ESSAYS ON THE CROSS-SECTION OF RETURNS

Lu Zhang

A DISSERTATION

in

Finance

for the Graduate Group in Managerial Science and Applied Economics

Presented to the Faculties of the University of Pennsylvania in Partial

Fulfillment of the Requirements for the Degree of Doctor of Philosophy

2002

_____

Supervisor of Dissertation

_____

Graduate Group Chairperson

Dedication

To my parents, for teaching me the value of knowledge.

Acknowledgments

I am grateful to my advisor, Craig MacKinlay, for his kindness, support, and exceptional guidance, and to the rest of my dissertation committee — Andrew Abel, Amir Yaron, Leonid Kogan, and especially Joao Gomes — for their collective insight and encouragement. I would also like to express my gratitude to several other faculty — Michael Brandt, Domenico Cuoco, Gary Gorton, Michael Gibbons, Skander Van den Heuvel, Donald Keim, Nick Souleles, and Rob Stambaugh — for their roles in creating a supportive and exciting environment for learning.

ABSTRACT

ESSAYS ON THE CROSS-SECTION OF RETURNS

Lu Zhang

A. Craig MacKinlay

My dissertation aims at understanding the economic determinants of the cross-section of equity returns. It contains three chapters.

Chapter One constructs a dynamic general equilibrium production economy to explicitly link expected stock returns to firm characteristics such as firm size and the book-to-market ratio. Despite the fact that stock returns in the model are characterized by an intertemporal CAPM with the market portfolio as the only factor, size and book-to-market play separate roles in describing the cross-section of returns. However, these firm characteristics appear to predict stock returns only because they are correlated with the true conditional market beta. Moreover, quantitative analysis suggests that these cross-sectional relations can subsist even after one controls for a typical empirical estimate of market beta. This lends support to the view that the documented ability of size and book-to-market to explain the cross-section of stock returns is consistent with a single-factor conditional CAPM model.

Chapter Two asks whether firms' financing constraints are quantitatively important in explaining asset returns. It has two main findings. First, for a large class of theoretical models, financing constraints have a parsimonious representation amenable to empirical analysis. Second, financing frictions lower both the market Sharpe ratio and the correlation between the pricing kernel and returns. Consequently, they significantly worsen the performance of investment-based asset pricing models. These findings question whether

financing frictions are important for explaining the cross-section of returns and whether they provide a realistic propagation mechanism in several macroeconomic models.

Chapter Three proposes a novel economic mechanism underlying the value premium, the average return difference between value and growth stocks in the cross-section. The key element emphasized is the asymmetric adjustment cost of capital. During recessions, value firms face more difficulty than growth firms in downsizing capital, and hence their dividend streams fluctuate more with economic downturns. The upshot is that value stocks are more risky than growth stocks in bad times. An industry equilibrium model shows that this mechanism, when combined with a countercyclical market price of risk, goes a long way in generating a value premium that is quantitatively comparable to that observed in the data.

Contents

List of Tables

List of Figures

Chapter 1

Equilibrium Cross-Section of Returns

with Joao F. Gomes and Leonid Kogan

1.1 Introduction

The cross-sectional properties of stock returns have attracted considerable attention in

recent empirical literature in financial economics. One of the best known studies, by Fama

and French (1992), uncovers the relations between factors such as book-to-market ratio

and firm size and stock returns, which appear to be inconsistent with the standard Capital

Asset Pricing Model (CAPM). Despite their empirical success, these simple statistical

relations have proved very hard to rationalize and their precise economic source remains

a subject of debate.[1] The challenge posed by the Fama and French (1992) findings to

traditional structural models has created a significant hurdle to the understanding of

---

[1]Cochrane (1999), Campbell (2000) and Campbell, Lo and MacKinlay (1997) review the related literature. Various competing interpretations of observed empirical regularities include, among others, Berk (1995), Berk, Green and Naik (1999), Fama and French (1993, 1995, 1996), Jagannathan and Wang (1996), Kothari, Shanken, and Sloan (1995), Lakonishok, Shleifer, and Vishny (1994), Lettau and Ludvigson (1999), Liew and Vassalou (2000), Lo and MacKinlay (1988) and MacKinlay (1995).

more complex, dynamic properties of the cross-section of stock returns.

In this work we construct a stochastic dynamic general equilibrium one-factor model in which firms differ in characteristics such as size, book value, investment and productivity among others, thus establishing an explicit economic relation between firm level characteristics and stock returns. We show that the simple structure of our model provides a parsimonious description of the firm level returns and makes it a natural benchmark for interpreting many empirical regularities.

Our findings can be summarized as follows. First, we show that our one-factor equilibrium model can still capture the ability of book-to-market and firm value to describe the cross-section of stock returns. These relations can subsist after one controls for typical empirical estimates of conditional market β. This lends support to the view that the documented ability of size and book-to-market to explain the cross-section of stock returns is *not* necessarily inconsistent with a single-factor conditional CAPM model and provides a possible rationalization for the Fama and French (1992) findings. Second, we also establish a number of additional properties of the cross-section of stock returns with important implications for optimal dynamic portfolio choice. In particular, we find that cross-sectional dispersion in individual stock returns is related to the aggregate stock market volatility and business cycle conditions. In addition, we show that the size and book-to-market return premia are inherently conditional in their nature and likely countercyclical.

Our theoretical approach builds on the work of Berk, Green, and Naik (1999). These authors construct a two-factor partial equilibrium model based on ideas of time-varying risks to explain cross-sectional variations of stock returns associated with book-to-market and market value. They show that their calibrated model is able to capture several of the Fama and French (1992) findings. Our work differs along several important dimensions. First, ours is a single-factor model in which the conditional CAPM holds. We can then identify separate roles of size and book-to-market without appealing to multiple sources

2

of risk. Second, the simple structure of our model allows us to illustrate the role of β mismeasurement in generating the cross-sectional relations between the Fama and French's factors and stock returns. Finally, the general equilibrium nature of our model allows us to present a self-consistent account of the business cycle properties of firm level returns.

Our work is also related to a variety of recent papers that explore the asset pricing implications of production and investment in an equilibrium setting. Examples of this line of research include Bossaerts and Green (1989), Cochrane (1991 and 1996), Jermann (1998), Kogan (2000a and 2000b), Naik (1994), Rouwenhorst (1995) and Coleman (1997). To the best of our knowledge, however, ours is the first work aiming directly at explaining the cross-sectional variations of stock returns from a structural general equilibrium perspective.

The rest of the paper is organized as follows. Section 1.2 describes the model economy and its competitive equilibrium and derives an explicit analytical relation between the systematic risk of stock returns and firm characteristics. Sections 1.3 and 1.4 examine the quantitative implications of our model. Section 1.5 concludes.

1.2 The Model

In this section we develop a general equilibrium model with heterogeneous firms to characterize individual returns and link them to underlying firm characteristics. There are two types of agents: firms and households. We keep the household sector very standard, summarized by a single representative household which makes the optimal consumption and portfolio allocation decisions. The heart of the model is the production sector, where a continuum firms are engaged in production of the consumption good. Each firm operates a number of individual projects of different characteristics. This firm level uncertainty is crucial to obtain a non-degenerate equilibrium cross-sectional distribution of firms, a

3

necessary condition for our analysis in sections 1.3 and 1.4. Subsection 1.2.1 details the structure of the economy, while subsection 1.2.2 describes the equilibrium aggregate asset prices and establishes the link between systematic risk of stock returns and firm characteristics.

1.2.1 The Economy and the Competitive Equilibrium

Technology

Production of the consumption good (numeraire) in this economy takes place in basic productive units, which we label *projects*. These projects expire at a randomly chosen time, defined by an idiosyncratic Poisson process with common arrival rate δ. They have three individual features: scale, productivity, and cost.

Let $\mathcal{I}_t$ denote the set of all projects existing at time t and let i be the index of an individual project and s denote the time of creation, or vintage. We make two simplifying assumptions with respect to the scale of the project, k_{it}^s. First, the scale of a project is determined when the project is created and it remains fixed throughout the life of the project. Second, all projects of the same vintage have identical scale. Given these assumptions, and when there is no possibility of confusion, we will use only $k_i = k_{it}^s$ to denote the scale of project i created at time $s(i) \leq t$.

Project's productivity is driven by an exogenous stochastic process X_{it}, resulting in a flow of output at rate $X_{it}k_i$. Specifically, we define $X_{it} = \exp(x_t)\,\epsilon_{it}$, where x_t is a systematic, economy-wide productivity measure common for all projects, while ϵ_{it} is the idiosyncratic, project-specific component. Furthermore, we assume that x_t follows a linear mean-reverting process

$$dx_t = -\theta_x \left(x_t - \overline{x}\right) dt + \sigma_x dB_{xt} \tag{1.1}$$

and ϵ_{it} is driven by a square-root process

$$d\epsilon_{it} = \kappa(1 - \epsilon_{it})\, dt + \sigma_\epsilon \sqrt{\epsilon_{it}} dB_{it} \qquad (1.2)$$

where B_{xt} and B_{it} are standard Brownian motions.[2] Naturally we will assume that the idiosyncratic productivity shocks of all projects are independent of the economy-wide productivity shock, i.e., $dB_{xt}\, dB_{it} = 0$ for all i. We will place one further restriction on the correlation structure of the shocks below. Initial productivity of new projects is unobserved and drawn from the long-run distribution implied by (1.2).

While specific nature of processes (1.1) and (1.2) is convenient but not essential to our purposes, the assumption of mean-reversion in productivity shocks is very important. This assumption, however, is supported by both aggregate and cross-sectional evidence. At the aggregate level, mean-reversion implies that the growth rate of output is not exploding, which is consistent with standard findings in the economic growth literature (e.g., Kaldor (1963)). At the firm level, this assumption is required to obtain a stationary equilibrium distribution of firms. This is consistent with the cross-sectional evidence on firm birth and growth, suggesting that growth rates decline with age and size (e.g., Hall (1987) and Evans (1987)).

Finally, projects of the same vintage differ in their unit cost, measured in terms of consumption goods as e_{it}. Specifically, a potential new project i can be adopted at time s with investment cost of $e_{is}k_i$, where k_i is the scale of all new projects at time s.

Together, our assumptions about productivity and cost imply that all new projects

---

[2]The process in (1.1) is chosen to possess a stationary long-run distribution with constant instantaneous volatility, so that aggregate stock returns are not heteroscedastic by assumption. The idiosyncratic component in (1.2) follows a different type of process. It also has a stationary distribution, but it is heteroscedastic. Since our focus in this paper is on the systematic component of stock returns, such heteroscedasticity is not problematic. The advantage of (1.2) is that the conditional expectation of ϵ_{it} is an exponential function of time and a linear function of the initial value ϵ_{i0}, which facilitates computation of individual stock prices. An additional advantage of this process is that its unconditional mean is independent of κ and σ_ϵ, which simplifies the calibration.

are *ex-ante* identical in terms of expected future output, differing only in their cost. As we will see below, these assumptions guarantee that individual investment decisions can be aggregated into a stochastic growth model with adjustment costs. In addition to its computational appeal, this feature is useful in providing a realistic setting for aggregate asset pricing (e.g., Jermann (1998)).

Firms

Firms in our economy are infinitely lived. We assume that the set of firms $\mathcal{F}$ is exogenously fixed and let f be the index of an individual firm. Each firm owns a finite number of individual projects. While we do not explicitly model entry and exit of firms, a firm can have zero projects, thus effectively exiting the market, and a new entrant can be viewed as a firm that begins operating its first project.

We make a further assumption that the idiosyncratic productivity shocks ϵ_{it} are firm-specific. Formally, let $\mathcal{I}_{ft}$ denote the set of projects owned by firm f at time t and let $f(i)$ denote the index of the firm owning project i. If (ongoing) projects i and j belong to the same firm, then dB_{it} and dB_{jt} are perfectly correlated, otherwise they are independent. Mathematically,

$$dB_{it}\, dB_{jt} = \begin{cases} dt, & j \in \mathcal{I}_{f(i),t} \\ 0, & j \notin \mathcal{I}_{f(i),t} \end{cases} \tag{1.3}$$

Firms are financed entirely by equity and outstanding equity of each firm is normalized to one share. We denote individual firm's stock price by V_{ft}. Stocks represent claims on the dividends, paid by firms to shareholders, and equal to the firm's output net of investment costs.[3] We specify the shareholders' problem below.

