has relatively higher marginal profitability and a larger opportunity set, its value of growth options given capital is greater than that of firm 2, inferred from the firm dynamics.

3.2 The Size Premium

This section explores the possibility of the size premium via dividend policy using a simple Gordon model. We assume that there exist only two firms with $K_{1t} < K_{2t}$ indexed by 1 and 2 in the economy. We further assume that $D_{1t} = \delta K_{1t}^{\theta}$ and $D_{2t} = \delta K_{2t}^{\theta}$, if dividends are distributed. Similar to the previous examples, for simplicity, we assume that the firm 1 decides not to pay dividends until $t+u^0+1$, and the firm 2 pays dividends now. In addition, firms continue to pay dividends once it started. Finally, to control for the value premium, two firms are assumed to have the same book-to-market equity ratio, i.e.,

$$\frac{K_{1t}}{P_{1t}} = \frac{K_{2t}}{P_{2t}}.$$

The firm 1 (2), by construction, is a small (large) firm in this simple example.[12]

[12] Because the marginal profitability of firm 1 is higher than the firm 2, this exogenous assumption that small firms postpone the initiation of dividends used in this section is consistent with our firm dynamics in section 2.

Stock prices of the firms are then written as

$$
\begin{aligned}
P_{1t} &= \sum_{u=u^0+1}^{\infty} \delta K_{1t}^{\theta} \left(\frac{1+\Delta P_1/P_1}{1+r} \right)^u = \frac{(1+\Delta P_1/P_1)^{u^0+1} \delta K_{1t}^{\theta}}{(1+r)^{u^0} (r - \Delta P_1/P_1)}, \\
P_{2t} &= \sum_{u=1}^{\infty} \delta K_{2t}^{\theta} \left(\frac{1+\Delta P_2/P_2}{1+r} \right)^u = \frac{(1+\Delta P_2/P_2) \delta K_{2t}^{\theta}}{(r - \Delta P_2/P_2)},
\end{aligned}
$$

in which $\Delta P_1/P_1$ and $\Delta P_2/P_2$ are long-term dividend growth rates in Gordon model, and r is a constant opportunity cost of capital in the financial market. $\Delta P_1/P_1$ and $\Delta P_2/P_2$ are less than r to have finite stock prices. We then solve for the difference of two long-term dividend growth rates under the assumption of equal book-to-market equity ratios to show that

$$
\frac{\Delta P_1}{P_1} - \frac{\Delta P_2}{P_2} \propto \left(\frac{K_{2t}}{K_{1t}} \right)^{\frac{\theta-1}{u^0}} .
$$

$(K_{2t}/K_{1t})^{\theta-1}$ is the ratio of marginal profitability of firm 2 to marginal profitability of firm 1, which must be less than 1 according to the assumption. The difference between two long-term dividend growth rates has the positive relation to u^0, which presents the firm 1's dividend policy. This property implies that in the stock market, the expected long-term dividend growth rate of the firm 1 is higher than the firm 2, even though the firm 1 has the same book-to-market equity ratio as the firm 2. While the firm 1 starts to pay dividends later, the condition of equal book-to-market equity ratios makes its long-term dividend growth rate expected to be higher. Plus, if the firm 1 has the smaller capital, it will invest more in capital, postpone further the initiation time of dividend payout by due to the increases in its marginal profitability. Consequently, the expected capital gain of firm 1 must be higher than before. If the firm 2's expected dividend price ratio is sufficiently low, then the firm 1's expected stock return must be higher than the firm 2, implying the small firm's size premium. One caveat of this explanation is the lack of risk adjustment, and further quantitative analysis is desired.

To summarize, time-varying, persistent, and stochastic extensive margins in dividend policy have potentials to explain cross-sectional variations in stock returns. However, as emphasized in the beginning of this section, dividend policy is a highly endogenous process depending on the fundamental and firm specific shocks. To delve into this issue, we now turn our attention back to the full model, and quantitatively analyze it.

4 Quantitative Analysis

4.1 Calibration

We calibrate 12 parameters $\left(\bar{\delta}, \rho_{x,}\sigma_{x,}\rho_z, \sigma_z, \phi, \delta, \beta, \mu_x, \theta, \xi_1, \xi_2\right)$ both in monthly and quarterly frequencies to facilitate the comparison of our results with those from the empirical literature. The monthly model is denoted as Model $\mathcal{M}$, and the quarterly model, as Model $\mathcal{Q}$. The first parameter set $\left(\bar{\delta}, \rho_{x,}\sigma_{x,}\rho_z, \sigma_z, \phi, \delta, \beta\right)$ is calibrated outside the models, and the second parameter set $(\mu_x, \theta, \xi_1, \xi_2)$ is calibrated inside the models following the idea of generalized method of moments. Table 1 reports the parameter values that we use to solve and simulate the models.

The monthly depreciation rate $\bar{\delta}$ is 1%, which implies an annual rate of 12% (Abel and Eberly (2002)). The persistence of the aggregate productivity process (ρ_x) is 0.95 and its conditional volatility (σ_x) is 0.007, which are quarterly values. 0.983 and 0.0023 are the monthly values of those parameters, ρ_x and σ_x, respectively. These values are consistent with Cooley and Prescott (1995). For the persistence (ρ_z) and conditional volatility (σ_z) of the firm-specific productivity shock, we set $\rho_z = 0.97$ and $\sigma_z = 0.10$ for monthly frequency. These values are chosen from the related literature (e.g., Zhang (2005), Li, Livdan, and Zhang (2009), and Livdan, Sapriza, and Zhang (2009)) to generate a plausible amount of dispersion in the cross-sectional distribution of firms. We set $\rho_z = 0.912$ and $\sigma_z = 0.30$ for quarterly frequency. The average payout ratio (δ) is calibrated to be 60%, borrowed from the average value of aggregate payout ratios from 1981 to 2007 in CRSP/COMPUSTAT merged data. We choose the time preference parameter, $\beta = 0.998$ on monthly and $\beta = 0.994$ on quarterly.

Regarding the second set of parameters $(\mu_x, \theta, \xi_1, \xi_2)$, we calibrate those parameters using the following procedure.[13]

1. Set initial values for group $\Lambda^0 = \left(\mu_x^0, \theta^0, \xi_1^0, \xi_2^0\right)$.

2. Solve the value function $V\left(S_{it}\right)$ such that,

   (a) Each grid on the compact set $\mathcal{S}$ represents each firm in the market.

   (b) By backward induction, solve the $V^0\left(S_{it}\right)$ and $V^1\left(S_{it}\right)$ on each iteration, and pick the bigger value.

[13] The convexity parameter of investment adjustment cost (ϕ) is normalized to one to conserve the number of parameters used in our numerical study.

3. From the converged value function $V(S_{it})$, we have the operating profit function (profitability) and endogenous policy functions (investment-capital ratio, disinvestment-capital ratio, and dividend policy) such that,

$$Z_{it} = \left(\frac{\Pi(S_{it})}{K_{it}}, \frac{I_{+}(S_{it})}{K_{it}}, \frac{I_{-}(S_{it})}{K_{it}}, a_{it}; \Lambda \right)',$$

   in which, I_{+} is the net investment and I_{-} is the net disinvestment.

4. Define the generalized moments of policy functions as,

$$\frac{1}{n} \sum_{i=1}^{n} Z_{it} - \mathbb{E}[Z_{it}] = \bar{\mathcal{Z}}(\Lambda) - \mathcal{Z},$$

   in which n is $n_K \times n_x \times n_z$, ($n_K$: number of grids on $\mathcal{S}_K$, n_x: number of states of x, and n_z: number of states of z).

5. By iterating the process from step 1 to step 4, we solve the minimization problem for the parameters as follows,

$$\hat{\Lambda} = \arg\min_{\Lambda} \left(\bar{\mathcal{Z}}(\Lambda) - \mathcal{Z} \right)' W \left(\bar{\mathcal{Z}}(\Lambda) - \mathcal{Z} \right),$$

   in which W is a weight matrix, and we use identity matrix in this study.

Our method is not the same as the conventional simulated method of moments since we do not generate random numbers to compute the moments of endogenous policy functions. Rather, generated moments are the average numbers of target values from points on all grids of the compact state space while iterating value functions regarding the values of Z. The average monthly profitability defined as the operating profit to the capital ratio (Π/K) is 1.25%, the value of the average monthly net investment ratio (I_{+}/K) is 1.25%, for the net disinvestment (I_{-}/K), 0.17% is the average value on monthly frequency. These values are reported by Abel and Eberly (2002). The average proportion of firms in CRSP that paid dividends in a period from 1926 to 1999 is 49%. This number is borrowed from Fama and French (2001). The calibrated parameters are $\Lambda = (-2.0, 0.68, 323, 0.084)$ for Model $\mathcal{M}$ and $(-1.71, 0.65, 458, 0.044)$ for Model Model $\mathcal{Q}$. The curvature parameter in the production function (θ) is monthly 0.68 and quarterly 0.65, close to the values suggested by Livdan, Sapriza, and Zhang (2009) or the average values estimated by Cooper and Ejarque (2001), Cooper and Ejarque (2003), Hennessy and Whited (2005), and Hennessy and Whited (2007). The long-run average level of aggregate productivity (μ_x) is -2.0 for Model $\mathcal{M}$ and -1.71 for Model $\mathcal{Q}$, which are higher than other free cash flow asset pricing models (e.g., Zhang (2005), Li, Livdan, and Zhang (2009), and Livdan, Sapriza, and Zhang (2009)). The difference may come from the way this parameter is retrieved. They calibrate μ_x and time-varying γ exogenously by fitting the first and second moments of risk-free rate data. Meanwhile, we use only firm characteristics data to calibrate it. Convexity parameter of the cash holding cost (ξ_1) and procyclical parameter of cash holding cost (ξ_2) are 323, 458, and 0.084, 0.044 on monthly and quarterly frequencies, respectively.

4.2 Investment, Dividend Policy, and Cash Holdings

Panel (A)'s in Figures 6-9 display optimal investment behaviors as a function of capital and productivity shocks (K, x, z). First, we observe that optimal investments increase until capital stock (K) reaches level at which dividend payouts begin. When this occurs, investment drops off by about the half of where it used to be. As discussed, x and z determine the amount of investments, and they have positive relationships with investment.

Panel (B)'s in Figures 6-9 show that the investment-capital ratios (I/K) given x and z decrease as capital stock increases, because of diminishing returns to scale of capital. This is consistent with the view that young firms are more likely to invest their resources in capital compared to mature firms. What is new in our model is that a firm will execute a one-time discrete reduction of investment in its life cycle, as it grows. The discrete shift in investment results from the lumpy behavior of dividend policy, as we discussed earlier, because dividend payment substitutes for the amount of investment. However, we want to emphasize that this does not necessarily imply that investment and dividend payout will feature negative correlation, since adjustment is costly for both investment and dividend and therefore, frequent changes of these variables are unlikely to prevail. Only a few observations show conditionally negative correlations, and the unconditional correlation between investment and dividend does not need to be negative, because it can depend more on firm profitability, level of capital, and related, firm age.

To verify if our policy function is a reasonable description of reality, Panels (A) and (B) in Figure 10 plot the investment-book equity ratios of all firms in SIC 3000-3099, sorted in terms of the book equity values[14]. One can see the clear resemblance of this

[14] We compare our numerical solutions with all patterns of dividend policies and investments of firms within 73 industries categorized by the first two digits of SIC code. It is confirmed that most of industries have the similar patterns to our model solutions for firm dynamics. For example, we plot

figure to the panel (B) of Figures 6-9. Firms with the smaller amount of capital invest more, and the investment decreases rather abruptly as capital is accumulated. Panel (C)'s in Figures 6-9 show the optimal dividend policies. Zero dividend equity ratios mean that firms are not paying dividends, and positive numbers show that the firms pay dividends. Firms with higher x and z are more likely to initiate their dividend payouts earlier, since those firms will accumulate capital faster, hence they tend to mature earlier than other firms. Recall that our model has the cash holding cost motivated in part by the agency cost resulting from the tension between equity holders and the firm manager. Thus, accumulated capital will lead to larger operating incomes which trigger dividend payout. Panels (C) and (D) in Figure 10 illustrate this lumpiness in dividend policy associated with the life cycle of firms. The pattern shown in this figure clearly confirms our theoretical prediction depicted in the panel (C)'s of Figures 6-9.

It is natural to think that the propensity to pay dividends (PPD), $\mathbb{P}\{a=1|S\}$ and firm-characteristics such as profitability, investment, and the market value of capital are associated with each other. Figures 6-9 suggest that our model indeed replicates the stylized facts in Fama and French (2001) as mentioned earlier: Asset prices generated by the model are consistent with the firm characteristics.

Panel (D)'s in Figures 6-9 plot the value functions of firm dynamics (2) which are the present values of cash flows into cash holdings. These figures show that as K, x, and z increase, a young firm's present value of cash flows increases, but the value of a mature firm which pays dividends decreases. For a young firm, the increases in x and z make its managerial cost increase due to the procyclical cash holding cost, and force the young firm to invest more cash in capital. Then, this procyclical effect expedites the young firm's initiation of dividend payouts earlier than when it has lower x and z. For a mature firm, although it also shares this procyclicality channel, it will face a relatively limited investment opportunity set due to its large capital. Thus, the mature firms are bound to solve a more constrained optimization to minimize the total managerial cost. This renders the mature firms to incur higher cash holding cost than when it was young. This also makes them to pay more dividends to reduce the cash holding cost. Thus, the present values of cash flows for the mature firms will become smaller than those of younger firms, and decrease in x and z, ceteris paribus.

---

dividend policies and investments of firms within SIC 3000-3099.

4.3 Analyzing Stock Returns

Now, to analyze the expected market return and cross sections of expected stock returns, we simulate 200 artificial panels, each of which has 2703 firms on each state x. Model $\mathcal{M}$ simulates average 2100 months to generate one panel, and Model $\mathcal{Q}$ simulates average 1200 quarters on each panel. The cash flows are discounted by the computed pricing kernels and summed up to P_{it}^{j} in each panel j, where $j = 1, ...200$, until all P_{it}^{j}'s converge, following the pricing formula (6). Then, the final stock price of firm i at time t, P_{it}, is computed by the average value of $\{P_{it}^{j}\}_{j=1}^{200}$. Finally, the expected stock returns are computed using Markov transition matrices of x and z approximated by the method in Adda and Cooper (2003). We calculate the expected value-weighted stock returns on each state, x, and consider them as the market portfolio returns, or the wealth portfolio returns. In addition, the risk-free rate is computed by the reciprocal of the average value of $M_{t,t+1}$ using the simulated data.