[3]Instead of assuming that investment is financed by retaining earnings, one can make an equivalent assumption that investment is financed by new equity issues. The exact form of financing has no effect on the firm market value.

While they do not control the scale or productivity of their projects, firms do make *investment* decisions by selecting which new projects to operate. Specifically, firms are presented with potential new projects over time. If a firm decides to invest in a new project, it must incur the required investment cost, which in turn entitles it to the permanent ownership of the project. These investment decisions are irreversible and investment cost cannot be recovered at a later date.[4] If the firm decides not to invest in a project, the project disappears from the economy.

The arrival rate of new projects is independent of the individual firm's past investment decisions. Specifically, all firms have an equal probability of receiving a new project in every period. This assumption guarantees that large firms do not adopt more projects than small firms, which is again consistent with the evidence on firm size and growth.[5] Moreover, it also implies that the decision to accept or reject a project has no effect on the individual firm's future investment opportunities.

Hence, current investment decisions do not depend on the nature of a specific firm — they are determined exclusively by the cost of new projects relative to the present value of projects' cash flows. Given these assumptions, the optimal investment decision of a firm faced with project i at time s is to invest if

$$V_{it}^a = E_t \left[\int_0^\infty e^{-\lambda s} M_{t,t+s} \left(e^{-\delta s} k_i X_{t+s} \right) ds \right] \geq e_{it} k_i \tag{1.4}$$

where V_{it}^a is the net present value of the future stream of cash flows associated with the project and $M_{t,t+s}$ is the stochastic discount factor between periods t and $t + s$, equal to the intertemporal marginal rate of substitution of the representative household in equilibrium.[6] Note also that we have used the fact that the idiosyncratic productivity

---

[4]Otherwise the assumption that initial productivity is unobserved would not matter.

[5]All that is required is that project arrival is less than proportional to firm size. This is the simplest way of meeting this requirement and it seems the natural one to start with. Results for alternative assumptions are substantially similar and are available upon request.

[6]Our treatment of the firm's problem can be related to the Arbitrage Pricing Theory of Ross (1976).

component ϵ_{it} is independent of all other processes in the economy and that, for any new project, ϵ_{it} is drawn from the steady state distribution of process (1.2). Hence, $E_t [X_{t+s} \epsilon_{it+s}] = E_t [\epsilon_{it+s}] E_t [X_{t+s}] = E_t [X_{t+s}]$.

Proposition 1 *(Optimal firm investment) A new project i is adopted if and only if*

$$e_{it} \le \overline{e}_t = \overline{e}(x_t)$$

Proof Given the stochastic process for aggregate productivity shocks (1.1), it follows that the present value of project's cash flows per unit production scale equals

$$\frac{V_{it}^a}{k_i} = E_t \left[\int_0^\infty e^{-\lambda s} M_{t,t+s} \left(e^{-\delta s} X_{t+s} \right) ds \right]$$

which in turn depends only on the current state of the economy x_t. Equation (1.4) implies then that a new project is adopted if and only if

$$e_{it} \le V_{it}^a(x_t)/k_i = \overline{e}_t = e(x_t)$$

∎

Proposition 1 establishes a simple, but crucial, property that optimal investment decisions by firms at any time t are independent of the firms' identity and only rely on the unit cost of new projects. Specifically, firms adopt new projects with unit cost below the threshold $\overline{e}(x)$, which is only a function of the aggregate state variable. Note that this result hinges on the convenient assumption that projects are ex-ante identical in their

---

Even though cash flows of individual projects and firms are not spanned by a small number of traded assets, their idiosyncratic components are perfectly diversifiable. Therefore, the only stochastic components of cash flows and returns that are priced by the market are those associated with market-wide risk factors, which are common to all firms. In our model, x_t is the only systematic risk factor, which in equilibrium is spanned by the market portfolio. Thus, the associated risk premium is uniquely determined by absence of arbitrage. Alternatively, in the framework of a representative household, consumption-based asset pricing model, the aggregate consumption process can be used as a single systematic risk factor which is sufficient for pricing all risky assets (e.g., Breeden (1979)).

productivity and allows for the simple aggregation results below.

The value of the firm can then be viewed as a sum of two components, the present value of output from existing projects and the present value of dividends (output net of investment) from future projects. Using the terminology from Berk et al. (1999), the former component represents the value of *assets-in-place*, V_{ft}^a, while the second can be interpreted as the value of *growth options*, V_{ft}^o. We can then compute the value of a firm's stock as a sum of these two components

$$V_{ft} = V_{ft}^a + V_{ft}^o \tag{1.5}$$

where the value of assets in place can be constructed as

$$V_{ft}^a = \sum_{i \in \mathcal{I}_{ft}} V_{it}^a \tag{1.6}$$

Finally, it is useful for future use to define the book value of a firm as the sum of book values of the firm's (active) individual projects

$$B_{ft} = \sum_{i \in \mathcal{I}_{ft}} e_{i,s(i)} k_{it}^{s(i)}$$

and the book value of a project is defined as the associated investment cost $e_{is} k_{it}^s$.

Heterogeneity and Aggregation

To facilitate aggregation, we assume that there exists a large number (a continuum) of firms in the economy. In our informal construction we appeal to the law of large numbers, which simplifies the analysis and clarifies economic intuition, albeit at a cost of some mathematical rigor. Thus, one might view the results based on the law of large numbers

9

as an approximation to an economy with a very large number of firms.[7]

Let $\int_{\mathcal{I}_t} \cdot \, di$ and $\int_{\mathcal{F}} \cdot \, df$ denote aggregation operators over projects and firms respectively. The aggregate scale of production in the economy, K_t, is

$$K_t \equiv \int_{\mathcal{I}_t} k_i \, di = \int_{-\infty}^{t} k_{it}^{s} \left(\int_{\mathcal{I}_t} \chi_{\{i:\, s(i) \in [\tau, \tau + d\tau)\}} \, di \right) ds$$

where $\chi_{\{\cdot\}}$ denotes the indicator function and $\int_{\mathcal{I}_t} \chi_{\{i:\, s(i) \in [\tau, \tau+d\tau)\}} \, di$ is the number (measure) of projects created during $[\tau, \tau + d\tau)$ that remain in existence at time t. Similarly, aggregate output Y_t is given by

$$
\begin{aligned}
Y_t &= \int_{\mathcal{I}_t} X_{it} k_i \, di = \int_{-\infty}^{t} k_{it}^{s} \left(\int_{\mathcal{I}_t} X_{it}\, \chi_{\{i:\, s(i) \in [\tau, \tau+d\tau)\}} di \right) ds \\
&= \exp\left(x_t\right) \int_{-\infty}^{t} k_{it}^{s} \left(\int_{\mathcal{I}_t} \chi_{\{i:\, s(i) \in [\tau, \tau+d\tau)\}} \epsilon_{it} \, di \right) ds \\
&= \exp(x_t) \int_{-\infty}^{t} k_{it}^{s} \left(\int_{\mathcal{I}_t} \chi_{\{i:\, s(i) \in [\tau, \tau+d\tau)\}} \, di \right) ds = \exp(x_t) K_t
\end{aligned}
\tag{1.7}
$$

where the fourth equality follows from the law of large numbers, since by (1.2) random variables ϵ_{it}s are identically distributed with unit mean and are independent across a continuum of firms, with each firm owning a finite number of projects. Equation (1) is consistent with our interpretation of x_t as the aggregate productivity shock.

New potential projects are continuously arriving in the economy. To ensure balanced growth, we assume that the arrival rate of new projects is proportional to the total scale of existing projects in the economy K_t and independent of project unit cost. Formally, the arrival rate (measured by production scale) of new projects with cost less than e_t equals $ZK_t e_t$. Alternatively, $ZK_t e_t dt$ is the total scale of projects with the cost parameter less than e_t arriving between t and $t + dt$. The parameter Z governs the quality of the

---

[7]Feldman and Gilles (1985) formalize the law of large numbers in economies with countably infinite numbers of agents by aggregating with respect to a finitely-additive measure over the set of agents. Judd (1985) demonstrates that a measure and the corresponding law of large numbers can be meaningfully introduced for economies with a continuum of agents.

investment opportunity set. Given our definition of the arrival rate, the total scale of projects in the economy evolves according to

$$dK_t = -\delta K_t dt + Z K_t \bar{e}_t dt \qquad (1.8)$$

where δ is the rate at which existing projects expire. The aggregate investment spending, I_t, is then given by

$$I_t = I\left(\bar{e}_t\right) \equiv \int_0^{\bar{e}_t} e_{it} Z K_t de_{it} = \frac{1}{2} Z K_t \bar{e}_t^2 \qquad (1.9)$$

Aggregate dividends are defined as the aggregate output net of aggregate investment, or

$$D_t = Y_t - I_t \qquad (1.10)$$

In addition, we define the value of the aggregate stock market V_t, which is the market value of a claim on aggregate dividends, as

$$V_t = \int_{\mathcal{F}} V_{ft} \, df \qquad (1.11)$$

Finally, given (1.10) and (1.11) we can define the process for cumulative aggregate stock returns as

$$\frac{dR_t}{R_t} = \frac{dV_t + D_t dt}{V_t} \qquad (1.12)$$

Households

There is a single consumption good in the economy, which is produced by the firms. The economy is populated by identical competitive households, who derive utility from the consumption flow C_t. The entire population can then be modeled as a single representative household. We assume that this household has standard time-separable isoelastic

preferences:

$$E_0 \left[\frac{1}{1-\gamma} \int_0^\infty e^{-\lambda t} C_t^{1-\gamma} dt \right] \tag{1.13}$$

Households do not work and derive income from accumulated wealth only.[8] We let W_t denote the individual wealth at time t. Financial markets in our model consist of risky stocks and an instantaneously riskless bond in zero net supply that earns a rate of interest r_t. Financial markets are perfect: there are no frictions and no constraints on short sales or borrowing.

The representative household then maximizes her expected utility of consumption (1.13), subject to the constraints

$$dW_t = -C_t \, dt + W_{bt} r_t \, dt + W_{st} \frac{dR_t}{R_t} \tag{1.14}$$

$$W_t = W_{bt} + W_{st} \tag{1.15}$$

$$W_t \geq 0 \tag{1.16}$$

where W_{bt} and W_{st} is the amount of wealth invested in the bond and stocks, respectively.[9] The returns processes on bonds, r_t, and stocks, R_t, are taken as exogenous by households and will be determined in equilibrium. The nonnegative-wealth constraint (1.16) is used to rule out arbitrage opportunities, as shown in Dybvig and Huang (1989).[10]

---

[8]Since labor is not productive, this assumption is innocuous.

[9]We are assuming that households invest directly in the aggregate stock market portfolio. Combined with the assumption that firms' value is computed using the economy-wide stochastic discount factor to discount their dividends, this formulation is not restrictive and allowing households to invest in individual securities would lead to identical implications for equilibrium prices and policies.

[10]To make sure that the wealth process is well defined by (1.14), we assume that both the consumption policy and the portfolio policy are progressively measurable processes, satisfying standard integrability conditions:

$$\int_0^{\tau_n} C_t + \left| W_{bt} r_t + W_{st} \frac{E_t \, [dR_t]}{R_t \, dt} \right| dt < \infty$$

$$\int_0^{\tau_n} \left\langle W_{st} \frac{dR_t}{R_t}, W_{st} \frac{dR_t}{R_t} \right\rangle < \infty$$

for a sequence of stopping times $\tau_n \nearrow \infty$, where $\langle \cdot, \cdot \rangle_t$ denotes the quadratic variation process.

The Competitive Equilibrium

With the description of the economic environment complete we are now in a position to state the definition of the competitive equilibrium.

Definition 1 *(Competitive equilibrium) A competitive equilibrium is summarized by stochastic processes for optimal household decisions $C_t^\star$, $W_{bt}^\star$, $W_{st}^\star$, and firm investment policy $\overline{e}_t^\star$, such that*

(a) Optimization

(i) Given security returns, households maximize their expected utility (1.13), subject to constraints (1.14–1.16);

(ii) Given the stochastic discount factor

$$M_{t,t+s} = e^{-\lambda s} \left(\frac{C_t^\star}{C_{t+s}^\star} \right)^\gamma$$

firms maximize their market value (10).