4.3.1 Unconditional Moments of Aggregate Market Values

Table 2 reports the unconditional moments of market variables such as equity premium, risk-free rate, price-dividend ratio, book-to-market equity ratio, aggregate dividend growth, and the volatilities of those variables as well. The U.S. historical data are collected from various sources. The average equity premium ranges from 4% to 8% according to the related literature. The volatility of market return is 19.4%, according to Guvenen (2009), computed using Standard and Poor's 500 index during the period of 1890-1991. The empirical Sharpe ratio is 0.50, which is from Cochrane (2005). The fitted model generates reasonable values for the expected equity premiums compared to the data. Model $\mathcal{M}$ with $\gamma = 3$ and Model $\mathcal{Q}$ with $\gamma = 3.5$ have the volatilities of stock market returns of 20.1% and 20.9%, which is close to 19.4%. The cross-sectional volatility of individual stock returns is from 25% to 32% reported by Zhang (2005). Model $\mathcal{M}$ reports higher volatilities than this, but Model $\mathcal{Q}$ reports somewhat lower values. The actual rates of capital gain and dividend price ratio are 2.1% and 4.70% from Fama and French (2002) covering from 1872 to 2000. This means that the dividend price ratio is about 123% larger than the capital gain in constituting the market equity return. The dividend price to capital gain ratios $(D_{t+1}/\Delta P_{t+1})$ from all models are around 2.5, which is close to the empirical value. Turning to the risk-free rate, annualized U.S. risk-free rate is 1.8% with volatility of 3.0% according to Zhang (2005). The corresponding risk-free rates in Model $\mathcal{M}$ are relatively higher than data. The historical standard deviation of risk-free rate is more volatile compared to the simulated values in Model $\mathcal{M}$. Model $\mathcal{Q}$ creates more volatile risk-free rates than Model $\mathcal{M}$. However, there are several other studies reporting that the risk-free rate has much lower volatility, and therefore, we believe that our result on the risk-free rate is reasonable.

Regarding the Sharpe ratio, the simulated Sharpe ratios have a wide range from 0.23 to 0.59 in Model $\mathcal{M}$, and the range from 0.27 to 0.54 in Model $\mathcal{Q}$. When relative risk aversion is between 3 and 4, our results are consistent with the empirical counterpart. Model $\mathcal{M}$ generates aggregate dividend growths around 2.40%, with the aggregate dividend growth volatilities at about 11.4%. Model $\mathcal{Q}$ generates somewhat higher aggregate dividend growth volatilities of around 13%, which is close to the data, 13.4%, reported by Guvenen (2009), and the aggregate dividend growths are around 2.56%, which is fairly close to the empirical value, 2.5%. The average book-to-market equity ratio (BE/ME) is 0.67 and its standard deviation is 0.23 according to Zhang (2005). Model $\mathcal{Q}$ represents values of 0.62 and 0.24 respectively, which are again fairly close to the data. We also compute $(BE/ME)^{\text{payer}} / (BE/ME)^{\text{non-payer}}$ which is the ratio of average book-to-market equity ratios of dividend payers and non-payers. The empirical value is 1.32, computed using the CRSP/COMPUSTAT merged data. This value states that high book-to-market firms are more likely to pay dividends (Smith and Watts (1992) and Baker and Wurgler (2004)), and it is consistent with our theory for the value premium. All of our simulated values are greater than 1, and Model $\mathcal{Q}$ has values close to the data. For the average price-dividend ratio, as relative risk aversion increases, models yield lower rates of price dividend ratios, but higher volatilities of them. Results suggest that our model explains the data reasonably well with relative risk aversion around 3.

Thus, the model replicates the moments of key financial market variables in both monthly and quarterly versions. Note that all the parameters are fitted by matching the moments of variables related to firm characteristics, not financial variables. The model can capture the level and volatility of the historic equity premium, while keeping plausible values for the first two moments of the risk-free rates. However, we must mention that our model does not resolve the equity premium puzzle. Since our model uses the equilibrium condition that the aggregate consumption is entirely financed by the sum of dividend shares, the aggregate consumption volatility coincides with that of aggregate dividend. This makes consumption growth highly volatile compared to empirical evidence. To break this tight link between consumption and dividend, one can include labor income or other types of capital or source of income not directly traded in the market. Alternatively, modifying the preference function to generate a sufficient volatility size of the stochastic discount factor would be desired, following Santos and Veronesi (2009).

4.3.2 Conditional Moments of Risk Premiums and Betas

Cochrane, Longstaff, and Santa-Clara (2008) and Martin (2009) extend Lucas (1978) to set the aggregate endowment process as the sum of multiple trees exogenously given, and show that the implied aggregate dividend (or consumption) volatility is time-varying, and dependent on the dividend share. The intuition behind this result is that the idiosyncratic shocks of each dividend process (tree) will induce dividend shares to fluctuate via the binding equilibrium condition that consumption equals the sum of these dividends. Then, the conditional moments of the aggregate consumption and dividend growth will vary over time, as the relative contribution of each tree to the aggregate process changes over time.

The model also has the feature of multiple dividend processes with the binding equilibrium condition, and therefore, time-varying and stochastic volatilities of the aggregate dividend can prevail as dividend shares vary. However, this does not provide much economic insight regarding the nature of this time-variability because dividend processes are exogenously given. In contrast, endogenous evolution of dividend processes, especially from the infrequent adjustments, arises in the model and this has additional implications on the time-series behaviors of the conditional moments on the related stochastic discount factor.

We illustrate the point using a simple version of our model developed in the previous section. Suppose that there exist only two firms indexed by 1 and 2 in the economy, and these firms have only an aggregate shock, x_t without firm-specific shocks, and zero depreciation rate. In addition, K_{1t} is assumed to be less than K_{2t} ($K_{1t} < K_{2t} \ll \bar{K}$). Then, the proposition 2.1.3 and the equilibrium condition state that the second firm must pay the dividend, while the first firm may or may not pay the dividend. That is, the condition $(a_1, a_2) \in \{(0,1),(1,1)\}$ holds in equilibrium and $(a_1, a_2) = (1,1)$ occurs only when x is sufficiently high. Therefore, we can view $(a_1, a_2) = (0,1)$ as the recession in this simple example. It is easy to show that the conditional mean and variance of the aggregate dividend growth vary over time.

$$\mathbb{E}_t\left[\left(\frac{D_{1t+1}+D_{2t+1}}{D_{1t}+D_{2t}}\right)\right] = \frac{e^{(1-\rho_x)(\mu_x - x_t)}\left\{\left(K_{1t}+I_{1t}\right)^{\theta}\mathbb{E}_t\left[e^{\epsilon_{xt+1}}a_{1t+1}\right] + \left(K_{2t}+I_{2t}\right)^{\theta} e^{\frac{\sigma_x^2}{2}}\right\}}{a_{1t}K_{1t}^{\theta}+K_{2t}^{\theta}},$$

$$\mathbb{V}ar\left[\left(\frac{D_{1t+1}+D_{2t+1}}{D_{1t}+D_{2t}}\right)\right]=\frac{e^{2\mu_x(1-\rho_x)+2(\rho-1)x_t}\left(K_{1t}+I_{1t}\right)^{2\theta}}{(a_{1t}K_{1t}^{\theta}+K_{2t}^{\theta})^2}\mathbb{V}ar\left[e^{\epsilon_{xt+1}}a_{1t+1}\right].$$

This implies that the maximum conditional Sharpe ratio varies over time given the definition of the stochastic discount factor. How do they change in response to changes in business condition?

Figures 11 and 12 plot the countercyclical business cycle patterns of risk premiums, Sharpe ratios, and the quantities of risks (β) for the full version of the model. The expected stochastic discount factor in Panel (A) $\mathbb{E}_t\left[M_{t,t+1}\right]$, increases in the aggregate shock by intertemporal substitution, which means that the shareholder's marginal utility for the future is low in good times and high in bad times. So, the risk-free rate is lower in good times, but higher in bad times, meaning that investors in good times are likely to save more than in bad times from the argument of consumption smoothing. Panels (B) and (C) numerically show that the volatility of the stochastic pricing kernel and the maximum conditional Sharpe ratio have the countercyclical pattern.

Panels (A) to (C) in Figure 12 depict that conditional expected market returns in Models $\mathcal{M}$ and $\mathcal{Q}$ have countercyclical variations over time, while Panels (B) and (D) in Figure 12 demonstrate the same relationship via the quantities of risk (β_M) in Models $\mathcal{M}$ and $\mathcal{Q}$.

4.3.3 Cross-Sectional Analysis of Stock Returns

Factor Portfolios and Size-BE/ME Portfolios To examine the value and size premia, we follow the Fama-French method: Pool the factor portfolios, SMB (small minus big) and VMG (value minus growth, previously HML), and the six value-weighted size$-BE/ME$ portfolios (Fama and French (2006)). Firms below the median size are defined as small (S) and those above are big (B). We assign firms to growth (G), neutral (N), and value (V) groups if their BE/ME is in the bottom 30%, middle 40%, or top 30% of simulated firms. The six portfolios, small and big growth (SG and BG), small and big neutral (SN and BN), and small and big value (SV and BV) are the intersections of these sorts. SMB is the simulated expected return on the three small-size stock portfolios minus the expected returns on the three big-size stock portfolios such that,

$$\mathrm{SMB}=\frac{\mathrm{SG}+\mathrm{SN}+\mathrm{SV}}{3}-\frac{\mathrm{BG}+\mathrm{BN}+\mathrm{BV}}{3},$$

and

$$\mathrm{VMG}=\frac{\mathrm{SV}+\mathrm{BV}}{2}-\frac{\mathrm{SG}+\mathrm{BG}}{2},$$

in which the value-growth factor, VMG is the difference between the expected returns on the two value portfolios and the two growth portfolios. The empirical values in Table 3 are from Fama and French (2007), and Chen, Petkova, and Zhang (2008). Fama and French (2007) analyze the financial market data from 1927 to 2006, and Chen, Petkova, and Zhang (2008) deal with the data from 1945 to 2005. In both studies, value and size premia prevail, and the average value premium is about twice bigger than the average size premium, and the capital gain effect outperforms the dividend price effect on both value and size premiums. On a more disaggregated level with six portfolios, Fama and French (2007) report that the expected capital gain (long-term dividend growth) has stronger effect than the expected dividend price, while Chen, Petkova, and Zhang (2008) show that the latter has stronger effect when they use the return with and without dividends. When they use annuity formula to measure long-term dividend growth, they find that the expected long-run dividend growth effect is bigger than the expected dividend-price effect in smaller stocks, and the latter is bigger in larger stocks.

Our model produces results broadly consistent with data and Model $\mathcal{Q}$ shows better performances. In case of Model $\mathcal{M}$, Table 3 says that VMGs have the range from 14.3% to 19.9%, and SMBs have the range from 3.85% to 12.6%. Although these numbers are bigger than the empirical data, the dividend price effect is relatively smaller than the long-term dividend growth effect, which is consistent with the empirical pattern described above. In addition, on small stock portfolios (SG, SN, SV), the simulated capital gain effects are bigger than the simulated dividend price effects, while the opposite is true for bigger stocks (BG, BN, BV), consistent with Fama and French (2007), and Chen, Petkova, and Zhang (2008). Quarterly version of our model (Model $\mathcal{Q}$) is similar to Model $\mathcal{M}$, but these fit the data better. One caveat is that the model generates negative returns for big growth (BG) firms in monthly models. One can observe that the negativity of the equity returns from big growth firms comes from the expected capital gain effect ($\Delta P/P$). We suspect that this is partly due to the fact that our setup has a stochastic discount factor formed by a simple power utility function, given that this problem appears to be mitigated as risk aversion increases.

To demonstrate the countercyclical time-variation in VMG and value spread[15], we plot VMGs and value spreads at quarterly frequency with $\gamma = 3.5$ in Figure 13. Zhang (2005) argues that the exogenous countercyclical price of risk and the value firm's inflexibility of reversibility make the value premium and value spread countercyclical.

[15] The value spreads are computed by the difference between the log BE/ME of the value portfolio and the log BE/ME of the growth portfolio (Cohen, Polk, and Vuolteenaho (2003)).

Meanwhile, we focus on the extensive margins of dividend policy and investment behaviors and these frictions in financial and real sectors lead to countercyclical value premiums and value spreads as well as other cross-sectional variations in stock returns.

Decile Portfolios sorted by BE/ME, Size, and PPD We now pool decile portfolios sorted by the BE/ME and the size following the Fama-French method (Fama and French (1992)). The individual stocks are grouped into 10 portfolios sorted by book-to-market equity ratios and market values. The annualized expected portfolio returns are calculated as the equal-weighted portfolio returns. Additionally, we generate new 10 portfolios sorted by $\hat{a}_{it}$ values, which are fitted propensities to pay dividends such that,

$$\hat{\mathbb{P}}\left\{a_{it}=1|S_{it}\right\} \equiv \hat{a}_{it}=\hat{\alpha}_a+\hat{b}_{a1}\ln ME_{it}+\hat{b}_{a2}BE/ME_{it}+\hat{b}_{a3}I_{it}/K_{it}+\hat{b}_{a4}\Pi_{it}/K_{it},$$

in which $\hat{\alpha}_a=-1.42$, $\hat{b}_{a1}=0.30$, $\hat{b}_{a2}=0.73$, $\hat{b}_{a3}=-12.4$, and $\hat{b}_{a4}=10.7$ on average values.[16] Each coefficient is significant, and they represent the effects of the size, the book-to-market equity ratio, the investment, and the profitability, respectively. The signs of all coefficients are consistent with the empirical results in Fama and French (2001), and Denis and Osobov (2008).

Figure 14 based on Models $\mathcal{M}$ and $\mathcal{Q}$ shows that the value premium prevails even with a finer grid of value portfolios (Panels (A) and (D)). The value premiums projected by our model outperform the simulated data of Gomes, Kogan, and Zhang (2003) and Zhang (2005). The graphs show the robustness of our model for the value premium at each frequency. Panels (B) and (E) present that the size portfolios have the size premium correspondent to the empirical facts. Especially, the expected size premiums on the first and second decile portfolios are coherent to the movement of the empirical data due to non-existence of dividend price effect. However, the size premiums on other portfolios disappear as CRRA increases. The 10 portfolios sorted by PPD ($\hat{a}_{it}$) have risk premiums consistently increasing in PPD (Panels (C) and (F) in Figure 14). BE/ME and PPD are theoretically associated with each other since PPD represents the probability of occurrence of cash flow into shareholders, and the book-to-market equity is positively related to the cash flow risk as stated in Proposition 3.1.1. Thus, these Panels (C) and (F) demonstrate that the propensity to pay dividends captures the risk factor on cross-sectional behaviors of individual stock returns.

[16] $\ln ME_{it}$ denotes the log market size that is generated by $\ln P_{it}$ in our model.

5 Conclusion

This paper considers an important dimension of corporate decision, dividend policy, to construct a general equilibrium model of production and investment with many firms to analyze stock returns. This model not only links firm characteristics such as size and book-to-market equity ratio to stock returns, but also explains the relationships of these variables with the propensity to pay dividends and the profitability of firms. We focus on the extensive margin of dividend policy such that dividend policy can be highly persistent with occasional discrete shifts. According to the model, firms pay no dividend if they are either young and growing or showing very poor performances. On the other hand, dividend-paying firms are the mature firms with relatively high book-to-market equity ratios or showing very good performances. The existence of persistently zero dividend payouts makes cash flow risks smaller due to the lack of covariations of these risks with the stochastic discount factor. This can generate the value and size premia, because of the diverse patterns of dividend payment for those respective firms. Quantitative results show that the model can explain the cross sections of stock returns by interactions between dividend payouts and other firm variables, which cries out for more rigorous studies. Modeling more corporate financial features, such as stock repurchases and capital structure, is a natural next step to further analyze this aspect.

Finally, the potential lumpiness in individual cash flows has a few interesting implications for the time-series behaviors of aggregate variables as well. Specifically, the model, despite simple preferences used for the stochastic discount factor, generates countercyclical variations in market risk premiums and time-varying volatilities of aggregate dividend and asset returns. More realistic stochastic discount factors and the inclusion of labor or non-tradable goods market will be useful additions to better understand this feature. We leave these to future works.

References

Abel, A. B., and J. C. Eberly (2002): "Investment and q with Fixed Costs: An Empirical Analysis," *Working Paper*.

Adda, J., and R. Cooper (2003): *Dynamic Economics: Quantitative Methods and Applications*. The MIT Press.

Baker, M., and J. Wurgler (2004): "A Catering Theory of Dividends," *Journal of Finance*, 59(3), 1125–1165.

Bulan, L., N. Subramanian, and L. Tanlu (2007): "On the Timing of Dividend Initiations," *Financial Management*, 36(4), 31–65.

Chen, L., R. Petkova, and L. Zhang (2008): "The Expected Value Premium," *Journal of Financial Economics*, 87, 269–280.

Cochrane, J. H. (1996): "A Cross-Sectional Test of an Investment-Based Asset Pricing Model," *Journal of Political Economy*, 104(3), 572–621.

——— (2005): *Asset Pricing*. Princeton University Press, revised edn.

Cochrane, J. H., F. A. Longstaff, and P. Santa-Clara (2008): "Two Trees," *Review of Financial Studies*, 21(1).

Cohen, R. B., C. Polk, and T. Vuolteenaho (2003): "The Value Spread," *Journal of Finance*, 58, 609–641.

Cooley, T. F., and E. C. Prescott (1995): "Economic Growth and Business Cycles," in *Frontiers of Business Cycle Research*, ed. by T. F. Cooley. Princeton University Press.

Cooper, R., and J. Ejarque (2001): "Exhuming Q: Market Power Vs. Capital Market Imperfections," *NBER Working paper 8182*.

——— (2003): "Financial Frictions and Investment: Requiem in Q," *Review of Economic Dynamics*, 6, 710–728.

DeAngelo, H., L. DeAngelo, and D. J. Skinner (2004): "Are Dividends Disappearing? Dividend Concentration and the Consolidation of Earnings," *Journal of Financial Economics*, 72, 425–456.

DeAngelo, H., L. DeAngelo, and R. Stulz (2006): "Dividend Policy and the Earned/Contributed Capital Mix: A Test of the Life-Cycle Theory," *Journal of Financial Economics*, 81(2), 227–254.

Denis, D. J., and I. Osobov (2008): "Why Do Firms Pay Dividends? International Evidence on the Determinants of Dividend Policy," *Journal of Financial Economics*, 89, 62–82.

Easterbrook, F. H. (1984): "Two Agency-Cost Explanations of Dividends," *American Economic Review*, 74(4), 650–659.

Fama, E. F., and K. R. French (1992): "The Cross-Section of Expected Stock Returns," *Journal of Finance*, 47(2), 427–465.

——— (2001): "Disappearing Dividends: Changing Firm Characteristics or Lower Propensity to Pay?," *Journal of Financial Economics*, 60, 3–43.

——— (2002): "The Equity Premium," *Journal of Finance*, 57(2), 637–659.

——— (2006): "The Value Premium and the CAPM," *Journal of Finance*, 61(5).

——— (2007): "The Anatomy of Value and Growth Stock Returns," *Financial Analysts Journal*, 63(6).

Gomes, J., L. Kogan, and L. Zhang (2003): "Equilibirum Cross Section of Returns," *Journal of Political Economy*, 111, 689–732.

Grullon, G., R. Michealy, and B. Swaminathan (2002): "Are Dividend Changes a Sign of Firm Maturity?," *Journal of Business*, 75(3), 387–424.

Guvenen, F. (2009): "A Parsimonious Macroeconomic Model For Asset Pricing," *Econometrica*, 77(6), 1711–1750.

Harford, J. (1999): "Corporate Cash Reserves and Acquisitions," *Journal of Finance*, 54(6), 1969–1997.

Harford, J., S. A. Mansi, and W. F. Maxwell (2008): "Corporate Governance and Firm Cash Holdings in the US," *Journal of Financial Economics*, 87, 535–555.

Hennessy, C. A., and T. M. Whited (2005): "Debt Dynamics," *Journal of Finance*, 60, 1129–1165.

——— (2007): "How Costly is External Financing? Evidence from a Structural Estimation," *Journal of Finance*, 62, 1705–1745.

Hopenhayn, H. A. (1992): "Entry, Exit, and Firm Dynamics in Long Run Equilibrium," *Econometrica*, 60(5), 1127–1150.

Jensen, M. C. (1986): "Agency Costs of Free Cash Flow, Corporate Finance, and Takeovers," *Americal Economic Review*, 76(2), 323–329.

Jensen, M. C., and W. H. Meckling (1976): "Theory of the Firm: Managerial Behavior, Agency Costs and Ownership Structure," *Journal of Financial Economics*, 3(4), 305–360.

Lease, R., K. John, A. Kalay, U. Lowenstein, and O. H. Sarig (1999): *Dividend Policy: Its Impact on Firm Value*. The Harvard Business School Press.

Lettau, M., and J. A. Wachter (2007): "Appearing and Disappearing Dividends: The Link to Catering Incentives," *Journal of Finance*, 62(1), 55–92.

Li, E. X. N., D. Livdan, and L. Zhang (2009): "Anomalies," *Review of Financial Studies*, 22(11).

Livdan, D., H. Sapriza, and L. Zhang (2009): "Financially Constrained Stock Returns," *Journal of Finance*, 64(4).

Lucas, R. E. (1978): "Asset Prices in an Exchange Economy," *Econometrica*, 46.

Martin, I. (2009): "The Lucas Orchard," *Working paper.*

Menzly, L., T. Santos, and P. Veronesi (2004): "Understanding Predictablility," *Journal of Political Economy*, 112(1), 1–47.

Mueller, D. (1972): "A Life Cycle Theory of the Firm," *Journal of Industrial Economics*, 20(3), 199–219.

Nikolov, B., and T. M. Whited (2009): "Agency Conflicts and Cash: Estimates from a Structural Model," *Working Paper.*

Santos, T., and P. Veronesi (2009): "Habit Formation, the Cross Section of Stock Returns and the Cash-Flow Risk Puzzle," *Journal of Financial Economics.*

Smith, C. W., and R. L. Watts (1992): "The Investment Opportunity Set and Corporate Financing, Dividend, and Compensation Policies," *Journal of Financial Economics*, 32, 263–292.

Stokey, N. L., and R. E. Lucas (1989): *Recursive Mothods in Economic Dynamics.* Harvard University Press.

Zhang, L. (2005): "The Value Premium," *Journal of Finance*, 60(1).

A Technical Assumptions and Proofs

A.1 Technical Assumptions

Assumption A.1.1. We assume that

$$\mathbb{E}_t\left[V_{KK}\right] > -\frac{e^{x+z}K^{\theta-4}}{18\beta\left(1-\delta\right)}\left(\begin{array}{c}\xi_2 e^{(\xi_2+2)x+2z}K^{2\theta}+6K\left(K+\xi_1 e^{(1+\xi_2)x+z}K^{\theta}\left(\theta-1\right)\right)\\ +6\alpha_{21}e^{\xi_2 x}K^2\left(3\theta-2\right)\end{array}\right),$$

$$\mathbb{E}_t\left[V_{Kx}\right] > \frac{\xi_1 e^{(1+\xi_2)x+z}K^{\theta-3}\left(K+e^{x+z}K^{\theta}\right)\left(\xi_2 e^{x+z}K^{\theta}-2K\right)}{2\beta},$$

and

$$\mathbb{E}_t\left[V_{Kz}\right] > -\frac{\xi_1 e^{(1+\xi_2)x+z}K^{\theta-2}\left(3K+e^{x+z}K^{\theta}\right)}{3\beta}.$$

Assumption A.1.2. If S_{1t+u^1} and S_{2t+u^1} are in $\mathcal{S}^1$, and $P_{1t}=P_{2t}$, then

$$\frac{\sum_{u=u^1}^{\infty}\left(\mathbb{E}_{t+1}\left[M_{t,t+u}D_{it+u}\right]-\mathbb{E}_t\left[M_{t,t+u}D_{it+u}\right]\right)}{P_{1t}} - \frac{\sum_{u=u^1}^{\infty}\left(\mathbb{E}_{t+1}\left[M_{t,t+u}D_{jt+u}\right]-\mathbb{E}_t\left[M_{t,t+u}D_{jt+u}\right]\right)}{P_{2t}} = o\left(1\right).$$

A.2 Proof of Proposition 2.1.3

We suppress subscripts and superscripts for the simplicity. From (2) , the first order conditions for the investment is

$$-\frac{\left(I+K\right)\left(\xi_1 e^{\xi_2 x}I^2+2K\left(\xi_1 e^{\xi_2 x}I+K-\xi_1 e^{(\xi_2+1)x+z}K^{\theta}\right)\right)}{2K^3}+\beta\mathbb{E}_t\left[V_K\left(I+\left(1-\delta\right)K,x,z\right)\right]=0.$$

By the implicit function theorem, $\partial I/\partial K$ is

$$\frac{\left[\begin{array}{c}3\xi_1 e^{\xi_2 x}I^3+2IK\left(K+\xi_1 e^{\xi_2 x}\left(3I+e^{x+z}K^{\theta}\left(\theta-2\right)\right)\right)\\ +2\xi_1 e^{\xi_2 x}K^2\left(I+e^{x+z}K^{\theta}\left(\theta-1\right)\right)+2\beta\left(1-\delta\right)K^4\mathbb{E}_t\left[V_{KK}\right]\end{array}\right]}{K\left(3\xi_1 e^{\xi_2 x}I^2+2\xi_1 e^{\xi_2 x}K^2+2K\left(K-\xi_1 e^{\xi_2 x}\left(e^{x+z}K^{\theta}-3I\right)\right)-2\beta K^2\mathbb{E}_t\left[V_{KK}\right]\right)}. \tag{14}$$

We know that $\mathbb{E}_t\left[V_{KK}\right]<0$ since the value function of dynamic programming is concave in Stokey and Lucas (1989). From the assumption $e^{x+z}K^{\theta}<3I^0$ and the first condition of Assumption A.1.1, the numerator and denominator of (14) are positive, hence

$$\frac{\partial I}{\partial K}>0.$$

For $\partial I/\partial x$, we have

$$\frac{-\xi_1 e^{\xi_2 x}\left(I+K\right)\left(\xi_2 I^2-2K\left(-\xi_2 I+e^{x+z}K^{\theta}+\xi_2 e^{x+z}K^{\theta}\right)\right)+2\beta K^2\mathbb{E}_t\left[V_{Kx}\right]}{3\xi_1 e^{\xi_2 x}I^2+2\xi_1 e^{\xi_2 x}K^2+2K\left(K-\xi_1 e^{\xi_2 x}\left(e^{x+z}K^{\theta}-3I\right)\right)-2\beta K^2\mathbb{E}_t\left[V_{KK}\right]}. \tag{15}$$

From the second condition in Assumption A.1.1., the numerator of (15) is positive, thus

$$\frac{\partial I}{\partial x}>0.$$

For $\partial I/\partial z$, the following is

$$\frac{\partial I}{\partial z}=\frac{2K\left(\xi_1 e^{(1+\xi_1)x+z}K^{\theta}\left(I+K\right)+\beta K^2\mathbb{E}_t\left[V_{Kz}\right]\right)}{3\xi_1 e^{\xi_2 x}I^2+2\xi_1 e^{\xi_2 x}K^2+2K\left(K-\xi_1 e^{\xi_2 x}\left(e^{x+z}K^{\theta}-3I\right)\right)-2\beta K^2\mathbb{E}_t\left[V_{KK}\right]}. \tag{16}$$

We have the assumption such that

$$\mathbb{E}_t\left[V_{Kz}\right] > -\frac{\xi_1 e^{(1+\xi_2)x+z} K^{\theta-2}\left(3K + e^{x+z}K^{\theta}\right)}{3\beta},$$

then, the numerator of (16) is positive, therefore,

$$\frac{\partial I}{\partial z} > 0,$$

which completes the proof. □

A.3 Proof of Proposition 2.3.2

We refer to Proposition 2.1.1. Our recursive general equilibrium model is based on the industry equilibrium model (Hopenhayn (1992)). We can apply Theorem 2 in the proof of Proposition 2 of Appendix A in Zhang (2005). □

A.4 Proof of Proposition 3.1.1

First, the cash flow risk effect is,

$$-\frac{\mathbb{C}ov_t\left[M_{t,t+1}, \frac{D_{1t+1}}{P_{1t}}\right]}{\mathbb{E}_t\left[M_{t,t+1}\right]} + \frac{\mathbb{C}ov_t\left[M_{t,t+1}, \frac{D_{2t+1}}{P_{2t}}\right]}{\mathbb{E}_t\left[M_{t,t+1}\right]} = 0 + \frac{\mathbb{C}ov_t\left[\left(\frac{D_{2t}}{D_{2t+1}}\right)^{\gamma}, \frac{D_{2t+1}}{P_{2t}}\right]}{\mathbb{E}_t\left[M_{t,t+1}\right]} < 0.$$

For the discount risk effect, from (6) , by the law of iterated expectation, we can show the following:

$$\begin{aligned}
\mathbb{E}_t\left[P_{1t+1} - P_{1t}\right] &= \mathbb{E}_t\left[\sum_{u=1}^{\infty}\left(\beta^u \mathbb{E}_{t+1}\left[\left(\frac{C_{t+1}}{C_{t+u+1}}\right)^{\gamma} D_{1t+u+1}\right] - \beta^u \mathbb{E}_t\left[\left(\frac{C_t}{C_{t+u}}\right)^{\gamma} D_{1t+u}\right]\right)\right] \\
&= \sum_{u=u^0}^{\infty}\left(\beta^u \mathbb{E}_t\left[\left(\frac{C_{t+1}}{C_{t+u+1}}\right)^{\gamma} D_{1t+u+1}\right] - \beta^{u+1}\mathbb{E}_t\left[\left(\frac{C_t}{C_{t+u+1}}\right)^{\gamma} D_{1t+u+1}\right]\right).
\end{aligned}$$

The innovation between ΔP_{1t+1} and $\mathbb{E}_t\left[\Delta P_{1t+1}\right]$ is

$$\sum_{u=u^0}^{\infty}\beta^u\left(\mathbb{E}_{t+1}\left[\left(\frac{C_{t+1}}{C_{t+u+1}}\right)^{\gamma} D_{1t+u+1}\right] - \mathbb{E}_t\left[\left(\frac{C_{t+1}}{C_{t+u+1}}\right)^{\gamma} D_{1t+u+1}\right]\right).$$

Similarly, for the stock of the firm 2,

$$\Delta P_{2t+1} - \mathbb{E}_t\left[\Delta P_{2t+1}\right] = \sum_{u=1}^{\infty}\beta^u\left(\mathbb{E}_{t+1}\left[\left(\frac{C_{t+1}}{C_{t+u+1}}\right)^{\gamma} D_{2t+u+1}\right] - \mathbb{E}_t\left[\left(\frac{C_{t+1}}{C_{t+u+1}}\right)^{\gamma} D_{2t+u+1}\right]\right).$$

By Assumption A.1.2, the difference between the innovations of expected capital gains is

$$\begin{aligned}
\Delta\bar{P}_{t+1} - \mathbb{E}_t\left[\Delta\bar{P}_{t+1}\right] &\equiv \frac{\Delta P_{1t+1} - \mathbb{E}_t\left[\Delta P_{1t+1}\right]}{P_{1t}} - \frac{\left(\Delta P_{2t+1} - \mathbb{E}_t\left[\Delta P_{2t+1}\right]\right)}{P_{2t}} \\
&= -\frac{1}{P_{2t}}\sum_{u=1}^{u^0-1}\beta^u\left(\mathbb{E}_{t+1}\left[\left(\frac{D_{2t+1}}{D_{2t+u+1}}\right)^{\gamma} D_{2t+u+1}\right] - \mathbb{E}_t\left[\left(\frac{D_{2t+1}}{D_{2t+u+1}}\right)^{\gamma} D_{2t+u+1}\right]\right) \\
&\quad + o\left(1\right).
\end{aligned}$$