(b) Equilibrium

(i) Goods market clears:

$$C_t^\star = D_t = Y_t - I_t \tag{1.17}$$

(ii) Stock market clears:

$$W_{st}^\star = V_t = \int_{\mathcal{F}} V_{ft}\, df \tag{1.18}$$

(iii) Bond market clears:

$$W_{bt}^\star = 0 \tag{1.19}$$

∎

The following proposition establishes that the optimal policies $\overline{e}_t^\star$ and $C_t^\star$ can be characterized as the solution to a system of one differential equation and one algebraic equation.

Proposition 2 *(Equilibrium allocations) The competitive equilibrium allocations of consumption $C_t^\star$ and investment $\overline{e}_t^\star$ can be computed by solving the equations*

$$\overline{e}^\star(x) = [c^\star(x)]^\gamma \, p(x) \tag{1.20}$$

and

$$c^\star(x) = \exp(x) - \frac{1}{2} Z \left[\overline{e}^\star(x)\right]^2 \tag{1.21}$$

where function $p(x)$ satisfies

$$\frac{\exp(x)}{[c^\star(x)]^\gamma} = \left[\lambda + (1-\gamma)\delta + \gamma Z \overline{e}^\star(x)\right] p(x) + \theta_x \left(x - \overline{x}\right) p'(x) - \frac{1}{2}\sigma_x^2 p''(x) \tag{1.22}$$

and

$$\overline{e}_t^\star = \overline{e}^\star(x_t)$$

$$C_t^\star \equiv c^\star(x_t)\, K_t$$

Proof See Appendix 1.5. ∎

1.2.2 Asset Prices

With the optimal allocations computed we can now easily characterize the asset prices in the economy, including the risk-free interest rate and both the aggregate and firm-level stock prices.

Aggregate Prices

The following proposition summarizes the results for the equilibrium values of the risk-free rate and the aggregate stock market value.

Proposition 3 *(Equilibrium asset prices) The instantenous risk-free interest rate is determined by:*

$$r_t = -\frac{E_t[dM_{t,t+dt} - 1]}{dt} = \lambda + \gamma \left[Z\overline{e}^\star (x_t) - \delta \right] + \gamma \frac{[A(c^\star(x_t)]}{c^\star (x_t)} - \frac{1}{2}\gamma(\gamma + 1)\sigma_x^2 \left[\frac{c^\star (x_t)'}{c^\star (x_t)} \right]^2$$

(1.23)

where $A(c(x))$ satisfies

$$A(c(x)) = -\theta_x \left(x - \overline{x} \right) c' \left(x \right) + \frac{1}{2}\sigma_x^2 c'' \left(x \right)$$

The aggregate stock market value, V_t, can then be computed as

$$V_t = E_t \left[\int_0^\infty e^{-\lambda s} \left(\frac{C_t^\star}{C_{t+s}^\star} \right)^\gamma C_{t+s}^\star \, ds \right] = (c_t^\star)^\gamma \, \psi \left(x_t \right) K_t$$

(1.24)

where function $\psi \left(x \right)$ satisfies the differential equation

$$\lambda\psi \left(x \right) = [c^\star(x)]^{1-\gamma} + (1 - \gamma) \left[Z\overline{e}^\star(x) - \delta \right] \psi \left(x \right) - \theta_x(x - \overline{x})\psi' \left(x \right) + \frac{1}{2}\sigma_x^2\psi'' \left(x \right)$$

Proof See Appendix 1.5. ∎

While the exact conditions are somewhat technical, the intuition behind them is quite simple. As we would expect, the instantenous risk-free interest rate is completely determined by the equilibrium consumption process of the representative household, and its implied properties for the stochastic discount factor. Also, the aggregate stock market value represents a claim on the the future stream of aggregate dividends paid out by firms. In equilibrium, however, these must equal the consumption of the representative

household.

In addition to the definition above, value of the stock market can also be viewed as a sum of two components, the present value of output from existing projects and the present value of dividends (output net of investment) from all future projects. The value of assets-in-place is given by

$$V_t^a = E_t \left[\int_0^\infty e^{-\lambda s} \left(\frac{C_t^\star}{C_{t+s}^\star} \right)^\gamma \left(\int_{\mathcal{I}_t} X_{it+s} e^{-\delta s} k_i \, di \right) ds \right] \qquad (1.25)$$

Using arguments similar to (1), we can restate this as

$$V_t^a = K_t E_t \left[\int_0^\infty e^{-(\lambda+\delta)s} \left(\frac{C_t^\star}{C_{t+s}^\star} \right)^\gamma \exp\left(x_{t+s} \right) ds \right] = K_t \left(c_t^\star \right)^\gamma p\left(x_t \right) \qquad (1.26)$$

where $p\left(x_t \right)$ is defined by (1.22) above. By definition then, the value of aggregate growth options can be constructed as

$$V_t^o = V_t - V_t^a \qquad (1.27)$$

Firm-Level Stock Prices

Valuation of individual stocks is straightforward once the aggregate market value is computed. First, note that as we have seen above, the value of a firm's stock is the sum of assets-in-place and growth options, where the value of assets-in-place is the sum of present values of output from all projects currently owned by the firm. The value of an individual project i is given by the following Proposition.

Proposition 4 *(Project valuation) The present value of output of a project i is given by*

$$V_{it}^a = E_t \left[\int_0^\infty e^{-\lambda s} \left(\frac{C_t^\star}{C_{t+s}^\star} \right)^\gamma \left(e^{-\delta s} k_i X_{it+s} \right) ds \right] = \frac{k_i}{K_t} \left[\widetilde{V}_t^a \left(\epsilon_{it} - 1 \right) + V_t^a \right] \qquad (1.28)$$

where $\widetilde{V}_t^a$ is defined as

$$\widetilde{V}_t^a \equiv K_t \, E_t \left[\int_0^\infty e^{-(\lambda+\delta+\kappa)s} \left(\frac{C_t^\star}{C_{t+s}^\star} \right)^\gamma \exp\left(x_{t+s}\right) ds \right]$$

Proof See Appendix 1.5. ∎

Given the result in Proposition 4, the value of assets in place for the firm, V_{ft}^a, can be constructed as

$$V_{ft}^a = \int_{\mathcal{I}_{ft}} \frac{k_i}{K_t} \left[\widetilde{V}_t^a \left(\epsilon_{it} - 1 \right) + V_t^a \right] di \tag{1.29}$$

Now since future projects are distributed randomly across the firms with equal probabilities, all firms will derive the same value from growth options. Clearly then this implies that the value of growth options of each firm, V_{ft}^o, equals

$$V_{ft}^o = \frac{1}{\int_{\mathcal{F}} 1 \, df} V_t^o \tag{1.30}$$

We can then join these two components to obtain the total value of the firm, V_{ft}, as

$$V_{ft} = \int_{\mathcal{I}_{ft}} \frac{k_i}{K_t} \left[\widetilde{V}_t^a \left(\epsilon_{it} - 1 \right) + V_t^a \right] di + \frac{1}{\int_{\mathcal{F}} 1 \, df} V_t^o \tag{1.31}$$

By relating individual firm value to market aggregates, the decomposition (1.31) is extremely useful as it implies that the instantaneous market betas of individual stock returns can also be expressed as a weighted average of market βs of three *economy-wide* variables, V_t^a, $\widetilde{V}_t^a$, and V_t^o. Proposition 5 formally establishes this property.

Proposition 5 *(Market betas of individual stocks) Firm market βs are described by*

$$\beta_{ft} = \widetilde{\beta}_t^a + \frac{V_{ft}^o}{V_{ft}} \left(\beta_t^o - \widetilde{\beta}_t^a \right) + \frac{K_{ft}}{V_{ft}} \left(\frac{K_t}{V_t^a} \right)^{-1} \left(\beta_t^a - \widetilde{\beta}_t^a \right) \tag{1.32}$$

where

$$K_{ft} = \int_{\mathcal{I}_{ft}} k_i \, di$$

and

$$\beta_t^a = \frac{\partial \log (V_t^a)/\partial x}{\partial \log (V_t)/\partial x}, \qquad \widetilde{\beta}_t^a = \frac{\partial \log \left(\widetilde{V}_t^a\right)/\partial x}{\partial \log (V_t)/\partial x}, \qquad \beta_t^o = \frac{\partial \log (V_t^o)/\partial x}{\partial \log (V_t)/\partial x} \qquad (1.33)$$

Proof Since the market beta of a portfolio of assets is a value-weighted average of betas of its individual components, the expression for the value of the firm (1.31) implies that

$$\begin{aligned}
\beta_{ft} &= \left(1 - \frac{V_{ft}^o}{V_{ft}}\right) \beta_{ft}^a + \frac{V_{ft}^o}{V_{ft}} \beta_t^o \\
&= \left(1 - \frac{V_{ft}^o}{V_{ft}}\right) \left((1 - \pi_{ft})\widetilde{\beta}_t^a + \pi_{ft}\beta_t^a\right) + \frac{V_{ft}^o}{V_{ft}} \beta_t^o
\end{aligned}$$

where

$$\pi_{ft} = \frac{K_{ft}}{V_{ft}^a} \left(\frac{K_t}{V_t^a}\right)^{-1}$$

Simple manipulation then yields (1.32). ∎

Stock Returns and Firm Characteristics

Proposition 5 is extremely important. It shows that the weights on the "aggregate" betas, β_t^a, $\widetilde{\beta}_t^a$, and β_t^o, depend on economy-wide variables like K_t/V_t^a, and V_t^o, but also, and more importantly on firm-specific characteristics such as the size, or value, of the firm, V_{ft}, and the ratio of the firm's production scale to its market value, K_{ft}/V_{ft}.

The second term in (1.32) creates a relation between size and β, as the weight on the beta of growth options, β_t^o, depends on the value of the firm's growth options relative to its total market value. Firms with small production scale derive most of their value from growth options and their betas are close to β_t^o. Since all firms in our economy have

identical growth options, the cross-sectional dispersion of betas due to the loading on β_t^o is captured entirely by the size variable V_{ft}. Large firms, on the other hand, derive a larger proportion of their value from assets in place, therefore their betas are close to a weighted average of β_t^a and $\widetilde{\beta}_t^a$.

The last term in (1.32) also shows that part of the cross-sectional dispersion of market betas is explained by the firm-specific ratio of the scale of production to the market value, K_{ft}/V_{ft}, captured empirically to certain extent by the firm's book-to-market ratio.[11] To see the intuition behind this result consider two firms, A and B, with the same market value. Assume that firm A has larger scale of production but lower productivity than B. As a result, the two stocks would differ in their systematic risk due to the differences in the distribution of cash flows from the firms' existing projects. By assumption, such a difference is not reflected in the firms' market value, but it would be captured by the ratio K_{ft}/V_{ft}. Thus, while firm size captures the component of firm's systematic risk attributable to its growth options, the book-to-market ratio serves as a proxy for risk of existing projects.

Note that in this model the cross-sectional distribution of expected returns is determined entirely by the distribution of market βs, since returns on the aggregate stock market are perfectly correlated with the consumption process of the representative household (and hence the stochastic discount factor, e.g., Breeden (1979)). Thus, if conditional market βs were measured with perfect precision, no other variable would contain additional information about the cross-section of returns.

However, equation (1.32) implies that if for any reason market βs were mismeasured (e.g. because the market portfolio is not correctly specified), then firm-specific variables like firm size and book-to-market ratios could appear to predict the cross-sectional distribution of expected stock returns simply because they are related to true conditional βs. In

---

[11]The ratio K_{ft}/V_{ft} can also be approximated by other accounting variables, e.g., by the earnings-to-price ratio.

section 1.4 we generate an example within our artificial economy of how mismeasurement of βs can lead to a significant role of firm characteristics as predictors of returns.

1.3 Aggregate Stock Returns

In this section we evaluate our model's ability to reproduce key qualitative and quantitative features of empirical data. While it is not the objective of this paper, it seems appropriate to ensure that the model is reasonably consistent with the well documented aggregate findings before examining its cross-sectional implications. Thus, our methodology follows the approach of Kydland and Prescott (1982) and Long and Plosser (1983). First, we calibrate the model parameters using the unconditional moments of aggregate stock returns and the moments of the aggregate consumption process. We then provide evidence on other aggregate-level properties of the model regarding the predictability of aggregate stock returns by the book-to-market ratio documented by Pontiff and Schall (1998).