Then, by the independence of x_t, z_{1t} and z_{2t}, the difference of discount risk effects of firm 1 and 2 is

$$-\frac{\mathbb{C}ov_t\left[M_{t,t+1},\frac{P_{1t+1}}{P_{1t}}\right]}{\mathbb{E}_t\left[M_{t,t+1}\right]}-\left(-\frac{\mathbb{C}ov_t\left[M_{t,t+1},\frac{P_{2t+1}}{P_{2t}}\right]}{\mathbb{E}_t\left[M_{t,t+1}\right]}\right)=\frac{-\mathbb{C}ov_t\left[M_{t,t+1},\Delta\bar{P}_{t+1}\right]}{\mathbb{E}_t\left[M_{t,t+1}\right]},$$

and

$$\begin{aligned}&-\mathbb{C}ov_t\left[M_{t,t+1},\Delta\bar{P}_{t+1}\right]\\&=\frac{1}{P_{2t}}\sum_{u=1}^{u^0-1}\mathbb{C}ov_t\left[M_{t,t+1},\beta^u\left(\mathbb{E}_{t+1}\left[\left(\frac{D_{2t+1}}{D_{2t+u+1}}\right)^\gamma D_{2t+u+1}\right]-\mathbb{E}_t\left[\left(\frac{D_{2t+1}}{D_{2t+u+1}}\right)^\gamma D_{2t+u+1}\right]\right)\right]\\&+o\left(1\right).\end{aligned}$$

Then, for $1\leq u\leq u^0-1$,

$$\begin{aligned}&\mathbb{C}ov_t\left[M_{t,t+1},\beta^u\left(\mathbb{E}_{t+1}\left[\left(\frac{D_{2t+1}}{D_{2t+u+1}}\right)^\gamma D_{2t+u+1}\right]-\mathbb{E}_t\left[\left(\frac{D_{2t+1}}{D_{2t+u+1}}\right)^\gamma D_{2t+u+1}\right]\right)\right]\\&\quad=\mathbb{E}_t\left[\left(\frac{D_{2t}}{D_{2t+1}}\right)^\gamma\beta^{u+1}\left(\mathbb{E}_{t+1}\left[\left(\frac{D_{2t+1}}{D_{2t+u+1}}\right)^\gamma D_{2t+u+1}\right]-\mathbb{E}_t\left[\left(\frac{D_{2t+1}}{D_{2t+u+1}}\right)^\gamma D_{2t+u+1}\right]\right)\right],\end{aligned}$$

Let $\gamma=1$, then by Jensen's inequality,

$$\begin{aligned}&\mathbb{C}ov_t\left[M_{t,t+1},\beta^u\left(\mathbb{E}_{t+1}\left[D_{2t+1}\right]-\mathbb{E}_t\left[D_{2t+1}\right]\right)\right]\\&=\beta^{u+1}\mathbb{E}_t\left[\frac{D_{2t}}{D_{2t+1}}\mathbb{E}_{t+1}\left[D_{2t+1}\right]\right]-\beta^{u+1}\mathbb{E}_t\left[\frac{D_{2t}}{D_{2t+1}}\mathbb{E}_t\left[D_{2t+1}\right]\right]\\&=\beta^{u+1}D_{2t}\left(1-\mathbb{E}_t\left[\frac{1}{D_{2t+1}}\right]\mathbb{E}_t\left[D_{2t+1}\right]\right)<0,\end{aligned}$$

and

$$-\mathbb{C}ov_t\left[M_{t,t+1},\Delta\bar{P}_{t+1}\right]<0.$$

Now, we want to generalize the result above for $\gamma\geq 2$. Before the start, we assume that there do not exist firm-specific shock, $\mu_x=0$, and the investment process is determined at time t[17]. Then,

$$\begin{aligned}&\mathbb{C}ov_t\left[M_{t,t+1},\beta^u\left(\mathbb{E}_{t+1}\left[\left(\frac{D_{2t+1}}{D_{2t+u+1}}\right)^\gamma D_{2t+u+1}\right]-\mathbb{E}_t\left[\left(\frac{D_{2t+1}}{D_{2t+u+1}}\right)^\gamma D_{2t+u+1}\right]\right)\right]\\&\qquad=\beta^{u+1}D_{2t}\left(\mathbb{E}_t\left[D_{2t+1}^{\gamma-1}\mathbb{E}_{t+1}\left[\frac{1}{D_{2t+u+1}^{\gamma-1}}\right]\right]-\mathbb{E}_t\left[\frac{1}{D_{2t+1}}\right]\mathbb{E}_t\left[\frac{D_{2t+1}^\gamma}{D_{2t+u+1}^{\gamma-1}}\right]\right).\end{aligned}$$

For $u=1$,

$$\begin{aligned}&\mathbb{E}_t\left[D_{2t+1}^{\gamma-1}\mathbb{E}_{t+1}\left[\frac{1}{D_{2t+2}^{\gamma-1}}\right]\right]\\&=\mathbb{E}_t\left[e^{(\gamma-1)\rho_x x_t+(\gamma-1)\epsilon_{t+1}}\mathbb{E}_{t+1}\left[e^{-(\gamma-1)\rho_x^2x_t-(\gamma-1)\rho_x\epsilon_{x\,t+1}-(\gamma-1)\epsilon_{x t+2}}\right]\right]\left(\frac{K_{2t+1}^\theta}{K_{2t+2}^\theta}\right)^{\gamma-1}\\&=e^{(\gamma-1)\rho_x(1-\rho_x)x_t+\frac{(\gamma-1)^2(1-\rho_x)^2\sigma_x^2}{2}+\frac{(\gamma-1)^2\sigma_x^2}{2}}\left(\frac{K_{2t+1}^\theta}{K_{2t+2}^\theta}\right)^{\gamma-1},\end{aligned}$$

[17] The first two restrictions are assumed for only mathematical simplicity.

and

$$\begin{aligned}
&\mathbb{E}_t\left[\frac{1}{D_{2t+1}}\right]\mathbb{E}_t\left[\frac{D_{2t+1}^{\gamma}}{D_{2t+2}^{\gamma-1}}\right] \\
&= \mathbb{E}_t\left[e^{-\rho_x x_t-\epsilon_{xt+1}}\right]\mathbb{E}_t\left[e^{\gamma\rho_x x_t+\gamma\epsilon_{xt+1}}\right]\mathbb{E}_t\left[e^{-(\gamma-1)\rho_x^2 x_t-(\gamma-1)\rho_x\epsilon_{xt+1}-(\gamma-1)\epsilon_{xt+2}}\right]\left(\frac{K_{2t+1}^{\theta}}{K_{2t+2}^{\theta}}\right)^{\gamma-1} \\
&= e^{(\gamma-1)\rho_x(1-\rho_x)x_t+\frac{\sigma_x^2}{2}+\frac{\gamma^2\sigma_x^2}{2}+\frac{(\gamma-1)^2\rho_x^2\sigma_x^2}{2}+\frac{(\gamma-1)^2\sigma_x^2}{2}}\left(\frac{K_{2t+1}^{\theta}}{K_{2t+2}^{\theta}}\right)^{\gamma-1},
\end{aligned}$$

also, for $\gamma \geq 2$,

$$\frac{(\gamma-1)^2(1-\rho_x)^2\sigma_x^2}{2}-\left(\frac{\sigma_x^2}{2}+\frac{\gamma^2\sigma_x^2}{2}+\frac{(\gamma-1)^2\rho_x^2\sigma_x^2}{2}\right) = -\sigma_x^2\left(\gamma+\rho_x-2\gamma\rho_x+\gamma^2\rho_x\right) < 0,$$

which shows Jensen's inequality effect, and leads to

$$\mathbb{E}_t\left[D_{2t+1}^{\gamma-1}\mathbb{E}_{t+1}\left[\frac{1}{D_{2t+2}^{\gamma-1}}\right]\right]-\mathbb{E}_t\left[\frac{1}{D_{2t+1}}\right]\mathbb{E}_t\left[\frac{D_{2t+1}^{\gamma}}{D_{2t+2}^{\gamma-1}}\right] < 0.$$

This is the same as when $\gamma = 1$. In addition, by Jensen's inequality for $2 \leq u \leq u^0 - 1$,

$$\mathbb{E}_t\left[D_{2t+1}^{\gamma-1}\mathbb{E}_{t+1}\left[\frac{1}{D_{2t+u+1}^{\gamma-1}}\right]\right]-\mathbb{E}_t\left[\frac{1}{D_{2t+1}}\right]\mathbb{E}_t\left[\frac{D_{2t+1}^{\gamma}}{D_{2t+u+1}^{\gamma-1}}\right] < 0.$$

Thus, for $\gamma \geq 1$,

$$\frac{-\mathbb{C}ov_t\left[M_{t,t+1},\Delta\bar{P}_{t+1}\right]}{\mathbb{E}_t\left[M_{t,t+1}\right]} < 0.$$

This result implies that

$$-\frac{\mathbb{C}ov_t\left[M_{t,t+1},\frac{P_{1t+1}}{P_{1t}}\right]}{\mathbb{E}_t\left[M_{t,t+1}\right]} < -\frac{\mathbb{C}ov_t\left[M_{t,t+1},\frac{P_{2t+1}}{P_{2t}}\right]}{\mathbb{E}_t\left[M_{t,t+1}\right]} \equiv \beta_1^{\text{disc}} < \beta_2^{\text{disc}}.$$

The proof is completed. □

A.5 Proof of Proposition 3.1.2

From (6), for some integer $u^0 > 0$

$$\frac{P_{1t}}{K_{1t}} = \sum_{u=u_0+1}^{\infty}\beta^u\mathbb{E}_t\left[M_{t,t+u}\frac{D_{1t+u}}{K_{1t}}\right] \text{ and } \frac{P_{2t}}{K_{2t}} = \sum_{u=1}^{\infty}\beta^u\mathbb{E}_t\left[M_{t,t+u}\frac{D_{2t+u}}{K_{2t}}\right].$$

The dividend equity ratio is

$$\frac{a_{t+u}\delta e^{x_{t+u}}K_{t+u}^{\theta}}{K_t},$$

in which

$$\begin{aligned}
K_{t+u} &= I_{t+u-1}+\left(1-\bar{\delta}\right)\left(I_{t+u-2}+\left(1-\bar{\delta}\right)K_{t+u-2}\right)... \\
&= \sum_{w=1}^{u}\left(1-\bar{\delta}\right)^{w-1}I_{t+u-w}+\left(1-\bar{\delta}\right)^{u}K_t.
\end{aligned}$$

Then, dividend-equity ratios for firm $i = 1, 2$ are

$$\frac{D_{it+u}}{K_{it}} = \frac{a_{it+u}\delta e^{x_{t+u}}}{K_{it}^{1-\theta}} \left[\sum_{w=1}^{u} \left(1-\bar{\delta}\right)^{w-1} \frac{I_{it+u-w}}{K_{it}} + \left(1-\bar{\delta}\right)^{u} \right]^{\theta}.$$

For simplicity, let us assume that $\bar{\delta} = 0$. We know that K_{t+1} is determined at t. That is,

$$\mathbb{E}_t \left[M_{t,t+u} \frac{D_{1t+u}}{K_{1t}} \right] - \mathbb{E}_t \left[M_{t,t+u} \frac{D_{2t+u}}{K_{2t}} \right] = \mathbb{E}_t \left[M_{t,t+u} \left(\frac{D_{1t+u}}{K_{1t}} - \frac{D_{2t+u}}{K_{2t}} \right) \right] > 0,$$

since for any $u^0 + 1 \le w \le u$, I/K is monotone decreasing with respect to K because of the diminishing return to scale, i.e.,

$$\frac{I_{1t+u-w}}{K_{1t}} > \frac{I_{2t+u-w}}{K_{2t}}.$$

The minimum of expected difference between $\frac{D_{1t+u}}{K_{1t}}$ and $\frac{D_{2t+u}}{K_{2t}}$ is defined as, for any $u \ge 0$,

$$d \equiv \min \mathbb{E}_t \left[\frac{D_{1t+u}}{K_{1t}} - \frac{D_{2t+u}}{K_{2t}} \right] > 0.$$

Then, for $\gamma > 1$,

$$\begin{aligned} \left(\frac{P_{1t}}{K_{1t}} - \frac{P_{2t}}{K_{2t}} \right) \frac{1}{D_{2t}^{\gamma}} &\ge -\sum_{u=1}^{u_0} \beta^u \mathbb{E}_t \left[\frac{1}{C_{t+u}^{\gamma-1}} \frac{1}{K_{2t}} \right] + d \sum_{u=u^0+1}^{\infty} \beta^u \mathbb{E}_t \left[\frac{1}{C_{t+u}^{\gamma}} \right] \\ &= \frac{d \left(\beta + \underline{g}(\gamma)\right)^{u^0+1}}{1 - \beta - \underline{g}(\gamma)} \mathbb{E}_t \left[\frac{1}{C_{t+u^0+1}^{\gamma}} \right] - \frac{\beta \left(1 - \left(\beta + \bar{g}(\gamma-1)\right)^{u^0} \right)}{\left(1 - \beta - \bar{g}(\gamma-1)\right) K_{2t}} \mathbb{E}_t \left[\frac{1}{C_{t+1}^{\gamma-1}} \right], \end{aligned}$$

in which

$$\begin{aligned} \underline{g}(\gamma) &\equiv \min \left\{ \frac{\mathbb{E}_t \left[C_{t+u+1}^{-\gamma} \right] - \mathbb{E}_t \left[C_{t+u}^{-\gamma} \right]}{\mathbb{E}_t \left[C_{t+u}^{-\gamma} \right]} \right\}, \\ \bar{g}(\gamma-1) &\equiv \max \left\{ \frac{\mathbb{E}_t \left[C_{t+u+1}^{-(\gamma-1)} \right] - \mathbb{E}_t \left[C_{t+u}^{-(\gamma-1)} \right]}{\mathbb{E}_t \left[C_{t+u}^{-(\gamma-1)} \right]} \right\}. \end{aligned}$$

$\underline{g}(\gamma)$ is the minimum for growth rates of expected marginal utilities of consumption with γ, and $\bar{g}(\gamma-1)$ is the maximum for growth rates of expected marginal utilities of consumption with $\gamma - 1$. If parametric restrictions are held such that $\mathbb{E}_t \left[1/C_{t+u^0+1}^{\gamma} \right] > \mathbb{E}_t \left[1/\left(C_{t+1}^{\gamma-1} dK_{2t} \right) \right]$ and

$$\frac{\left(\beta + \underline{g}(\gamma)\right)^{u^0+1}}{1 - \beta - \underline{g}(\gamma)} > \frac{\beta \left(1 - \left(\beta + \bar{g}(\gamma-1)\right)^{u^0} \right)}{1 - \beta - \bar{g}(\gamma-1)},$$

then,

$$\frac{P_{1t}}{K_{1t}} - \frac{P_{2t}}{K_{2t}} > 0 \Rightarrow \frac{K_{1t}}{P_{1t}} < \frac{K_{2t}}{P_{2t}},$$

which completes the proof. $\square$

Table 1: Benchmark Parameter Value Sets

This table lists the benchmark parameters used to solve and simulate the model. We calibrate 12 parameters $(\bar{\delta}, \rho_x, \sigma_x, \rho_z, \sigma_z, \phi, \delta, \beta, \mu_x, \theta, \xi_1, \xi_2)$ in monthly frequency and quarterly frequency to be consistent with the empirical literature. The monthly model is denoted as Model $\mathcal{M}$, and the quarterly model, as Model $\mathcal{Q}$. We categorize all parameters into two groups. The first parameter group $(\bar{\delta}, \rho_x, \sigma_x, \rho_z, \sigma_z, \phi, \delta, \beta)$ is calibrated outside models, and the second parameter group $(\mu_x, \theta, \xi_1, \xi_2)$ is calibrated inside models by the method using the idea of generalized method of moments. The moments to be used in this calibration are from average values of policy functions converged on grids of compact state space: profitability $(\Pi(S_{it})/K_{it})$, investment-capital ratio $(I_+(S_{it})/K_{it})$, disinvestment-capital ratio $(I_-(S_{it})/K_{it})$, and paying or not-paying dividend policy (a_{it}).