1.3.1 Calibration

We first calibrate the aggregate-level preference and technology parameters. The values of γ, λ, δ, $\overline{x}$, and Z are chosen to match approximately the unconditional moments of the key aggregate variables. Table 3.1 reports the parameter values used in simulation and Table 1.2 compares the moments of some key aggregate variables in the model with corresponding empirical estimates. For completeness, we report two sets of moments from the model: population moments and sample moments. Population moments are estimated by simulating a $300,000$-month time series; the sample moments are computed based on 200 simulations, each containing 70 years worth of monthly data.[12] In addition

---

[12]The 70-year sample length is comparable to that of CRSP, which is the historical data set used in generating the two (Data) columns in Table 1.2.

to point estimates and standard errors, we also report 95% confidence intervals based on empirical distribution functions from 200 simulations. Population moments are close to their empirical counterparts and almost all the moments of historical series are within the 95% confidence intervals in the (Sample) columns.

Our model is able to capture the historical level of the equity premium, while maintaining plausible values for the first two moments of the risk-free rate. These results are due to the combination of sufficiently high risk aversion ($\gamma = 15$) of the representative household and a small amount of predictability in the consumption process (e.g., Kandel and Stambaugh (1991)).[13] Based on these results, we conclude that our model provides a satisfactory fit of the aggregate data.

To further illustrate the properties of our model, we plot some key economic variables against the state variable x in Figure 1.1. Panel A shows that the optimal investment policy, $\overline{e}^\star$, increases with x. In equilibrium, $\overline{e}^\star$ equals the present value of cash flows from a new project of unit size, V^a/K, which is increasing in productivity parameter x. Similarly, the market value per unit scale of a typical project, V/K, is increasing in x, as shown in Panel B. According to Panel C, the value of assets-in-place as a fraction of the total stock market value decreases slightly with x. Most of the time, assets-in-place account for 75–80% of the stock market value in the model. Finally, Panel D compares the instantaneous stock market betas, β^a and β^o. The beta of growth options is higher than that of assets in place.

1.3.2 Quantitative Results

We now examine some additional quantitative implications of the model for the relationship between aggregate returns and other aggregate variables. Table 1.3 Panel A reports

---

[13]Note that we are not arguing that this is the precise mechanism behind the observed equity premium and other aggregate-level properties of asset prices. The only objective of this analysis is to verify that our cross-sectional results are not undermined by unreasonable aggregate-level properties of the model.

the means, standard deviations, and 1- to 5-year autocorrelations of the dividend yield and book-to-market ratio. We estimate these statistics by repeatedly simulating 70 years of monthly data,.a sample size similar to that used in Pontiff and Schall (1998). The Data rows report the mean and standard deviation of the book-to-market ratio to be 0.668 and 0.23 respectively, the values taken from Pontiff and Schall, Table 1 Panel A. Our model produces similar values of 0.584 and 0.19. The autocorrelations of the book-to-market ratio are decreasing with the horizon, matching the pattern observed in the data. However, the ratio is more persistent in the model compared to the data, as indicated by higher magnitude of autocorrelations. The model also reproduces the decreasing pattern of autocorrelations of the dividend yield data. While the standard deviation of dividend yield is close to the empirical value, the average level exceeds the number reported by Pontiff and Schall (1998). Panel B in Table 1.3 examines the performance of the book-to-market ratio as a predictor of stock market returns. The slope in the regression of monthly value-weighted market returns on one-period lagged book-to-market ratios based on the model is 1.75%. The empirical value of 3.02% is within the 95% confidence interval around the simulation-based estimate. The adjusted R^2s are also comparable. The same analysis at annual frequency produces similar results.

It is also important to note that, in the model, instantaneous stock market returns are perfectly correlated with consumption growth and the stochastic discount factor. As a result, asset returns are characterized by a single-factor intertemporal CAPM. To determine how closely monthly stock returns satisfy the ICAPM with the market portfolio being the only factor, we regress market returns on the contemporaneous realization of the stochastic discount factor, given by $(C_{t+\triangle t}/C_t)^{-\gamma}e^{-\lambda\triangle t}$. As expected, the regression shows that 96% of the variation in market return can be explained by variation in the stochastic discount factor. The unconditional correlation between the stochastic discount factor and the market return is -0.98 and the conditional correlation between the two is,

effectively, -1. Thus, even at the monthly frequency, a single-factor ICAPM is, theoretically, highly accurate. In this respect our environment differs crucially from Berk, Green, and Naik (1999). By construction then, stock returns in their model cannot be described using market returns as a single risk factor, allowing variables other than market βs to play an independent role in predicting stock returns.

1.4 The Cross-Section of Stock Returns

This section establishes our key quantitative results. After outlining our numerical procedure, subsection 1.4.3 documents the ability of the model to replicate the empirical findings about the relation between firm characteristics and stock returns. It also establishes that these findings disappear after one controls for the theoretically correct measure of systematic risk. Subsection 1.4.4 describes the conditional, or cyclical, properties of firm level returns.

1.4.1 Calibration

To examine the cross sectional implications of the model we must choose the parameters of the stochastic process for firm-specific productivity shocks, κ and σ_ϵ. We restrict these values by two considerations. First, we want to be able to generate empirically plausible levels of volatility of individual stock returns, which directly affects statistical inference about the relations between returns and firm characteristics. Second, we also want the cross-sectional correlation between firm characteristics, i.e., the logarithms of firm value and book-to-market ratio, to match the empirically observed values. The value, and particularly the sign of this correlation, are critical in determining the *univariate* relations between firm characteristics and returns implied by the multivariate relation (1.32), due to the well-known omitted variable bias.

We can accomplish these goals by setting value of $\kappa = 0.51$ and $\sigma_\epsilon = 2.10$. These values imply an average annualized volatility of individual stock returns of approximately 25% and a correlation between size and book-to-market variables of about -0.26, the number reported by Fama and French (1992). Panel D of Figure 1.1 shows the behavior of $\widetilde{\beta}^a$ implied by our choice of κ. In particular, $\widetilde{\beta}^a$ is lower than the market beta of assets-in-place and is increasing in the state variable x.

According to equation (1.32), there exists a cross-sectional relation between the market βs of stock returns and firm characteristics. The sign of this relation depends on the aggregate-level variables $\beta_t^o - \widetilde{\beta}_t^a$ and $\beta_t^a - \widetilde{\beta}_t^a$ in (1.32). Under the calibrated parameter values, the long-run average values of $\beta_t^o - \widetilde{\beta}_t^a$ and $\beta_t^a - \widetilde{\beta}_t^a$ are 0.67 and 0.21 respectively.

These numbers suggest then a negative relation between market βs and firm size and a positive one between βs and book-to-market. Since size and book-to-market are negatively correlated in our model, coefficients in univariate regressions of returns on these variables should have the same sign as partial regression coefficients in a joint regression, i.e., returns should be negatively related to size and positively related to book-to-market. To further evaluate the quantitative significance of these effects, we repeatedly simulate a panel data set of stock returns based on our model and apply commonly used empirical procedures on the simulated panel.

We follow the empirical procedures used by Fama and French (1992). First, we present some descriptive statistics of the simulated panel in Tables 1.4 and 1.5, providing an informal summary of the relations between returns, size, and book-to-market. Our main results are presented in Tables 1.7, 1.8, and 1.9, where we detail the cross-sectional relations between stock returns and firm characteristics.

1.4.2 Simulation and Estimation

In our simulations, the artificial panel consists of 360 months of observations for 2,000 firms. This panel size is comparable to that in Fama and French (1992), who used an average of 2,267 firms for 318 months. We also adhere to Fama and French's timing convention in that we match the accounting variables at the end of the calendar year $t-1$ with returns from July of year t to June of year $t+1$. Moreover, we use the value of the firm's equity at the end of calendar year $t-1$ to compute its book-to-market ratios for year $t-1$, and we use its market capitalization for June of year t as a measure of its size.[14] Further details of our simulation procedure are summarized in Appendix 1.5.

Some of our tests use estimates of market βs of stock returns, which are obtained using the empirical procedure of Fama and French (1992).[15] Their procedure consists of two steps. First, *pre-ranking* βs for each firm at each time period are estimated based on previous 60 monthly returns. Second, for each month stocks are sorted into ten portfolios by market value. Within each size portfolio, stocks are sorted again into ten more portfolios by their pre-ranking βs. The *post-ranking* βs of each of these 100 portfolios are then calculated using the full sample. All portfolios are formed using equal weights and all βs are calculated by summing the slopes of a regression of portfolio returns on market returns in the current and prior months. In each month, we then allocate the portfolio βs to each of the stocks within the portfolio. To highlight the fact that these post-ranking βs are estimated, we will refer to them as Fama and French-βs.

Following Fama and French (1992), we form portfolios at the end of June each year and the equal-weighted returns are calculated for the next 12 months. In each of these sorts, we form 12 portfolios. The middle 8 portfolios correspond to the middle 8 deciles

---

[14]In this aspect our simulation procedure differs from that of Berk et al. (1999), since they use a straightforward and intuitive timing convention (one-period-lag values of explanatory variables), which does not however agree with the definitions in Fama and French (1992).

[15]For details of the beta estimation procedure, we refer readers to Fama and French (1992).

of the corresponding characteristics, with 4 extreme portfolios (1A, 1B, 10A, and 10B) splitting the bottom and top deciles in half. We repeat the entire simulation 100 times and average the results of the sorting procedure across the simulations. In tables 1.4, 1.5 and 1.6, Panel A is taken from Fama and French (1992) and Panel B is computed based on the simulated panels.

1.4.3 Size and Book-to-Market Effects

Tables 1.4 and 1.5 report post-ranking average returns for portfolios formed by a one-dimensional sort of stocks on firm size and book-to-market. When portfolios are formed on firm value (Table 1.4), the simulated panel exhibits a negative relation between size and average returns, similar to the one observed empirically.[16] Table 1.5 presents average returns for portfolios formed based on ranked values of book-to-market ratios. Similar to the historical data, our simulated panels on average also show a positive relation between book-to-market ratios and average returns. Thus, one-dimensional sorting procedures indicate cross-sectional relations between Fama and French factors and returns that are similar to those in the historical data.

Table 1.7 shows a summary of our results from the Fama-MacBeth (1973) regressions of stock returns on size, book-to-market, and conditional market βs.[17] For comparison, we also report empirical findings of Fama and French (1992) and simulation results of Berk et al. (1999) in columns 2 and 3 of the same table.

Our first univariate regression shows that the logarithm of firm market value appears to contain useful information about the cross-section of stock returns in our model. The

---

[16]The level of average returns is higher in Panel A than in Panel B. This difference is due to the fact that we are modeling real returns in our model, while Fama and French (1992) report the properties of nominal historical returns.

[17]For each simulation, we compute the slope coefficients as the time series average coefficients over the 360-month cross-sectional regressions, and the t-statistics are these averages divided by the standard deviations across the 360 months, which provide standard Fama-MacBeth (1973) tests for statistical significance of regression coefficients. We then average the results across 100 simulations. The market βs are exact conditional βs computed based on our theoretical model.

relation between returns and the size variable is significantly negative. The average slope coefficient as well as the corresponding t-statistic implied by the model are close to their empirical values reported by Fama and French (1992). Panel A of Figure 1.2 shows the histogram of realized t-statistics across simulations. The empirical value is well within the body of realizations produced by the model. Our second univariate regression confirms the importance of book-to-market ratio in explaining the cross-sectional properties stock returns. While our slope coefficient is smaller than the one obtained by Fama and French (1992), our estimate is also positive on average. Panel B of Figure 1.2 shows that the coefficient of book-to-market is often significant at traditional levels, however, the model is not able to produce the t-statistics as high as that reported by Fama and French (1992).

Next, we regress returns on size and book-to-market jointly. On average our coefficients have the same signs as in Fama and French (1992) and Berk et al. (1999) as returns exhibit negative dependence on size and positive dependence on book-to-market. While our average size slope and the corresponding t-statistic are close to the empirical values, the average slope on book-to-market is smaller than in Fama and French (1992). Panel C of Figure 1.2 illustrates the range of t-statistics in a joint regression of returns on size and book-to-market that could be obtained if the historical data were generated by our model. We present the results in the form of a scatter plot, where each point corresponds to a realization of two t-statistics obtained in a single simulation. The empirically observed t-statistic on the size variable is comparable to typical realizations produced by the model. However, the t-statistic on book-to-market is usually somewhat lower than in Fama and French (1992).