Parameter	Model $\mathcal{M}$ (Monthly)	Model $\mathcal{Q}$ (Quarterly)	Description
Calibrated Outside the Model			
δ	0.01	0.03	Capital depreciation rate
ρ_x	0.983	0.95	Persistence coefficient of aggregate productivity
σ_x	0.0023	0.007	Conditional volatility of aggregate productivity
ρ_z	0.97	0.912	Persistence coefficient of firm-specific productivity
σ_z	0.10	0.30	Conditional volatility of firm-specific productivity
ϕ	1	1	Convexity parameter of investment adjustment cost
δ	0.6	0.6	Long-run average level of dividend payout ratios
β	0.998	0.994	Time preference coefficient
Calibrated Inside the Model			
μ_x	-2.0	-1.71	Long-run average level of aggregate productivity
θ	0.68	0.65	Curvature in the production function
ξ_1	323	458	Convexity parameter of cash holding cost
ξ_2	0.084	0.044	Procyclical parameter of cash holding cost

Table 2: Unconditional Moments of Market Values

This table reports a set of key unconditional moments of market values under the benchmark models $\mathcal{M}$ and $\mathcal{Q}$ with parameters in Table 1. The empirical range of equity premium $(r_M - r_f)$ is from the variety of finance literature. The volatility of market return $(\sigma[r_M])$ is from Guvenen (2009) that is computed by Standard and Poor's 500 index covering 1890-1991. The cross-sectional volatility of individual stock returns $(\sigma[r_i])$ is reported by Zhang (2005). The capital gain $(\Delta P_{t+1}/P_t)$ and dividend price ratio (D_{t+1}/P_t) effects of market returns are reported by Fama and French (2002). The risk-free rate (r_f), its volatility $(\sigma[r_f])$, BE/ME, and $\sigma[BE/ME]$ are from Zhang (2005). $(BE/ME)^{\text{payer}} / (BE/ME)^{\text{non-payer}}$ is the ratio of average book-to-market equity ratios of dividend payers and non-payers. It is computed using CRSP/COMPUSTAT merged data. Empirical Sharpe ratio is reported by Cochrane (2005). The average of price dividend ratio (P/D), the average volatility of log of price dividend ratio $(\sigma[\log P/D])$, the aggregate dividend growth and its volatility $(\Delta \log D, \sigma[\Delta \log D])$ are reported by Guvenen (2009).

U.S. Data Annualized Data	(%)	Model $\mathcal{M}$ $\gamma=2$	$\gamma=3$	$\gamma=4$	Model $\mathcal{Q}$ $\gamma=2$	$\gamma=3$	$\gamma=3.5$
Market Return and Risk-Free Rate							
$r_M - r_f$	4 to 8	1.14	4.67	11.2	1.40	5.81	8.75
$\sigma[r_M]$	19.4	13.3	20.1	27.3	11.6	17.7	20.9
$\sigma[r_i]$	25 to 32	47.1	48.7	51.1	21.8	22.6	22.8
$\Delta P_{t+1}/P_t$	2.10	2.02	2.51	3.20	1.49	2.16	2.58
D_{t+1}/P_t	4.70	3.87	6.06	9.40	3.64	5.18	6.17
r_f	1.80	4.75	3.90	1.42	3.73	1.54	0.00
$\sigma[r_f]$	3.00	0.93	1.14	1.23	1.17	2.06	1.62
$\sigma[M]$	0.49	0.22	0.37	0.57	0.26	0.43	0.54
Sharpe Ratio	0.50	0.23	0.38	0.59	0.27	0.44	0.54
Aggregate Dividend Growth							
$\Delta \log D$	2.50	2.40	2.39	2.40	2.59	2.56	2.56
$\sigma[\Delta \log D]$	13.4	11.5	11.4	11.4	13.9	13.8	13.6
Book-to-Market							
BE/ME	0.67	0.16	0.24	0.37	0.37	0.52	0.62
$\sigma[BE/ME]$	0.23	0.07	0.10	0.15	0.15	0.20	0.24
$\frac{(BE/ME)^{\text{payer}}}{(BE/ME)^{\text{non-payer}}}$	1.32	2.34	2.27	2.20	1.70	1.65	1.63
Price-Dividend Ratio							
P/D	22.1	26.3	16.8	10.9	28.9	20.5	17.3
$\sigma[\log P/D]$	26.3	8.49	16.8	24.6	12.0	23.5	29.2

Table 3: Size and Value Factors and the Size-BE/ME Portfolios

This table lists the size and value factors and the six size-BE/ME portfolios. We follow the Fama-French method to pool the factor portfolios, SMB (small minus big) and VMG (value minus growth, previously HML), and the six value-weighted size-BE/ME portfolios (Fama and French (2006)). Firms below the median size are small (S) and those above are big (B) in simulated 2703 firms. We assign firms to growth (G), neutral (N), and value (V) groups if their BE/ME is in the bottom 30%, middle 40%, or top 30% of simulated firms. The six portfolios, small and big growth (SG and BG), small and big neutral (SN and BN), and small and big value (SV and BV) are the intersections of these sorts. In our models $\mathcal{M}$ and $\mathcal{Q}$, SMB is the simulated expected returns on the three small-size stock portfolios minus the returns on the three big-size stock portfolios such that, SMB$=\frac{\text{SG+SN+SV}}{3} - \frac{\text{BG+BN+BV}}{3}$,and VMG$=\frac{\text{SV+BV}}{2} - \frac{\text{SG+BG}}{2}$, which is the value-growth factor, VMG is the expected returns on the two value portfolios minus the expected returns on the two growth portfolios. The empirical capital gain ($\Delta P_{t+1}/P_t$), dividend price ratio (D_{t+1}/P_t) effects, and returns for size and value portfolios are reported by Fama and French (2007), and Chen, Petkova, and Zhang (2008).

		Factor Portfolios			Size-BE/ME Portfolios					
		$r_M - r_f$	SMB	VMG	SG	SN	SV	BG	BN	BV
Annualized Data(%)										
1927-2006 data (Fama and French (2007))										
		4 to 8	1.67	4.19	8.69	12.9	14.4	9.18	10.0	11.8
	$\Delta P_{t+1}/P_t$		2.51	3.43	5.70	9.04	11.2	5.80	5.54	7.13
	D_{t+1}/P_t		−0.84	0.77	2.98	3.87	3.21	3.39	4.49	4.69
1945-2005 data (Chen, Petkova, and Zhang (2008))										
		4 to 8	2.70	5.20	8.60	12.4	15.1	7.80	8.50	11.7
	$\Delta P_{t+1}/P_t$		3.00	4.95	2.10	5.50	10.3	2.70	1.80	4.40
	D_{t+1}/P_t		−0.30	0.25	6.50	6.90	4.80	5.10	6.70	7.30
Estimated Expected Values (Chen, Petkova, and Zhang (2008))										
		4 to 8	2.13	6.10	4.80	9.30	13.2	5.60	6.00	9.30
	$\Delta P_{t+1}/P_t$		3.06	4.40	2.10	5.50	9.40	2.40	1.50	3.90
	D_{t+1}/P_t		−0.93	1.70	2.70	3.80	3.80	3.10	4.50	5.40
Model $\mathcal{M}$ (Monthly)										
$\gamma = 2$		1.14	12.6	19.9	4.61	7.02	23.6	−14.4	5.19	6.54
	$\Delta P_{t+1}/P_t$		16.4	17.7	4.29	4.74	20.6	−18.7	−1.45	0.45
	D_{t+1}/P_t		−3.82	2.2	0.33	2.28	2.99	4.32	6.64	6.09
$\gamma = 3$		4.67	8.88	17.4	5.93	9.04	20.7	−9.97	8.90	10.1
	$\Delta P_{t+1}/P_t$		15.6	15.2	5.39	5.65	16.0	−19.0	−1.48	0.69
	D_{t+1}/P_t		−6.74	2.22	0.54	3.39	4.62	9.03	10.4	9.38
$\gamma = 4$		11.2	3.85	14.3	8.05	11.7	20.9	−0.62	14.7	15.1
	$\Delta P_{t+1}/P_t$		14.3	11.1	7.19	6.77	13.8	−14.7	−1.44	0.88
	D_{t+1}/P_t		−10.4	3.21	0.86	4.92	7.12	14.1	16.1	14.2
Model $\mathcal{Q}$ (Quarterly)										
$\gamma = 2$		1.40	3.28	6.17	4.06	5.78	9.99	−0.50	4.59	5.91
	$\Delta P_{t+1}/P_t$		7.40	2.62	4.02	5.53	9.09	−0.50	−2.73	−0.34
	D_{t+1}/P_t		−4.13	3.55	0.04	0.25	0.90	0.00	7.32	6.25
$\gamma = 3$		5.81	1.60	6.97	4.94	7.08	10.5	0.76	7.91	9.09
	$\Delta P_{t+1}/P_t$		7.47	2.17	4.89	6.76	9.44	0.76	−2.61	0.53
	D_{t+1}/P_t		−5.87	4.80	0.05	0.32	1.10	0.00	10.5	8.56
$\gamma = 3.5$		8.75	0.48	6.90	5.65	7.82	11.4	2.68	10.1	10.7
	$\Delta P_{t+1}/P_t$		7.43	1.48	5.59	7.41	10.2	2.67	−2.73	0.99
	D_{t+1}/P_t		−6.95	5.41	0.06	0.41	1.21	0.01	12.8	9.69

Figure 1: Lumpy Dividend Policies

This figure shows that eight firms have lumpy dividend policies with fixed payout ratios for examples. DPS is dividends per share (DVPSX_F) and EPS is earnings per share (EPSPX) at fiscal year, t in CRSP/COMPUSTAT merged data. Star-solid line shows DPS and diamond-solid line plots EPS.

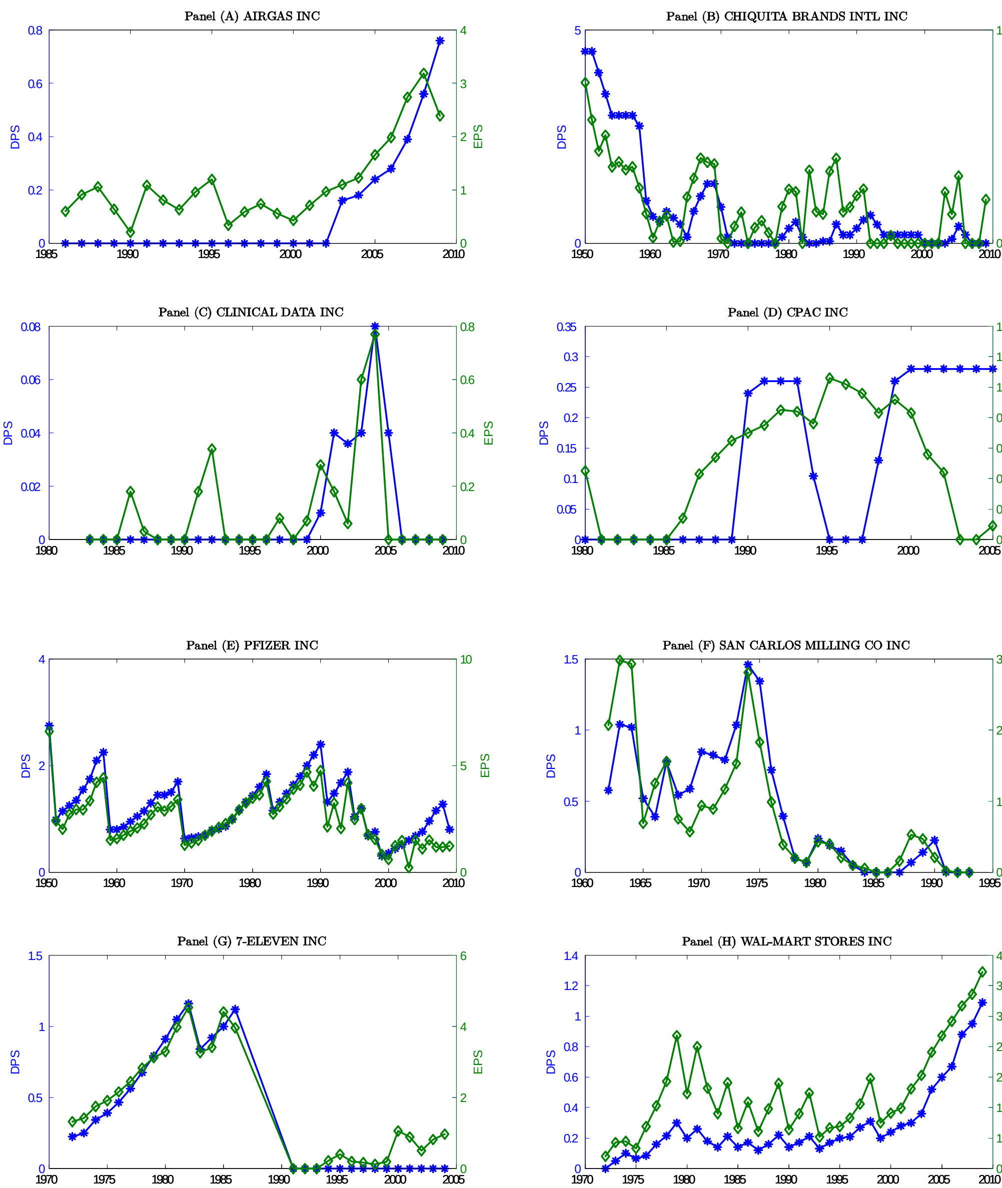


Figure 2: Aggregate Payout Ratios

This figure shows aggregate payout ratios in the U.S. stock market, and the regression coefficients such that

$$\delta_t = a_\delta + \beta_{\delta 1} \frac{\Pi_t}{A_t} + \beta_{\delta 2} \frac{I_t}{A_t} + e_t.$$

δ_t is computed as the aggregation of dividends per share (DVPSX_F) over the aggregation of earnings per share (EPSPX) at the fiscal year, t in CRSP/COMPUSTAT merged data (CCM). Π_t/K_t is the aggregated profitability as $\sum_{i=1}^{n} \Pi_{it} / \sum_{i=1}^{n} A_{it}$ using EBITDA (Π_{it}, earnings before interest, taxes, depreciation and amortization) and AT (A_{it}, assets - total) in CCM, and I_t/A_t is the aggregated investment ratio as $\sum_{i=1}^{n} I_{it} / \sum_{i=1}^{n} A_{it}$, in which I_{it} is CAPX (capital expenditures) minus SPPE (sale of property) in CCM. Panel (A) shows payout ratios, Panels (B) and (C) show $\beta_{\delta 1}$ and $\beta_{\delta 2}$ with moving windows of 27 years, which have starting years from 1950 to 1980. Dashed lines show 99-percent bandwidth using Newey-West autocorrelation consistent covariance estimators.