The first three regressions in Table 1.7 conform to the intuition derived from our theoretical relation (1.32) that size and book-to-market are related to systematic risks of stock returns and therefore have explanatory power in the cross-section. However, within our theoretical framework, firm characteristics add no explanatory power to the

conditional market βs of stock returns.[18] To illustrate this point, we regress returns on size while controlling for market β. The fourth row of Table 1.7 shows that the average coefficient on size and the corresponding t-statistic are close to zero.

Fama and French (1992) find that the estimated market βs show no explanatory power when used individually or jointly with Fama and French factors. This could be because in practice returns on the market portfolio are not perfectly correlated with the stochastic discount factor and additional risk factors are necessary to describe expected returns. Such mechanism lies beyond the scope of our *single-factor* model. To reconcile our results with poor empirical performance of Fama and French-βs one must take into account the fact that so far we have been using the exact conditional βs, which are not observable in practice. Instead, βs must be estimated, which leaves room for measurement error. Potential sources of errors are, among others, the fact that the market-proxy used in estimation is not the mean-variance efficient portfolio (Roll (1977)) or the econometric methods employed in estimation do not adequately capture the *conditional* nature of the pricing model (e.g., Ferson, Kandel and Stambaugh (1987), Jaganathan and Wang (1996), Campbell and Cochrane (2000), and Lettau and Ludvigson (2000)). Our artificial economy provides an example of how significance of firm characteristics as predictors of returns can persist due to β mismeasurement.

In our simulations we use the true market portfolio. However, in the model conditional market βs are time-varying, which can potentially lead to measurement problems. To illustrate the impact of β mismeasurement, we apply Fama and French (1992) estimation procedure to our simulated data. First, we form 100 portfolios by sorting on size and then on pre-ranking βs. Table 1.6 provides evidence on the relation between βs and average returns. After stocks have been sorted by size, the second-pass β sort produces

---

[18]Theoretically, market βs are sufficient statistics for instantaneous expected returns in our model. As shown in section 1.3, even at monthly frequency, the market portfolio is almost perfectly correlated with the stochastic discount factor.

little variation in average returns. Table 1.8 shows results of the joint regression of returns on firm value and Fama and French-β. On average, the size variable remains negative and significant, while the average t-statistic on Fama and French-β is close to zero. The scatter plot in Panel D of Figure 1.2 shows that the t-statistic on Fama and French-β is usually less than 1.96, while the coefficient on size would often appear significant. In a univariate regression, the slope coefficient and the t-statistic on Fama and French-β reported in Table 1.8 are relatively low compared to those on the exact conditional β, as reported in Table 1.7.

Table 1.9 presents a measure of estimation noise in Fama and French-β, the average correlation matrix of the true conditional βs, Fama and French-βs, size, and book-to-market. For every simulation, we calculate the correlations between true β, Fama and French-β, book-to-market, and size every month and then report the averages of the correlation coefficients and their corresponding standard deviations across simulations. Table 1.9 shows that size is highly negatively correlated with the exact conditional β. The correlation between Fama and French-β and the true β is lower. Not surprisingly, size serves as a more accurate measure of systematic risk than Fama and French-β and hence outperforms it in a cross-sectional regression. Moreover, imperfect correlation between the true β and Fama and French-β in our model lowers the coefficient and the t-statistic in the univariate regression of returns on Fama and French-βs due to the errors-in-variables bias. This illustrates how mismeasurement of β can have an effect on all of the cross-sectional results, bringing out firm characteristics such as size and book-to-market as predictors of expected returns.

Sensitivity Analysis

Finally, it is interesting to take some measure of the sensitivity of our findings to choices of the key parameters, κ and σ_ϵ, governing the cross-sectional properties of stock returns.

Tables 1.10 and 1.11 report the results of these experiments.

We consider two alternative combinations of parameters. First, we look at the effects of increasing the cross-sectional dispersion of stock returns to 30%, which corresponds to a value for σ_ϵ of 2.82. The results are reported in the columns labeled "High Variance" of Tables 1.10 and 1.11. Next, we study the effects of changing the persistence of the idiosyncratic productivity shocks by raising the value of κ to 0.4, while keeping the cross sectional variance of returns at 25%. The "Low Persistence" columns show the results of these simulations.

Comparison between columns 2 and 3 in Table 1.10 and 1.11 shows that the inference from the benchmark model carries, without any significant change, both to the High Variance and the Low Persistence variants of the model, as both the signs and significance of all the coefficients are preserved. Our main results appear to be quite robust with respect to perturbations of main parameter values.

1.4.4 Business Cycle Properties

The theoretical characterization of stock prices and systematic risk, as given by (1.31) and (1.32), highlights the fact that the properties of the cross-section of stock prices and stock returns depend on the current state of the economy. This dependence is captured by the economy-wide variables V_t^a, $\widetilde{V}_t^a$, and V_t^o and their market βs. Thus, our model also gives rise to a number of predictions about the variation of the cross-section of stock prices and returns over the business cycle. These properties of the cross-section of stock returns may have important implications for optimal dynamic portfolio choice.

Firm Characteristics

To help understand the relation between the cross-section of firm characteristics and the business cycle, we first characterize the cross-sectional dispersion of firm market values.

To this end, let **var** (h) denote the variance of the cross-sectional distribution of a firm-specific variable h. According to our characterization of firm market value (1.31), it follows immediately that

$$\textbf{var}\left(\frac{V_{ft}}{V_t}\right) = \left(\frac{\widetilde{V}_t^a}{V_t}\right)^2 \textbf{var}\left(\int_{\mathcal{I}_{ft}} (\epsilon_{it} - 1)\frac{k_i}{K_t}\,di\right) + \left(\frac{V_t^a}{V_t}\right)^2 \textbf{var}\left(\int_{\mathcal{I}_{ft}}\frac{k_i}{K_t}\,di\right) \quad (1.34)$$

The right-hand side of (1.34) captures the cross-sectional dispersion of relative firm size. This dispersion can be attributed to: (i) the cross-sectional variation of project-specific productivity shocks ϵ_{it} as well as project-specific and firm-specific production scale, and (ii) economy-wide variables V_t^a/V_t and $\widetilde{V}_t^a/V_t$.

The contribution of the first source of heterogeneity, captured by **var** $\left(\int_{\mathcal{I}_{ft}} k_i/K_t di\right)$ and **var** $\left(\int_{\mathcal{I}_{ft}} (\epsilon_{it} - 1)k_i/K_t di\right)$, is clearly path-dependent in theory, since the scale of new projects depends on the current aggregate scale of production K_t. Intuitively however this dependence is fairly low when the average life-time of individual projects is much longer than the average length of a typical business cycle.[19]

It falls then on the aggregate components, characterized by $V^a(x_t)/V(x_t)$ and $\widetilde{V}^a(x_t)/V(x_t)$, to determine the cross-sectional variance in market value. Given the properties of our environment, it is easy to see that this implies that the cross-sectional dispersion of firm size is countercyclical, that is, it expands in recessions and it becomes compressed in expansions. We can see this by looking at Panel D of Figure 1.1. Since the market βs of V_t^a and $\widetilde{V}_t^a$ are less than one, the ratios V_t^a/V_t and $\widetilde{V}_t^a/V_t$ should be negatively related to the state variable x_t. Figure 1.3 confirms this finding.

To quantify this relation, we simulate our artificial economy over a 200-year period and compute the cross-sectional standard deviation of the logarithm of firm values and book-to-market ratios on a monthly basis. Since the state variable x_t is not observable

---

[19]Note that the average project life is about $1/\delta = 25$ years, given our calibration.

empirically, we choose to capture the current state of the economy by the price-to-dividend ratio of the aggregate stock market.[20]

Figure 1.3 presents scatter-plots of the cross-sectional dispersion of firm characteristics against the logarithm of the aggregate price-dividend ratio. In both cases the relation is clearly negative. Note that cross-sectional dispersion is not a simple function of the state variable. This is partially due to the fact that we are using a finite number of firms and projects in our simulation, therefore our theoretical relations hold only approximately. Moreover, as suggested by the above theoretical argument, such relations are inherently history-dependent.

Stock Returns

Next we study how the cross-sectional distribution of actual stock returns depends on the state of the aggregate economy. First, we analyze the degree of dispersion of returns, $RD_t = \sqrt{\mathbf{var}\left(R_{ft}\right)}$, where R_{ft} denotes monthly returns on individual stocks. We construct a scatter-plot of RD_t versus contemporaneous values of the logarithm of the aggregate price-dividend ratio.

According to Figure 1.5, our model predicts a negative contemporaneous relation between return dispersion and the price-dividend ratio. This can be attributed to the countercyclical nature of both aggregate return volatility, as shown in Panel A of Figure 1.4, and of the dispersion in conditional market β, as shown in Panel B.

Since investment in our model is endogenously procyclical, an increase in aggregate productivity shock leads to an increase in the scale of production as well as an increase in stock prices. On the other hand, since investment is irreversible, the scale of production cannot be easily reduced during periods of low aggregate productivity, increasing volatility of stock prices.[21]

---

[20]In the model, the unconditional correlation between x_t and $\log\left(V_t/D_t\right)$ is approximately 99.3%

[21]Qualitatively, the impact of the irreversibility on conditional volatility of stock returns in our model

32

The countercyclical dispersion of conditional βs follows from the characterization of the systematic risk of stock returns (1.32) and the pattern observed in Figure 1.1, Panel D. During business cycle peaks, the dispersion of aggregate βs, i.e., β_t^a, $\widetilde{\beta}_t^a$, and β_t^o, is relatively low, contributing to lower dispersion of firm-level market βs. This effect is then reinforced by the countercyclical behavior of dispersion of firm characteristics.

An interesting empirical finding by Stivers (2000) is the ability of return dispersion to forecast future aggregate return volatility, even after controlling for the lagged values of market returns. We conduct a similar experiment within our model, by simulating 1000 years of monthly stock returns and regressing absolute values of aggregate market returns on lagged values of return dispersion and market returns. As in Stivers (2000), we allow for different slope coefficients depending on the sign of lagged market returns. As shown in Table 1.12, both lagged market returns and return dispersion predict future conditional volatility of returns. Return dispersion retains significant explanatory power even after controlling for market returns in the regression. This is due to the fact that lagged market returns provide only a noisy proxy for the current state of the economy, and return dispersion contains independent information such as the current dispersion of market βs.

Conditional Size and Book-to-Market Effects

The fact that dispersion of returns on individual stocks in our model changes counter-cyclically suggests that the size and book-to-market effects analyzed in subsection 1.4.3 are also conditional in nature.

To capture this cyclical behavior of cross-sectional patterns in returns and its impli-cations for dynamic portfolio allocation, we analyze the conditional performance of alter-native size- and value-based strategies. Specifically, we simulate 1,000 years of monthly

is similar to that in Kogan (2000a, 2000b).

individual stock returns and then form zero-investment portfolios by taking a long posi-
tion in bottom-size-decile stocks and a short position in top-size-decile stocks, as sorted
by size, with monthly rebalancing. We also construct alternative portfolios by doing the
opposite for book-to-market deciles. We then regress portfolio returns on the logarithm
of the aggregate price-dividend ratio.

Our model predicts an average annualized value (book-to-market) premium of 1.45%
and an average annualized size premium of 1.93%. Moreover, both strategies exhibit
significant countercyclical patterns in their expected returns. In particular, we find that
a 10% decline in the price-dividend ratio below its long-run mean implies approximately
a 12% and 9% increase in expected returns on the size and book-to-market strategies,
respectively, measured as a fraction of their long-run average returns.

1.5   Conclusion

This paper analyzes a general equilibrium production economy with heterogeneous firms.
In the model, the cross-section of stock returns is explicitly related to firm characteristics
such as size and book-to-market. Firms differ in the share of their total market value
derived from their assets, as opposed to future growth opportunities, which is captured
by their characteristics. Since these two components of firm value have different market
risk, firm characteristics are closely related to market β.

To the best of our knowledge, our paper is the first to explain the cross-section
of stock returns from a general equilibrium perspective. Our model demonstrates that
size and book-to-market can explain the cross-section of stock returns because they are
correlated with the true conditional β. We also provide an example of how empirically
estimated β can perform poorly relative to firm characteristics due to measurement errors.

Our model also gives rise to a number of additional implications for the cross-section

of returns. In this paper, we focus on the business cycle properties of returns and firm characteristics. Our results appear consistent with the limited existing evidence and provide a natural benchmark for future empirical studies.

Bibliography

[1] Abel, Andrew B., 1983, Optimal Investment under Uncertainty, *American Economic Review*, 73 (1), 228–233.

[2] Berk, Jonathan B., 1995, A Critique of Size Related Anomalies, *Review of Financial Studies*, 8, 275–286.

[3] Berk, Jonathan B, Richard C. Green and Vasant Naik, 1999, Optimal Investment, Growth Options and Security Returns, *Journal of Finance*, 54, 1153-1607.

[4] Bossaerts, Peter, and Richard C. Green, 1989, A General Equilibrium Model of Changing Risk Premia: Theory and Tests, *Review of Financial Studies*, 4, 467–493.

[5] Breeden, Douglas T., 1979, An Intertemporal Asset Pricing Model with Stochastic Consumption and Investment Opportunities, *Journal of Financial Economics*, 7, 265–296.

[6] Campbell, John Y., 1999, Asset Prices, Consumption, and the Business Cycle, in John B. Taylor and Michael Woodford eds., *Handbook of Macroeconomics*, Amsterdam: North-Holland.

[7] Campbell, John Y., 2000, Asset Pricing at the Millennium, *Journal of Finance*, 55, 1515–1567.

[8] Campbell, John Y., and John H. Cochrane, 2000, Explaining the Poor Performance of Consumption-Based Asset Pricing Models, *Journal of Finance*, 55, 2863–2878.

[9] Campbell, John Y., and Andrew W. Lo, and A. Craig MacKinlay, 1997, *The Econometrics of Financial Markets*, Princeton University Press, Princeton, New Jersey.

[10] Cochrane, John H., 1991, Production-Based Asset Pricing and the Link Between Stock Returns and Economic Fluctuations, *Journal of Finance*, XLVI (1), 209–237.

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

[12] Coleman II, W., 1997, "Behavior of Interest Rates in a General Equilibrium Multisector Model with Irreversible Investment," *Macroeconomic Dynamics* 1, 206–227.

[13] Duffie, Darrel, 1996, *Dynamic Asset Pricing Theory*, Princeton University Press, Princeton, New Jersey.

[14] Dybvig, Phillip, and Chifu Huang, 1989, Nonnegative Wealth, Absence of Arbitrage, and Feasible Consumption Plans, *Review of Financial Studies* 4, 377–401.

[15] Evans, David, 1987, The Relationship Between Firm Growth, Size and Age: Estimates for 100 Manufacturing Industries, *Journal of Industrial Economics* 35, 567–581.

[16] Fama, Eugene F., and Kenneth R. French, 1992, The Cross-Section of Expected Stock Returns, *Journal of Finance*, XLVII (2), 427–465.

[17] Fama, Eugene F., and Kenneth R. French, 1993, Common Risk Factors in the Returns on Stocks and Bonds, *Journal of Finance*, 33, 3–56.

[18] Fama, Eugene F., and Kenneth R. French, 1995, Size and Book-to-Market Factors in Earnings and Returns, *Journal of Finance*, 50, 131–155.

[19] Fama, Eugene F., and Kenneth R. French, 1996, Multifactor Explanations of Asset Pricing Anomalies, *Journal of Finance*, 51, 55–84.

[20] Fama, Eugene F., and James MacBeth, 1973, Risk, Return, and Equilibrium Empirical Tests, *Journal of Political Economy*, 71, 607–636.

[21] Feldman, Mark, and Christian Gilles, 1985, An Expository Note on Individual Risk without Aggregate Uncertainty, *Journal of Economic Theory*, 1985, 23–32.

[22] Ferson, Wayne E., Shmuel Kandel, and Robert F. Stambaugh, 1987, Tests of Asset Pricing with Time-Varying Expected Rsik Premiums and Market Betas, *Journal of Finance*, 42, 210–220.

[23] Hall, Bronwyn, H., 1987, The Relationship Between Firm Size and Firm Growth in the US Manufacturing Sector, *Journal of Industrial Economics* 35, 583–606.

[24] Jagannathan, Ravi, and Zhenyu Wang, 1996, The Conditional CAPM and the Cross-Section of Expected Returns, *Journal of Finance*, 51, 3–54.

[25] Jermann, Urban J., 1998, Asset Pricing in Production Economies, *Journal of Monetary Economics*, 41, 257–275.

[26] Judd, Kenneth, 1985, The Law of Large Numbers with a Continuum of IID Random Variables, *Journal of Economic Theory*, 35, 19–25.

[27] Kaldor, Nicholas, 1963, Capital Accumulation and Economic Growth, in Friedrich A. Lutz and Douglas C. Hague, eds., *Proceedings of a Conference Held by the International Economics Association,* London, Macmillan.

[28] Kandel, Shmuel, and Robert Stambaugh, 1991, Asset Returns and Intertemporal Preferences, *Journal of Monetary Economics*, 50, 27–71.

[29] Kogan, Leonid, 2000a, An Equilibrium Model of Irreversible Investment, Working Paper, the Wharton School, University of Pennsylvania.

[30] Kogan, Leonid, 2000b, Asset Prices and Irreversible Real Investment, Working Paper, the Wharton School, University of Pennsylvania.

[31] Kothari, S. P., Jay Shanken, and Richard G. Sloan, 1995, Another Look at the Cross-Section of Expected Stock Returns, *Journal of Finance*, 50, 185–224.

[32] Kydland, Finn E., and Edward C. Prescott, 1982, Time to Build and Aggregate Fluctuations, *Econometrica*, 50, 1345–1370.

[33] Lakonishok, Josef, Andrei Shleifer, and Robert W. Vishny, 1994, Contrarian Investment, Extrapolation, and Risk, *Journal of Finance*, XLIX (5), 1541–1578.

[34] Lettau, Martin, and Sydney Ludvigson, 2001, Resurrecting the (C)CAPM: A Cross-Sectional Test When Risk Premia Are Time-Varying, forthcoming, *Journal of Political Economy*.

[35] Liew, Jimmy, and Maria Vassalou, 2000, Can Book-to-Market, Size, and Momentum Be Risk Factors That Predict Economic Growth? *Journal of Financial Economics*, 57, 221–245.

[36] Lo, Andrew W., and A. Craig MacKinlay, 1988, Stock Market Prices Do Not Follow Random Walks: Evidence From a Simple Specification Test, *Review of Financial Studies*, 1, 41–66.

[37] Long, John B. Jr., and Charles I. Plosser, 1983, Real Business Cycles, *Journal of Political Economy*, 91 (1), 39–69.

[38] MacKinlay, A. Craig, 1995, Multifactor Models Do Not Explain Deviations From the CAPM, *Journal of Financial Economics*, 38, 3–28.

[39] Merton, Robert C., 1973, An Intertemporal Capital Asset Pricing Model, *Econometrica*, 41, 867–887.

[40] Pontiff, Jeffrey and Lawrence Schall, 1998, Book to Market as a Predictor of Market Returns, *Journal of Financial Economics*, 49, 141–160.

[41] Prescott, Edward, 1986, Theory Ahead of Business Cycle Measurement, *Carnegie-Rochester Conference on Public Policy*, 24, 11–44.

[42] Roll, Richard, 1977, A Critique of Asset Pricing Theory's Test: Part I, *Journal of Financial Economics* 4, 129–176.

[43] Ross, Stephen, 1976, The Arbitrage Theory of Capital Asset Pricing, *Journal of Economic Theory*, 13, 341–360.

[44] Rouwenhorst, G., 1995, "Asset Pricing Implications of Equilibrium Business Cycle Models," in Frontiers of Business Cycle Research, T. Cooley, Ed., Princeton University Press, Princeton, NJ.

[45] Stivers, Christopher T., 2000, Firm-level Return Dispersion and the Future Volatility of Aggregate Stock Market Returns, Working Paper, University of Georgia.

Proofs and Technical Results for Chapter One

Proof of Proposition 2

The equilibrium conditions imply that the optimal firm investment policy $\overline{e}^\star(x)$ satisfies the condition

$$V_{it}^a = E_t \left[\int_0^\infty e^{-\lambda s} \left(\frac{C_t^\star}{C_{t+s}^\star} \right)^\gamma \left(e^{-\delta s} k_i X_{t+s} \right) ds \right] = \overline{e}^\star(x) k_i \qquad (A1)$$

where we impose that optimal consumption decisions are used in determining the stochastic discount factor in equilibrium. In words, optimality of firms' investment decisions requires that the most expensive project undertaken has a present value of cash flows equal to its cost.

Using the fact that k_i is independent of t and equation (1.8), we obtain that:

$$\overline{e}^\star(x) k_i = (C_t^\star)^\gamma k_i E_t \left[\int_0^\infty e^{-(\lambda+\delta)s} \frac{X_{t+s}}{(C_{t+s}^\star)^\gamma} ds \right] =$$

$$= (C_t^\star)^\gamma k_i E_t \left[\int_0^\infty e^{-(\lambda+\delta)s} \frac{X_{t+s}}{(c_{t+s}^\star)^\gamma K_{t+s}^\gamma} ds \right] =$$

$$= (C_t^\star)^\gamma k_i E_t \left[\int_0^\infty e^{-(\lambda+\delta)s} \frac{X_{t+s}}{(c_{t+s}^\star)^\gamma K_t^\gamma \exp \left(\int_0^s -\gamma\delta + \gamma Z \overline{e}_\tau^\star d\tau \right)} ds \right] =$$

$$= (c_t^\star)^\gamma k_i p(x_t)$$

or, as in equation (1.20)

$$\overline{e}^\star(x) = (c_t^\star)^\gamma p(x_t)$$

where the Feynman-Kac theorem implies then that $p(x)$ satisfies the differential equation:[22]

$$[\lambda + (1-\gamma)\delta + \gamma Z \overline{e}^\star(x)] p(x) - \mathcal{A}[p(x)] - \frac{\exp(x)}{[c^\star(x)]^\gamma} = 0$$

and $\mathcal{A}[p(x)]$ is the infinitesimal generator of the diffusion process x_t:

$$\mathcal{A}[p(x)] \equiv -\theta_x (x - \overline{x}) p(x) + \frac{1}{2} \sigma_x^2 p''(x)$$

In addition, optimal consumption and investment policies are also related by the resource constraint (4). Using equations (1) and (1.9) this can be easily transformed into equation (1.21)

$$c^\star(x) = \frac{Y_t}{K_t} - \frac{I_t}{K_t} = \exp(x) - \frac{1}{2} Z [\overline{e}^\star(x)]^2$$

thus completing the proof of the Proposition.

Computation of Equilibrium

We solve for the equilibrium iteratively. First, we use equation (1.21) to eliminate $c(x)$ in (1.22). We then approximate the resulting differential equation for $p(x)$ with a system of linear equations

---

[22]See, for example, Duffie (1996) Appendix E.

upon discretizing the state space of x:

$$[\lambda + (1 - \gamma)\delta + \gamma Z\overline{e}_i]\, p_i = \widehat{\mathcal{A}}(p)_i + \frac{\exp(x_i)}{\left[\exp(x_i) - \frac{1}{2}Z\left(\overline{e}_i\right)^2\right]^\gamma}$$

where $\widehat{\mathcal{A}}(p)$ is the finite-difference approximation to the infinitesimal generator $\mathcal{A}(p)$. We then solve this system together with (1.20). We do this by using the following iterative procedure:

$$p_i^{(n+1)} = p_i^{(n)} + \triangle t^{(n)} \left[\frac{\exp(x_i)}{\left[\exp(x_i) - \frac{1}{2}Z\left(\overline{e}_i^{(n)}\right)^2\right]^\gamma} + \widehat{\mathcal{A}}(p)_i^{(n)} - \left[\lambda + (1-\gamma)\delta + \gamma Z\overline{e}_i^{(n)}\right] p_i^{(n)} \right]$$

$$\overline{e}_i^{(n+1)} = \overline{e}_i^{(n)} - \triangle t^{(n)} \left[\frac{\overline{e}_i^{(n)} - p_i^{(n)} \left[\exp(x_i) - \frac{1}{2}Z\left(\overline{e}_i^{(n)}\right)^2\right]^\gamma}{1 + \gamma Z\overline{e}_i^{(n)} p_i^{(n)} \left[\exp(x_i) - \frac{1}{2}Z\left(\overline{e}_i^{(n)}\right)^2\right]^{\gamma-1}} \right]$$

where the step-size $\triangle t^{(n)}$ is adjusted to ensure convergence.