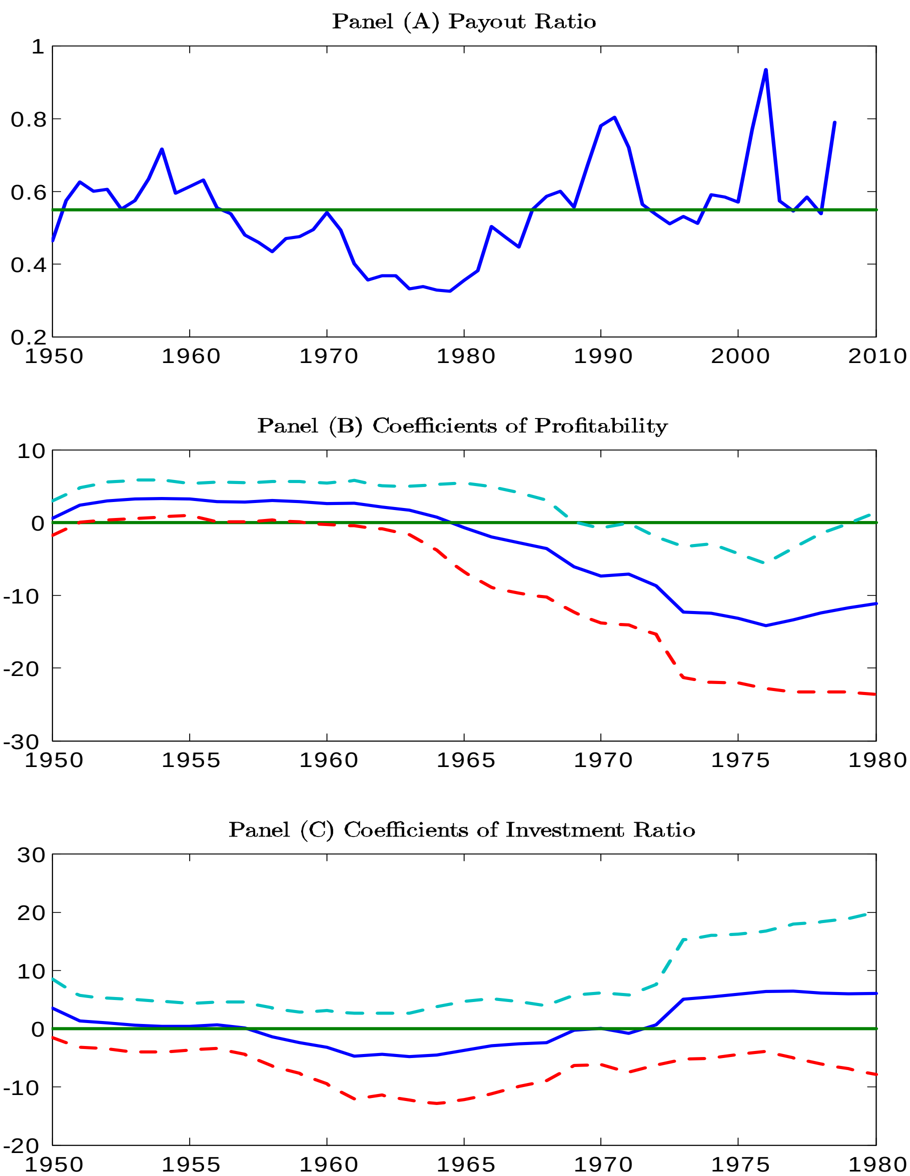


Figure 3: Regression about Payout Ratios

This figure shows the regression analysis of DPS on EPS and investment per share for individual firms which have more than 20 annual observations such that

$$\mathrm{DPS}_{it} = a_i + \beta_{i1}\mathrm{EPS}_{it} + \beta_{i2}\frac{I_{it}}{\varphi_{it}} + e_{it}.$$

DPS_{it} is firm i's dividends per share (DVPSX_F) and EPS_{it} is firm i's earnings per share (EPSPX) at the fiscal year, t in CRSP/COMPUSTAT merged data (CCM). I_{it} is CAPX (capital expenditures) minus SPPE (sale of property) and φ_{it} is common shares outstanding (CSHO) in CCM. Panels (A), (B) and (C) show distributions of estimated α_i, β_{i1} and β_{i2}, respectively. Vertical lines show means of estimated α_i, β_{i1} and β_{i2}. Panels (D), (E) and (F) show distributions of t values of α_i, β_{i1} and β_{i2}, respectively. t values are computed by Newey-West autocorrelation consistent covariance estimators. Vertical lines show the significance level (5%) for two-sided test.

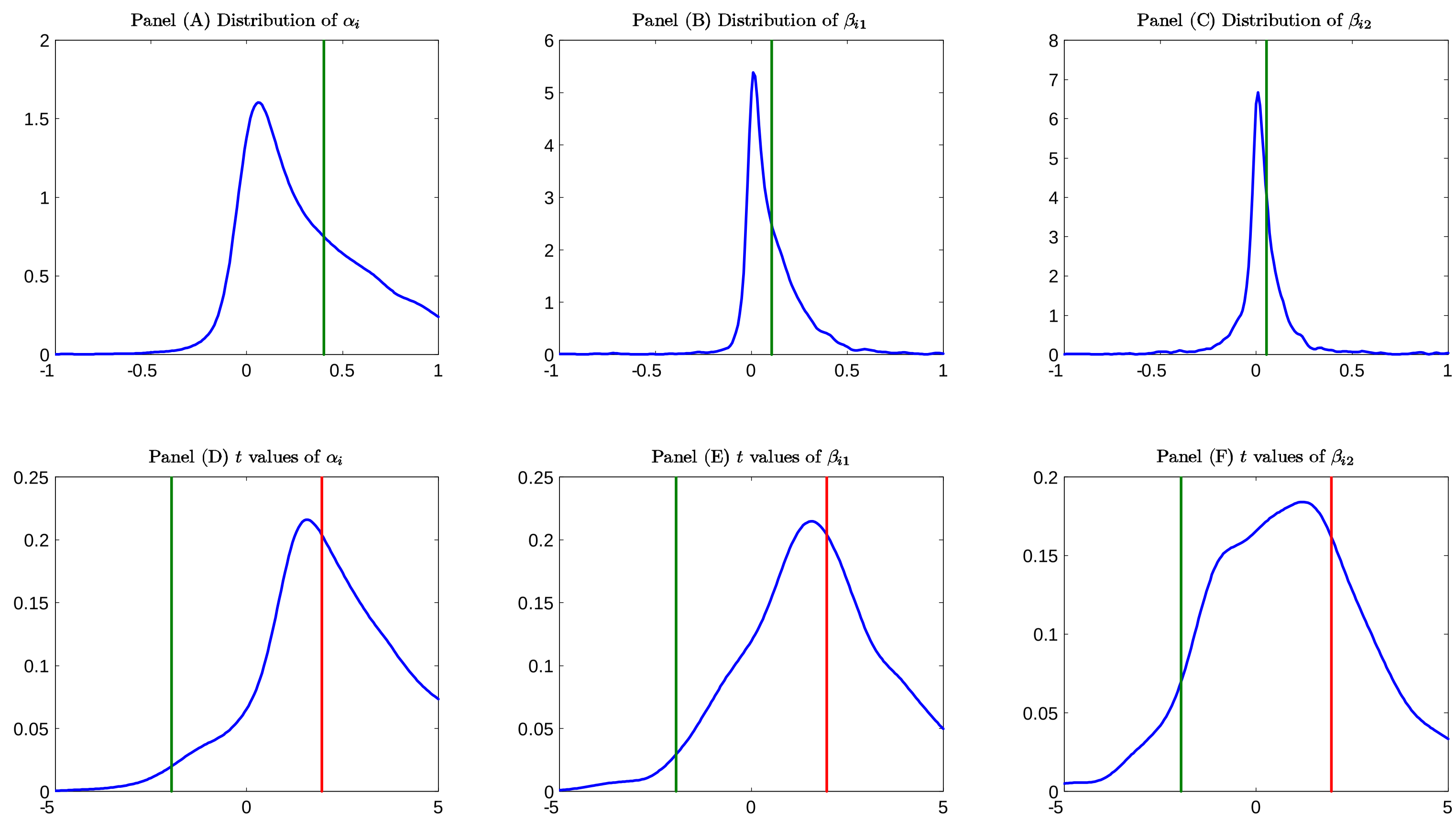
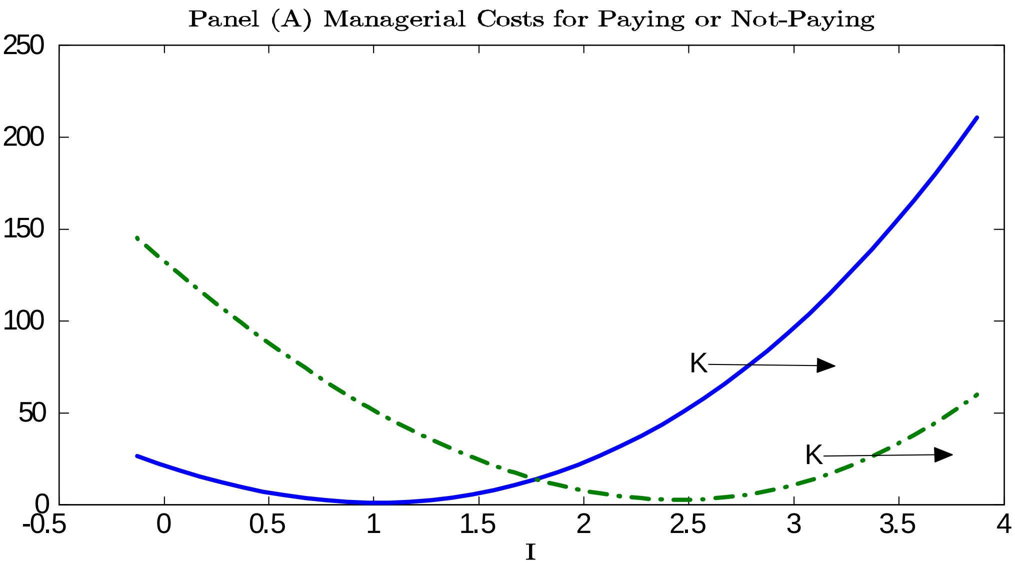


Figure 4: Non-convex Managerial Cost: $\Pi + \Xi$

This figure shows the managerial costs for paying or not-paying dividends. In Panel (A), the dash-dot line represents the managerial cost without paying dividends $\left(\Phi^0 + \Xi^0\right)$, and the solid line shows the managerial cost while paying dividends to shareholders $\left(\Phi^1 + \Xi^1\right)$. Arrows show directions of the increase in capital. Panel (B) shows the total managerial cost which a firm faces in dynamic programming problem.



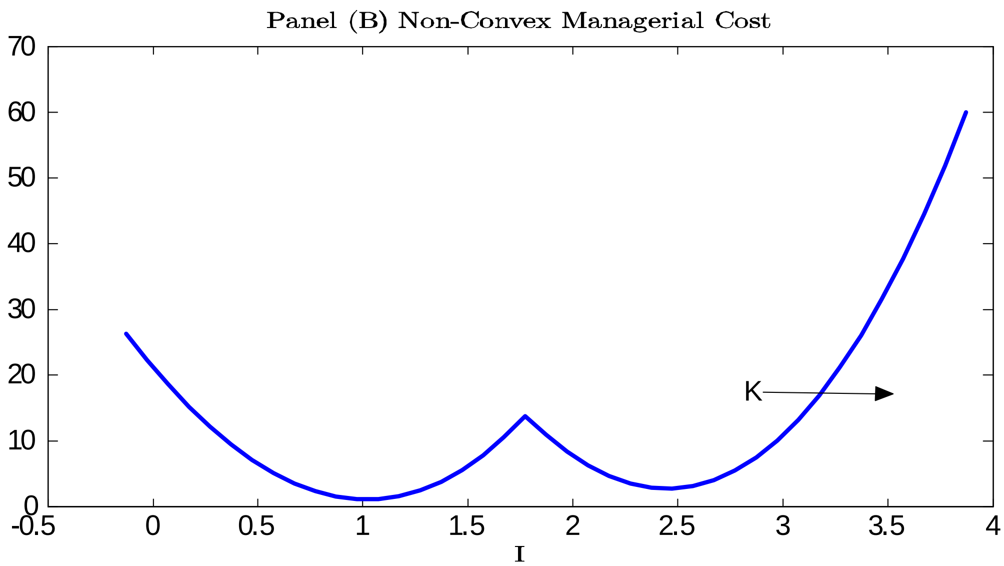


Figure 5: V^0 and V^1 according to Increase in x

This figure shows movements of V^0 and V^1 according to the increase in x_t. The graph is drawn from the fitted model with quarterly frequency. For the ease of exposition, only the local neighborhood around the point at which $V_1 = V_0$ is displayed. The arrow indicates the direction along which x_t increases, and the ellipses present levels of the capital and the present value of cash flows where firms initiate dividend payouts.

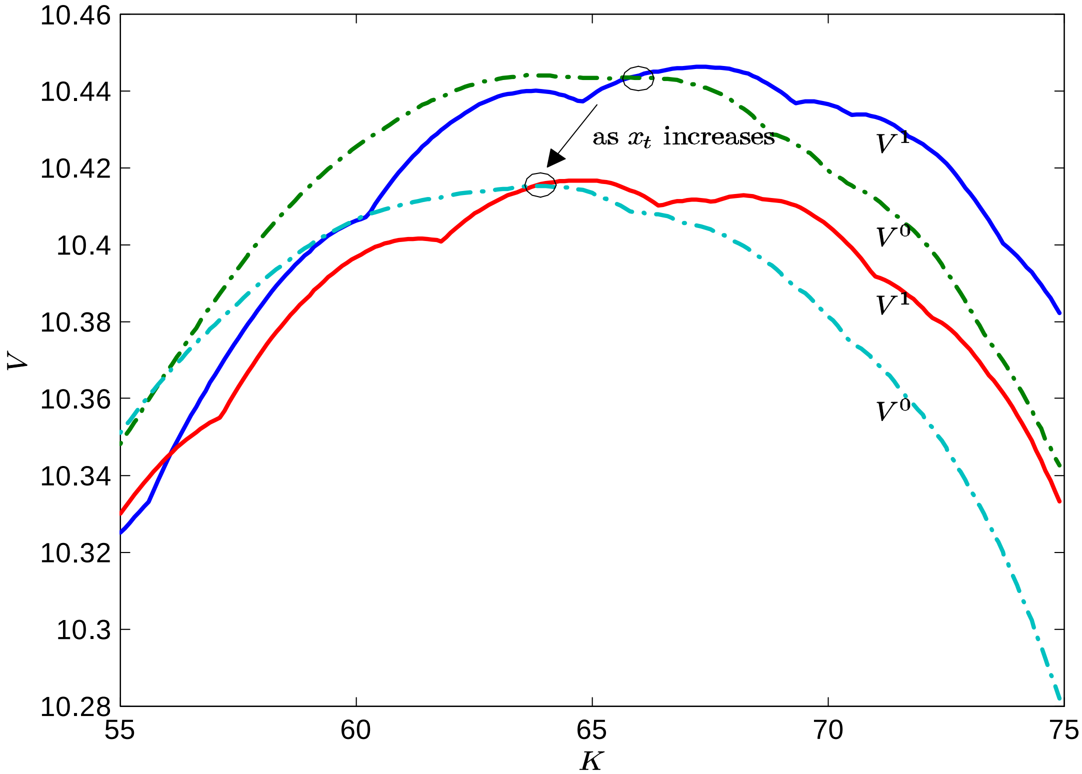


Figure 6: The Optimal Policy Functions Given x_t: Model $\mathcal{M}$

This figure plots the optimal policy and value functions on monthly frequency: Panel (A): I_{it}, Panel (B): I_{it}/K_{it}, Panel (C): D_{it}/K_{it}, and Panel (D): V_{it}, given the aggregate productivity x_t. Panels have a class of curves with angles at the initiation of dividend payouts, corresponding to different z_{it}. The arrow in each panel indicates the direction along which z_{it} increases.