Proof of Proposition 3

Let $m_t = (C_t^\star)^{-\gamma}$. Then $M_{t,t+s} = e^{-\lambda s} m_{t+s}/m_t$ and by Ito's Lemma,

$$M_{t,t+dt} - 1 = \left.\frac{\partial M}{\partial s}\right|_{s=0} dt + \frac{\partial m_t}{m_t \partial C_t^\star} dC_t^\star + \frac{1}{2} \frac{\partial^2 m_t}{m_t \partial \left(C_t^\star\right)^2} [dC_t^\star]^2$$

$$= -\lambda m_t\, dt - \frac{\gamma}{C_t^\star} m_t\, dC_t^\star + \frac{1}{2} \frac{\gamma(\gamma+1)}{(C_t^\star)^2} m_t\, [dC_t^\star]^2$$

Thus,

$$E[M_{t,t+dt} - 1] = -\lambda\, dt - \frac{\gamma}{C_t^\star}\, E[dC_t^\star] + \frac{1}{2} \frac{\gamma(\gamma+1)}{(C_t^\star)^2} [dC_t^\star]^2$$

Next, since $C_t^\star = K_t\, c^\star(x_t)$, another application of Ito's Lemma yields

$$E[dC_t^\star] = c^\star(x_t)\, dK_t + K_t\, E[d\, c^\star(x_t)] = c^\star(x_t)[Z\overline{e}^\star(x_t) - \delta]K_t\, dt + K_t\mathcal{A}[c^\star(x_t)]\, dt$$

$$[dC_t^\star]^2 = K_t^2 \left[c^\star(x_t)'\right]^2 \sigma_x^2$$

where $\mathcal{A}[c^\star(x)] \equiv \mu_x c^\star(x)' + \frac{1}{2}\sigma_x^2 c^\star(x)''$. As a result,

$$r_t = \lambda + \gamma\left[Z\overline{e}^\star(x_t) - \delta\right] + \gamma \frac{\mathcal{A}[c^\star(x)]}{c^\star(x_t)} - \frac{1}{2}\gamma(\gamma+1)\sigma_x^2 \left(\frac{c^\star(x_t)'}{c^\star(x_t)}\right)^2$$

Second, the value of the aggregate stock market, V_t, can be computed as

$$
V_t = E_t \left[\int_0^\infty e^{-\lambda s} \left(\frac{C_t^\star}{C_{t+s}^\star} \right)^\gamma C_{t+s}^\star \, ds \right] = (C_t^\star)^\gamma E_t \left[\int_0^\infty e^{-\lambda s} (c_{t+s}^\star)^{1-\gamma} K_{t+s}^{1-\gamma} \, ds \right]
$$

$$
= (C_t^\star)^\gamma E_t \left[\int_0^\infty e^{-\lambda s} (c_{t+s}^\star)^{1-\gamma} K_t^{1-\gamma} \exp \left(\int_0^s -(1-\gamma)\delta + (1-\gamma) Z \overline{e}_\tau^\star \, d\tau \right) ds \right]
$$

$$
= (c_t^\star)^\gamma \, \psi(x_t) \, K_t
$$

where, $\psi(x_t)$ is defined by

$$
\psi(x_t) \equiv E_t \left[\int_0^\infty e^{-\lambda s} (c_{t+s}^\star)^{1-\gamma} \exp \left(\int_0^s -(1-\gamma)\delta + (1-\gamma) Z \overline{e}_\tau^\star \, d\tau \right) ds \right]
$$

which, by Feynman-Kac theorem, satisfies the following differential equation:

$$
\lambda \psi(x) = [c^\star(x)]^{1-\gamma} + (1-\gamma) [Z\overline{e}^\star(x) - \delta] \psi(x) - \theta_x(x - \overline{x}) \psi'(x) + \frac{1}{2} \sigma_x^2 \psi''(x)
$$

Proof of Proposition 4

The present value of output from a specific project i, denoted V_{it}^a, is given by

$$
V_{it}^a = E_t \left[\int_0^\infty e^{-\lambda s} \left(\frac{C_t^\star}{C_{t+s}^\star} \right)^\gamma \left(e^{-\delta s} k_i X_{it+s} \right) ds \right]
$$

$$
= k_i (C_t^\star)^\gamma \int_0^\infty e^{-(\lambda+\delta)s} E_t \left[\frac{X_{t+s}}{(C_{t+s}^\star)^\gamma} \right] E_t \left[\epsilon_{it+s} \right] ds
$$

where the last equality follows from mutual independence of X_t and ϵ_{it}. The square-root process (1.2) has the property

$$
E_t \left[\epsilon_{it+s} \right] = \epsilon_{it} e^{-\kappa s} + (1 - e^{-\kappa s})
$$

which implies that

$$
V_{it}^a = k_i (C_t^\star)^\gamma \int_0^\infty e^{-(\lambda+\delta)s} E_t \left[\frac{X_{t+s}}{(C_{t+s}^\star)^\gamma} \right] \left[\epsilon_{it} e^{-\kappa s} + (1 - e^{-\kappa s}) \right] ds
$$

$$
= \frac{k_i}{K_t} K_t (C_t^\star)^\gamma \left[\int_0^\infty e^{-(\lambda+\delta+\kappa)s} E_t \left[\frac{X_{t+s}}{(C_{t+s}^\star)^\gamma} \right] (\epsilon_{it} - 1) \, ds + \int_0^\infty e^{-(\lambda+\delta)s} E_t \left[\frac{X_{t+s}}{(C_{t+s}^\star)^\gamma} \right] ds \right]
$$

$$
= \frac{k_i}{K_t} \left[\widetilde{V}_t^a (\epsilon_{it} - 1) + V_t^a \right]
$$

where $\widetilde{V}_t^a$ is defined as

$$
\widetilde{V}_t^a \equiv K_t E_t \left[\int_0^\infty e^{-(\lambda+\delta+\kappa)s} \left(\frac{C_t^\star}{C_{t+s}^\star} \right)^\gamma \exp(x_{t+s}) \, ds \right]
$$

41

Computation

Discretization and Simulation

We use a finite number of firms in the numerical implementation. While the number of firms is fixed, the total number of projects in the economy is time-varying and stationary. We let the scale of new projects be proportional to the aggregate production scale in the economy, which ensures stationarity of the cross-sectional distribution of the number of projects per firm. Thus, $k_{it}^t = K_t/\varphi$ where the constant φ controls the long-run average number of projects in the economy. On average, projects expire at the total rate $\delta N^\star$. The arrival rate of new projects is $Z\bar{e}_t\varphi$. Therefore, $ZE[\bar{e}_t]\varphi = \delta N^\star$, where $N^\star$ is the long-run average number of projects in the economy.

In the simulation, time increment is discrete. The unit cost of a new project are spaced out evenly over the interval $[0, \bar{e}_t]$. The investment of individual firm at time t is computed as the total amount the firm spends on its new projects at time t. The dividend paid out by a given firm during period t is defined as the difference between the cash flows generated by the firm's existing projects and its investment. Finally, the individual firm's book value is measured as the cumulative investment cost of the firm's projects that remain active at time t.

In our simulation, we first generate 200 years worth of monthly data, to allow the economy to reach steady state. After that, we repeatedly simulate a 420-month panel data set consisting of the cross-sectional variables (360 months of data constitute the main panel and 60 extra months are used for pre-ranking β estimation).

Quality of the Aggregation

We appeal to the law of large number in our theoretical analysis of the economy. Discretization of the economy introduces approximation error, the magnitude of which we evaluate by comparing the aggregate series to their exact analytical counterparts. We simulate the corresponding quantities for 10,080 firms over 420 months and record the aggregation results, the corresponding theoretical values, and the difference between the two. In all cases, the difference between these variables and their analytical counterparts is very close to zero.[23] We thus conclude that the quality of aggregation in our simulation is sufficiently high.

---

[23]Complete results are available upon request.

Table 1.1 : Parameter Values Used in Simulation

The table lists the values of all model parameters used in simulation: the risk aversion coefficient (γ), the time preference parameter (λ), the rate of project expiration (δ), the long run mean of the aggregate productivity variable ($\overline{X}$), the quality of investment opportunities (Z), the volatility (σ_x) and the rate of mean-reversion (θ_x) of the productivity variable, the rate of mean-reversion (κ) and the volatility (σ_ϵ) of the idiosyncratic productivity component.

Parameters	γ	λ	δ	$\bar{x}$	Z	σ_x	θ_x	κ	σ_ϵ
Values	15	0.01	0.04	$\log(0.01)$	0.50	0.08	0.275	0.51	2.10

Table 1.2 : Moments of Key Aggregate Variables

This table reports unconditional means and standard deviations of consumption growth ($C_{t+1}/C_t - 1$), real interest rate (r_t), equity premium ($\log R_t - \log r_t$), and the mean of the Sharpe ratio ($E(\log R_t - \log r_t)/\sigma(\log R_t - \log r_t)$). The numbers reported in columns denoted (Data) are from Campbell, Lo, and MacKinlay (1997). The numbers reported in columns denoted (Population) are population moments. These statistics are computed based on 300,000 months of simulated data. The two columns denoted (Sample) report the finite-sample properties of the corresponding statistics. We simulate 70-year long monthly data sets, which is comparable to the sample length typically used in empirical research. Simulation is repeated 200 times and the relevant statistics are computed for every simulation. Then we report the averages across the 200 replications. The numbers in parenthesis are standard deviations across these 200 simulations and the two numbers in brackets are 2.5% and 97.5% percentiles of the resulting empirical distribution, respectively. All numbers except those in the last three rows are in percentages.

	Data Mean	Data Std	Population Mean	Population Std	Sample Mean	Sample Std
$C_{t+1}/C_t - 1$	1.72	3.28	0.85	3.22	0.84	3.06
					(0.28)	(0.26)
					[0.22 1.33]	[2.56 3.50]
r_t	1.80	3.00	1.30	4.33	1.34	3.98
					(1.30)	(0.85)
					[-0.63 4.23]	[2.55 5.73]
$\log R_t - \log r_t$	6.00	18.0	6.00	14.34	5.89	15.28
					(1.32)	(1.73)
					[2.97 8.13]	[11.80 18.58]
Sharpe Ratio	0.33		0.42		0.39	
					(0.11)	
					[0.17 0.62]	

Table 1.3 : Book-To-Market As a Predictor of Market Returns

This table examines our model's ability to match the empirical regularities documented by Pontiff and Schall (1998). Panel A reports means, standard deviations, and autocorrelations of dividend yield (DIV) and book-to-market ratio (B/M), both from historical data and from simulation output. The numbers in columns denoted (Data) are from last two rows in Table 1 Panel A of PS. Panel B reports the properties of the regression of value-weighted market returns, both at monthly and annual frequency, on one-period lagged book-to-market. The columns denoted (Data) are from Table 2 of PS. In both Panels, the columns denoted (Model) report the statistics from 200 simulations, each of which has the same length as that of the data set used in PS. The numbers in parenthesis are standard deviations across 200 simulations and the two numbers in brackets are 2.5th and 97.5th percentiles, respectively. All numbers, except autocorrelations and adjusted R^2s, are in percentages.