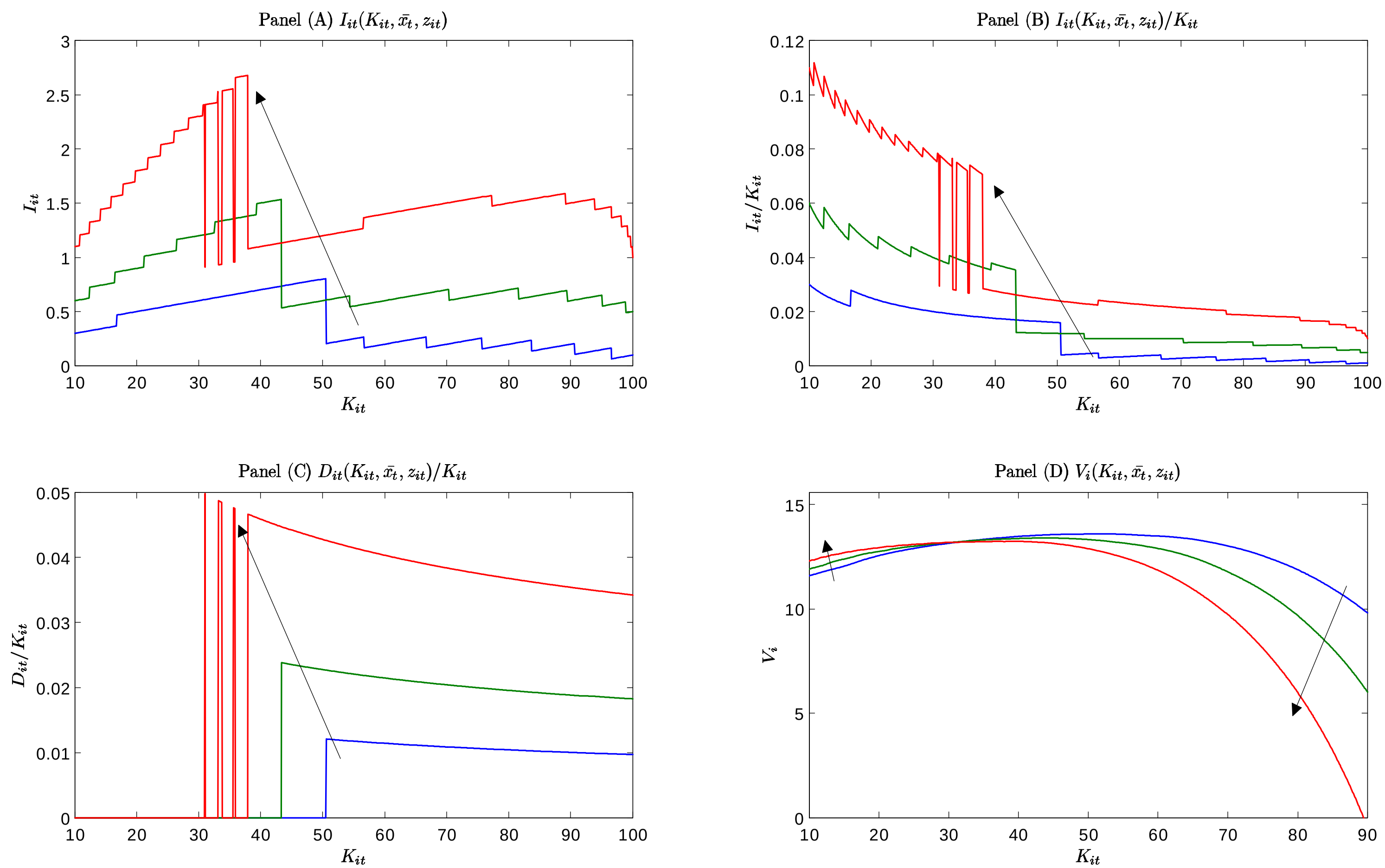


Figure 7: The Optimal Policy Functions Given z_{it}: Model $\mathcal{M}$

This figure plots the optimal policy and value functions on monthly frequency: Panel (A): I_{it}, Panel (B): I_{it}/K_{it}, Panel (C): D_{it}/K_{it}, and Panel (D): V_{it}, given the firm-specific productivity z_{it}. Panels have a class of curves with angles at the initiation of dividend payouts, corresponding to different x_t. The arrow in each panel indicates the direction along which x_t increases.

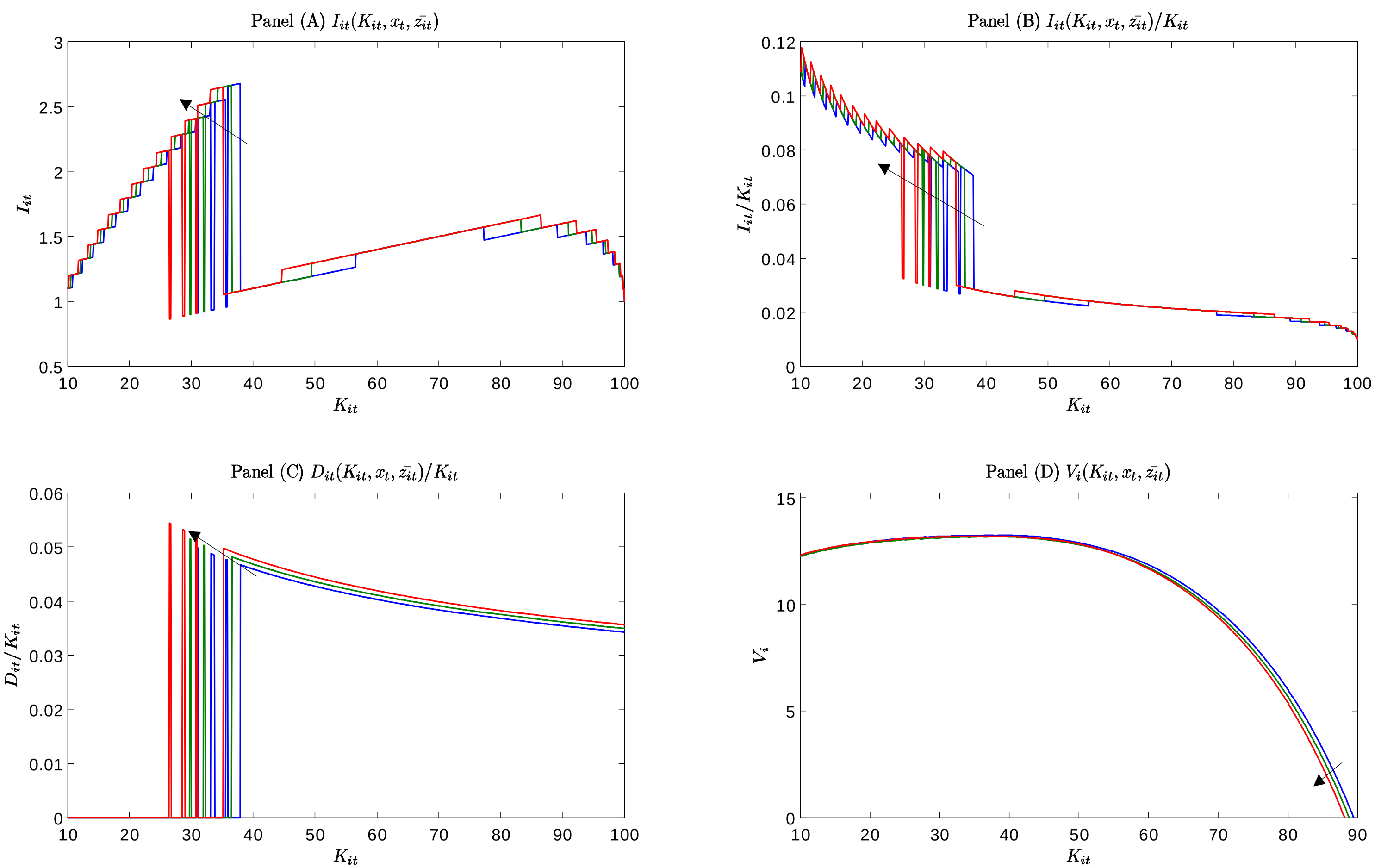


Figure 8: The Optimal Policy Functions Given x_t: Model $\mathcal{Q}$

This figure plots the optimal policy and value functions on quarterly frequency: Panel (A): I_{it}, Panel (B): I_{it}/K_{it}, Panel (C): D_{it}/K_{it}, and Panel (D): V_{it}, given the aggregate productivity x_t. Panels have a class of curves with angles at the initiation of dividend payouts, corresponding to different z_{it}. The arrow in each panel indicates the direction along which z_{it} increases.

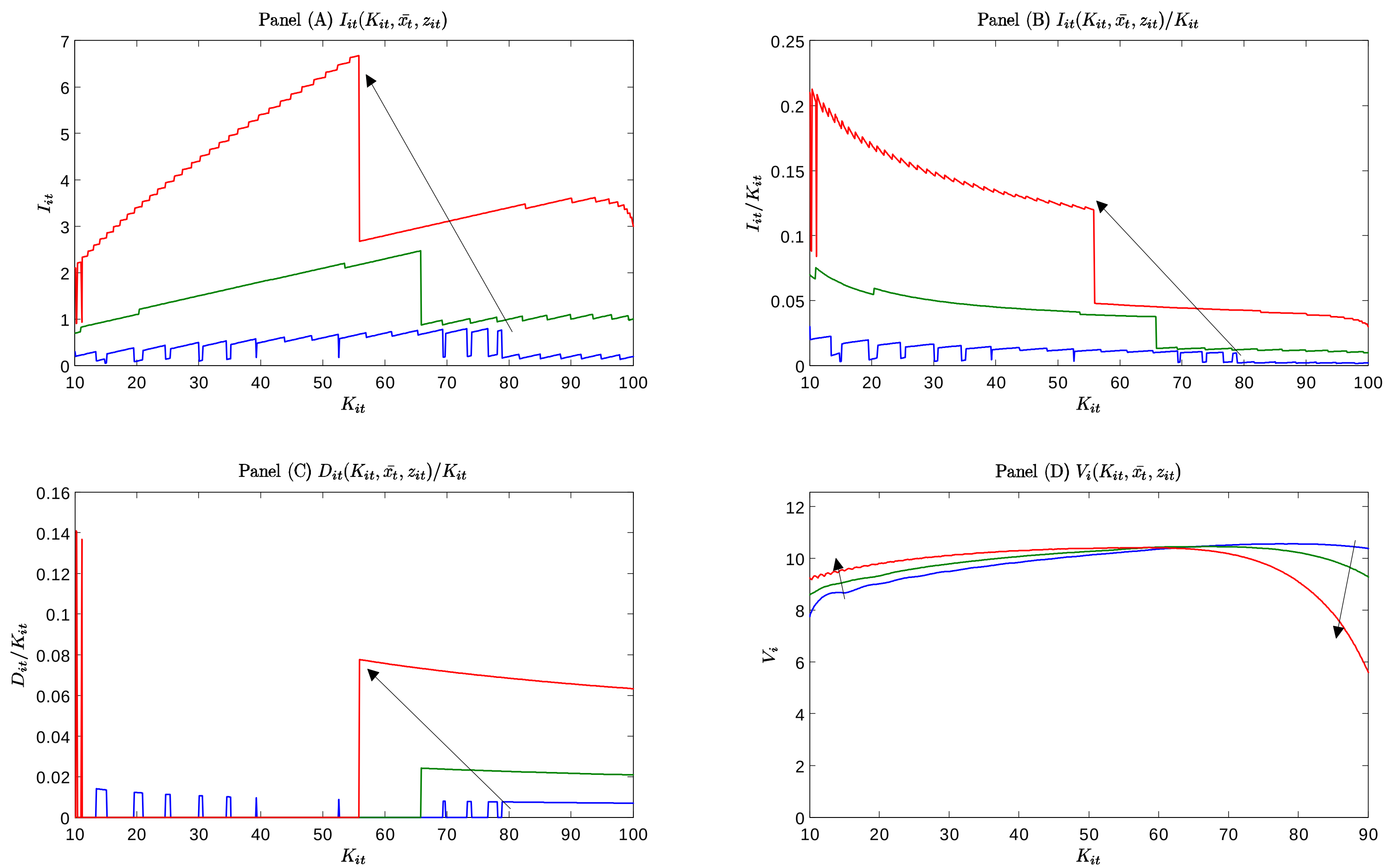


Figure 9: The Optimal Policy Functions Given z_{it}: Model $\mathcal{Q}$

This figure plots the optimal policy and value functions on quarterly frequency: Panel (A): I_{it}, Panel (B): I_{it}/K_{it}, Panel (C): D_{it}/K_{it}, and Panel (D): V_{it}, given the firm-specific productivity z_{it}. Panels have a class of curves with angles at the initiation of dividend payouts, corresponding to different x_t. The arrow in each panel indicates the direction along which x_t increases.

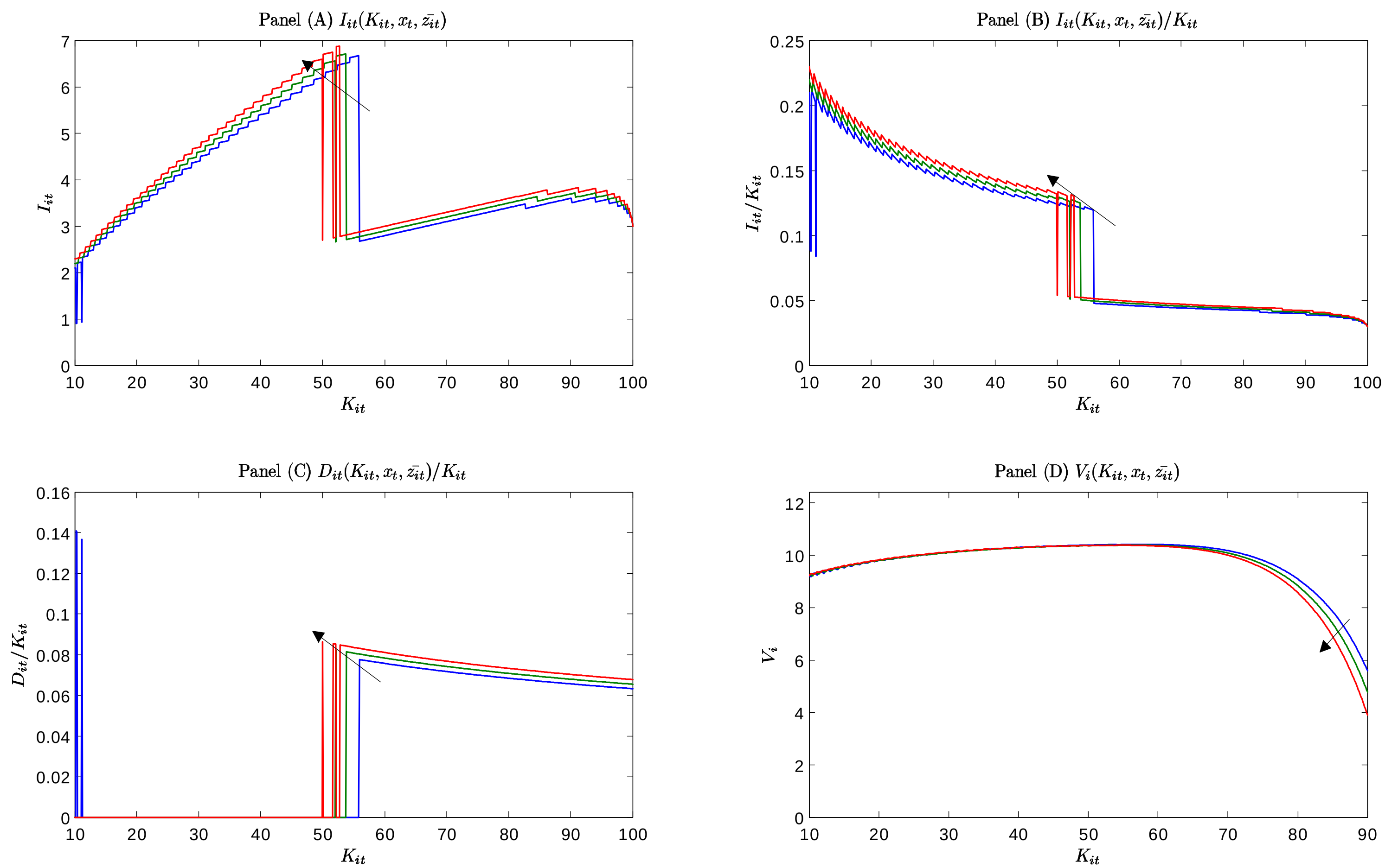


Figure 10: The Firm Life Cycle Theory of Dividends in SIC 3000-3099

This figure shows the characteristics of dividend payouts in SIC 3000-3099 (Rubber and Miscellaneous Plastic Products) as an example of the firm life cycle theory of dividends. Panel (A) shows investment ratio (I/CEQ), in which I is CAPX (capital expenditures), and CEQ is CEQ (common equity - total) in CRSP/COMPUSTAT merged data (CCM). Panel (B) shows investment ratio ($I/PPEGT$), in which $PPEGT$ is PPEGT (property, plant and equipment - total) in CCM. Panel (C) presents dividend on common equity ratios (DVC/CEQ), in which DVC is DVC (dividend common/ordinary) in CCM. Panel (D) also presents dividend on common equity ratios ($DVC/PPEGT$).

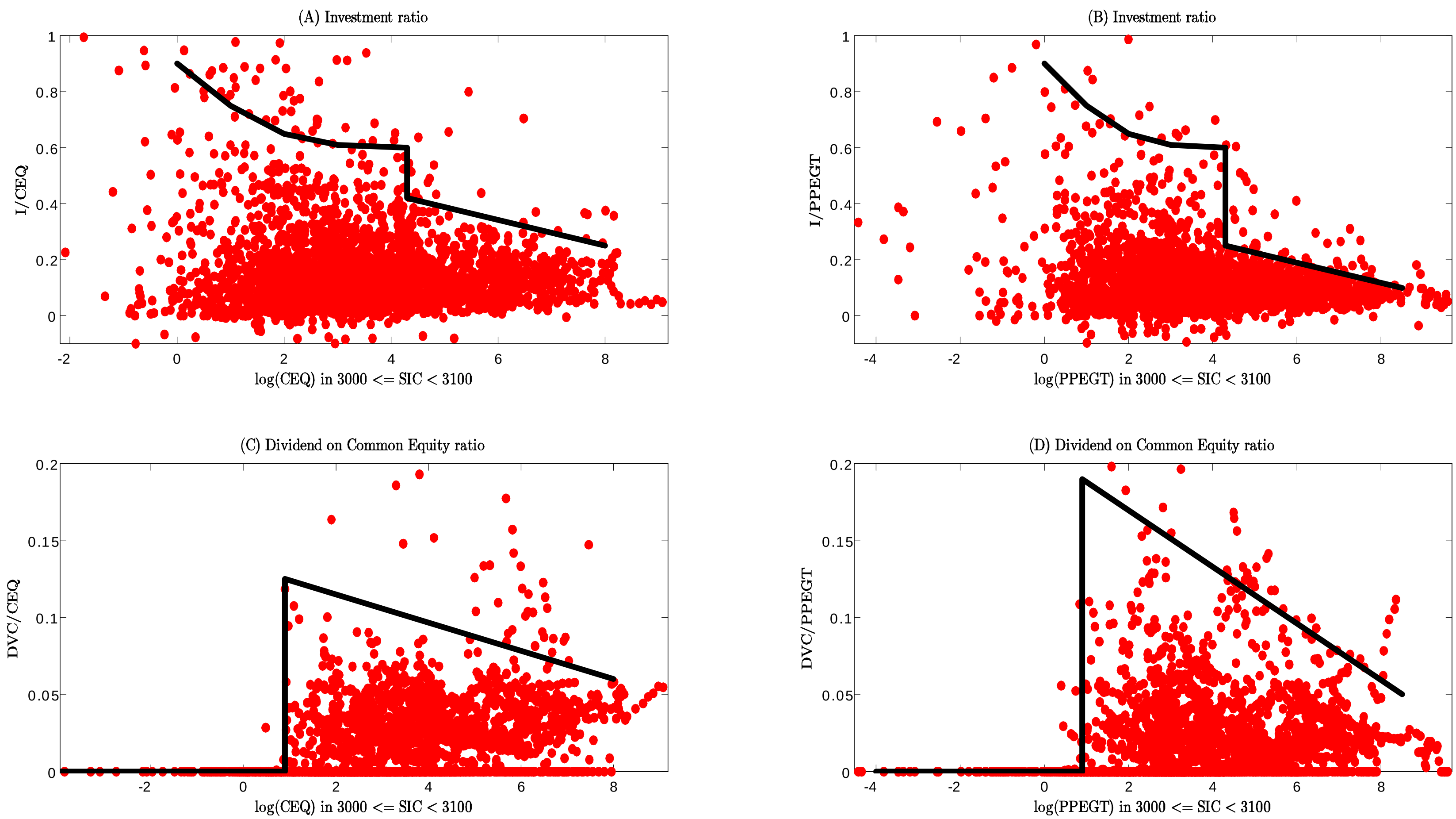


Figure 11: Countercyclical Time-Varying Risks

This figure plots the key moments of the stochastic pricing kernel, $M_{t,t+1}$. Panel (A) shows the conditional mean $(\mathbb{E}_t [M_{t,t+1}])$. Panel (B) and Panel (C) present the conditional volatility $(\sigma_t [M_{t,t+1}])$ and the conditional Sharpe ratio $(\sigma_t [M_{t,t+1}] / \mathbb{E}_t [M_{t,t+1}])$, respectively, based on Model $\mathcal{Q}$ with $\gamma = 3.5$.

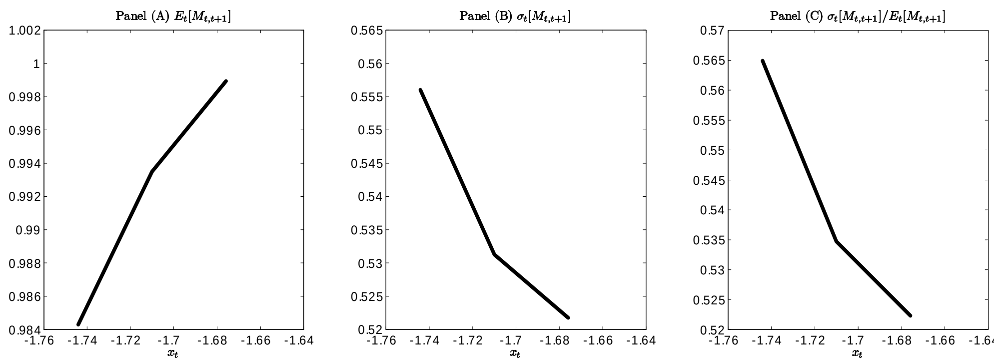


Figure 12: Conditional Expected Market Returns and β_M as Quantities of Risks

This figure plots the conditional expected market returns and the conditional quantities of risks β_Ms on monthly frequency (Panels (A) and (B)) and quarterly frequency (Panels (C) and (D)) according to CRRA (γ).

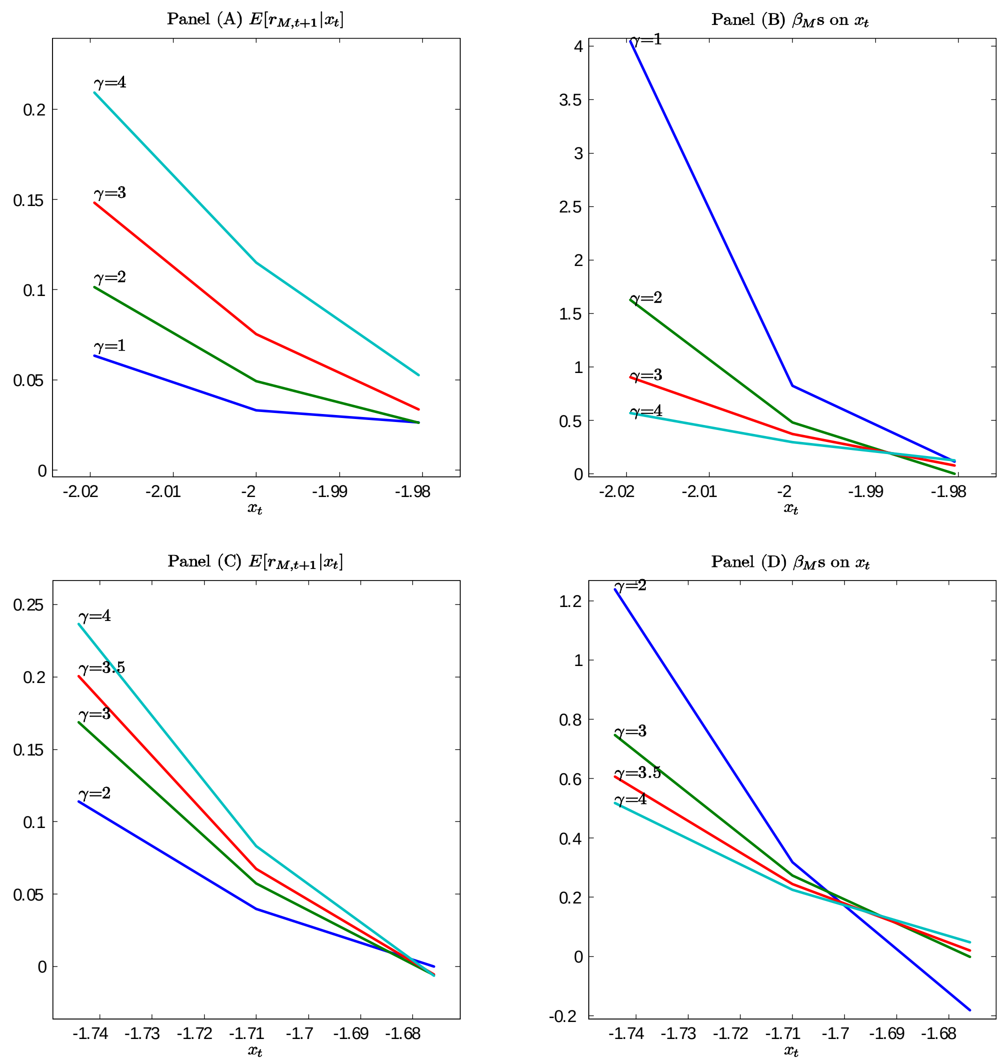


Figure 13: Time-Varying VMG and Value Spread in Quarterly Frequency

This figure plots expected VMGs and value spreads at quarterly frequency model with $\gamma = 3.5$. We follow the Fama-French method to pool VMG (value minus growth, previously HML) portfolios (Fama and French (2006)). Expected VMGs are the expected returns on the two value portfolios minus the expected returns on the two growth portfolios such that VMG$=\frac{\text{SV+BV}}{2} - \frac{\text{SG+BG}}{2}$ (see Table 3). The value spreads are computed by the difference between the $\log BE/ME$ of the value portfolio and the $\log BE/ME$ of the growth portfolio (Cohen, Polk, and Vuolteenaho (2003)). Panel (A) shows VMGs according to states, x_t. Panel (B) shows value spreads according to states, x_t.

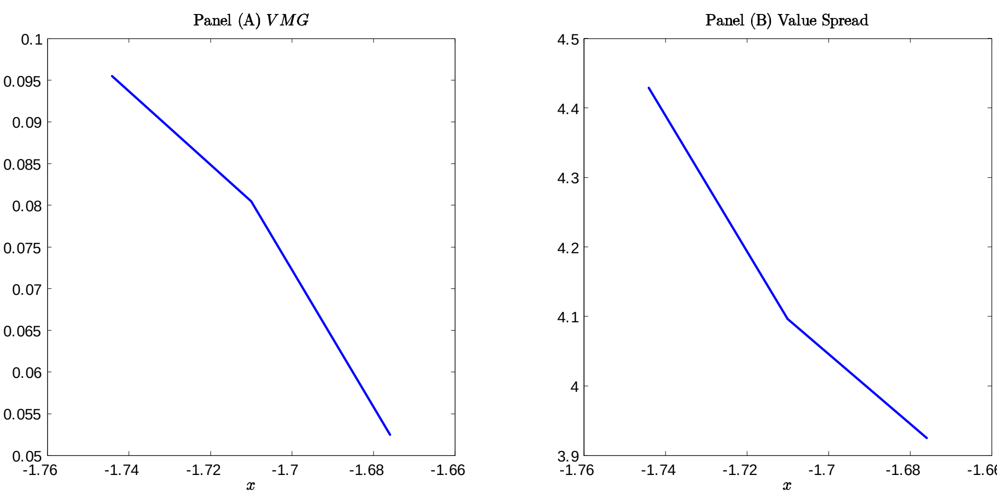


Figure 14: Monthly Expected Returns on 10 Portfolios Formed on BE/ME, Size, and fitted PPD

This figure plots monthly expected returns (%) on 10 portfolios formed on BE/ME, size, and fitted PPD ($\hat{a}$) according to CRRA. Panels (A), (B), and (C) are based on Model $\mathcal{M}$, and Panels (D), (E), and (F) are based on Model $\mathcal{Q}$. The break points for portfolios are determined by ranked values of BE/ME, size, and fitted PPD, and the pooling time of portfolios is time t. Line FF stands for the empirical data from Fama and French (1992). Lines GKZ and Z present the simulated data from Gomes, Kogan, and Zhang (2003) and Zhang (2005), respectively. The horizontal axes show the ascending order of 10 portfolios by BE/ME, size, and fitted PPD.