		Panel A: Means, Standard Deviations, and Autocorrelations						
	Source	mean	std	1 year	2 yrs	3 yrs	4 yrs	5 yrs
DIV	Data	4.267	1.37	0.60	0.36	0.26	0.23	0.25
	Model	6.407	0.97	0.69	0.46	0.31	0.19	0.11
		(0.321)	(0.22)	(0.08)	(0.14)	(0.17)	(0.18)	(0.18)
		[5.789 7.084]	[0.61 1.45]	[0.51 0.82]	[0.17 0.70]	[-0.05 0.61]	[-0.16 0.51]	[-0.22 0.45]
B/M	Data	0.668	0.23	0.68	0.43	0.23	0.08	0.00
	Model	0.584	0.19	0.88	0.80	0.73	0.68	0.64
		(0.052)	(0.04)	(0.03)	(0.07)	(0.09)	(0.12)	(0.13)
		[0.495 0.707]	[0.12 0.28]	[0.81 0.93]	[0.63 0.89]	[0.48 0.86]	[0.38 0.84]	[0.31 0.83]

	Panel B: Regressions on Book-To-Market			
	Data slope	Data adj.-R^2	Model slope	Model adj.-R^2
monthly	3.02	0.01	1.75	0.00
			(0.79)	(0.00)
			[0.68 3.65]	[0.00 0.01]
annual	42.18	0.16	19.88	0.04
			(10.46)	(0.04)
			[6.57 46.09]	[0.00 0.14]

Table 1.4 : Properties of Portfolios Formed on Size

At the end of June of each year t, 12 portfolios are formed on the basis of ranked values of size. Portfolios 2-9 cover corresponding deciles of the ranking variables. The bottom and top 2 portfolios (1A, 1B, 10A, and 10B) split the bottom and top deciles in half. The breakpoints for the size portfolios are based on ranked values of size. Panel A is from Fama and French (1992) Table II, Panel A. Panel B is constructed from the simulated panel. The average returns are the time-series averages of the monthly equal-weighted portfolio returns, in percent. $\log(V_f)$ and $\log\left(\frac{B_f}{V_f}\right)$ are the time-series averages of the monthly average values of these variables in each portfolio. β is the time-series average of the monthly portfolio post-ranking βs.

	1A	1B	2	3	4	5	6	7	8	9	10A	10B
					Panel A: Historical Data							
Return	1.64	1.16	1.29	1.24	1.25	1.29	1.17	1.07	1.10	0.95	0.88	0.90
β	1.44	1.44	1.39	1.34	1.33	1.24	1.22	1.16	1.08	1.02	0.95	0.90
$\log(V_f)$	1.98	3.18	3.63	4.10	4.50	4.89	5.30	5.73	6.24	6.82	7.39	8.44
$\log\left(B_f/V_f\right)$	-0.01	-0.21	-0.23	-0.26	-0.32	-0.36	-0.36	-0.44	-0.40	-0.42	-0.51	-0.65
					Panel B: Simulated Panel							
Return	0.74	0.72	0.72	0.71	0.71	0.70	0.69	0.68	0.67	0.65	0.62	0.57
β	1.05	1.05	1.03	1.02	1.02	1.01	1.00	0.99	0.97	0.95	0.89	0.89
$\log(V_f)$	4.23	4.40	4.48	4.53	4.56	4.60	4.64	4.68	4.73	4.82	4.95	5.21
$\log\left(B_f/V_f\right)$	-0.92	-0.86	-0.84	-0.84	-0.84	-0.85	-0.86	-0.89	-0.96	-1.08	-1.24	-1.50

Table 1.5 : Properties of Portfolios Formed on Book-to-Market

At the end of June of each year t, 12 portfolios are formed on the basis of ranked values of book-to-market, measured by $\log\left(\frac{B_f}{V_f}\right)$. The pre-ranking β's use 5 years of monthly returns ending in June of t. Portfolios 2-9 cover deciles of the ranking variables. The bottom and top 2 portfolios (1A, 1B, 10A, and 10B) split the bottom and top deciles in half. The breakpoints for the book-to-market portfolios are based on ranked values of book-to-market equity. Panel A is from Fama and French (1992) Table IV, Panel A. Panel B is from the simulated panel. The average returns are the time-series averages of the monthly equal-weighted portfolio returns, in percent. $\log(V_f)$ and $\log\left(\frac{B_f}{V_f}\right)$ are the time-series averages of the monthly average values of these variables in each portfolio. β is the time-series average of the monthly portfolio post-ranking βs.

	1A	1B	2	3	4	5	6	7	8	9	10A	10B
				Panel A: Historical Data								
Return	0.30	0.67	0.87	0.97	1.04	1.17	1.30	1.44	1.50	1.59	1.92	1.83
β	1.36	1.34	1.32	1.30	1.28	1.27	1.27	1.27	1.27	1.29	1.33	1.35
$\log(V_f)$	4.53	4.67	4.69	4.56	4.47	4.38	4.23	4.06	3.85	3.51	3.06	2.65
$\log\left(B_f/V_f\right)$	-2.22	-1.51	-1.09	-0.75	-0.51	-0.32	-0.14	0.03	0.21	0.42	0.66	1.02

	1A	1B	2	3	4	5	6	7	8	9	10A	10B
				Panel B: Simulated Panel								
Return	0.61	0.67	0.69	0.71	0.71	0.71	0.71	0.72	0.72	0.71	0.72	0.72
β	0.95	0.98	1.01	1.02	1.02	1.03	1.03	1.03	1.02	1.02	1.02	1.02
$\log(V_f)$	4.92	4.66	4.53	4.47	4.45	4.45	4.46	4.47	4.48	4.49	4.51	4.52
$\log\left(B_f/V_f\right)$	-1.54	-1.28	-1.15	-1.04	-0.97	-0.92	-0.87	-0.82	-0.77	-0.72	-0.66	-0.58

Table 1.6 : Average Returns For Portfolios Formed on Size (Down) and then β (Across)

Panel A is identical to Fama and French(1992) Table I Panel A, in which the authors report average returns for 100 size-β portfolios using all NYSE, AMEX, and NASDAQ stocks from July 1963 to December 1990 that meet certain CRSP-COMPUSTAT data requirements. Panel B is produced using our simulated panel data set. The portfolio-sorting procedure is identical to that used in Fama and French(1992). In particular, portfolios are formed yearly. The breakpoints for the size deciles are determined in June of year t using all the stocks in the panel. All the stocks are then allocated to the 10 size portfolios using the breakpoints. Each size decile is further subdivided into 10 β portfolios using pre-ranking βs of individual stocks, estimated with 5 years of monthly returns ending in June of year t. The equal-weighted monthly returns on the resulting 100 portfolios are then calculated for July of year t to June of year $t+1$. The pre-ranking βs are the sum of the slopes from a regression of monthly returns on the current and prior month's market returns. The average return is the time-series average of the monthly equal-weighted portfolio returns, in percent. The (ALL) column shows statistics for equal-weighted size-decile (ME) portfolios and the (ALL) row shows statistics for equal-weighted portfolios of the stocks in each β group.

	All	Low-β	β-2	β-3	β-4	β-5	β-6	β-7	β-8	β-9	High-β
Panel A: Average Monthly Returns (in Percent) from Fama and French(1992)											
All	1.25	1.34	1.29	1.36	1.31	1.33	1.28	1.24	1.21	1.25	1.14
Small-ME	1.52	1.71	1.57	1.79	1.61	1.50	1.50	1.37	1.63	1.50	1.42
ME-2	1.29	1.25	1.42	1.36	1.39	1.65	1.61	1.37	1.31	1.34	1.11
ME-3	1.24	1.12	1.31	1.17	1.70	1.29	1.10	1.31	1.36	1.26	0.76
ME-4	1.25	1.27	1.13	1.54	1.06	1.34	1.06	1.41	1.17	1.35	0.98
ME-5	1.29	1.34	1.42	1.39	1.48	1.42	1.18	1.13	1.27	1.18	1.08
ME-6	1.17	1.08	1.53	1.27	1.15	1.20	1.21	1.18	1.04	1.07	1.02
ME-7	1.07	0.95	1.21	1.26	1.09	1.18	1.11	1.24	0.62	1.32	0.76
ME-8	1.10	1.09	1.05	1.37	1.20	1.27	0.98	1.18	1.02	1.01	0.94
ME-9	0.95	0.98	0.88	1.02	1.14	1.07	1.23	0.94	0.82	0.88	0.59
Large-ME	0.89	1.01	0.93	1.10	0.94	0.94	0.89	1.03	0.71	0.74	0.56
Panel B: Average Monthly Returns (in Percent) from Simulated Panel											
All	0.69	0.68	0.69	0.69	0.69	0.69	0.69	0.69	0.68	0.68	0.69
Small-ME	0.73	0.73	0.73	0.73	0.73	0.73	0.73	0.73	0.74	0.73	0.74
ME-2	0.72	0.72	0.73	0.72	0.72	0.70	0.73	0.72	0.72	0.72	0.71
ME-3	0.71	0.70	0.72	0.71	0.72	0.72	0.70	0.72	0.71	0.70	0.71
ME-4	0.71	0.70	0.71	0.70	0.71	0.69	0.71	0.72	0.71	0.72	0.70
ME-5	0.70	0.71	0.69	0.71	0.70	0.71	0.71	0.70	0.70	0.71	0.71
ME-6	0.70	0.69	0.68	0.70	0.71	0.70	0.70	0.71	0.71	0.68	0.67
ME-7	0.69	0.68	0.68	0.69	0.68	0.69	0.70	0.67	0.67	0.70	0.68
ME-8	0.67	0.64	0.68	0.68	0.66	0.69	0.70	0.68	0.66	0.66	0.68
ME-9	0.65	0.65	0.67	0.65	0.64	0.64	0.68	0.64	0.62	0.62	0.67
Large-ME	0.59	0.56	0.59	0.59	0.61	0.61	0.58	0.62	0.58	0.60	0.59

Table 1.7 : Exact Regressions

This table lists summary statistics for the coefficients and the t-statistics of Fama-MacBeth regressions using exact conditional β on the simulated panel sets. The dependent variable is the realized stock return and independent variables are market β, the logarithm of the market value $(\log(V_t))$, and the logarithm of the book-to-market ratio $(\log(B_t/V_t))$. The column denoted (FF) gives the empirical results obtained by Fama and French (1992), Table III, using the historical returns of 2,267 firms over 318 months. The column denoted (BGN) gives the results obtained by Berk et al. (1999). The column denoted (Model) reports the results from our model. The coefficients in the columns are in percentage terms. The numbers in parenthesis are their corresponding t-statistics. Both coefficients and t-statistics are averaged across 100 simulations.

	FF	BGN	Model
$\log(V_t)$	-0.15	-0.035	-0.139
	(-2.58)	(-0.956)	(-2.588)
$\log[B_t/V_t]$	0.50	—	0.079
	(5.71)	—	(1.845)
$\log(V_t)$	-0.11	-0.093	-0.127
	(-1.99)	(-2.237)	(-2.476)
$\log[B_t/V_t]$	0.35	0.393	0.043
	(4.44)	(2.641)	(1.119)
β	-0.37	0.642	1.076
	(-1.21)	(2.273)	(2.602)
$\log(V_t)$	-0.17	0.053	0.038
	(-3.41)	(1.001)	(0.601)
β	—	—	0.916
	—	—	(2.992)
$\log[B_t/V_t]$	—	—	0.010
	—	—	(0.257)
β	0.15	0.377	0.932
	(0.46)	(1.542)	(3.052)

Table 1.8 : Fama-French Regressions

This table lists summary statistics for the coefficients and the t-statistics of Fama-MacBeth regressions using exact conditional β on the simulated panel sets. The dependent variable is the realized stock return and independent variables are market β, the logarithm of the market value ($\log(V_t)$), and the logarithm of the book-to-market ratio ($\log(B_t/V_t)$). The column denoted (FF) gives the empirical results obtained by Fama and French (1992), Table III, using the historical returns of 2,267 firms over 318 months. The column denoted (BGN) gives the results obtained by Berk et al. (1999). The column denoted (Model) reports the results from our model. The coefficients in the columns are in percentage terms. The numbers in parenthesis are their corresponding t-statistics. Both coefficients and t-statistics are averaged across 100 simulations.

	FF	BGN	Model
$\log(V_t)$	-0.15	-0.035	-0.139
	(-2.58)	(-0.956)	(-2.588)
$\log[B_t/V_t]$	0.50	−	0.079
	(5.71)	−	(1.845)
$\log(V_t)$	-0.11	-0.093	-0.127
	(-1.99)	(-2.237)	(-2.476)
$\log[B_t/V_t]$	0.35	0.393	0.043
	(4.44)	(2.641)	(1.119)
β	-0.37	0.642	0.100
	(-1.21)	(2.273)	(0.318)
$\log(V_t)$	-0.17	0.053	-0.126
	(-3.41)	(1.001)	(-2.091)
β	0.15	0.377	0.572
	(0.46)	(1.542)	(2.081)

Table 1.9 : Cross-Sectional Correlations

We calculate the cross-sectional correlations of exact conditional β, FF-β, book-to-market, and size for every simulated panel every month and then report the average correlations across 100 simulations. The numbers in parentheses are cross-simulation standard deviations.

	True β	FF-β	$\log[B_t/V_t]$	$\log(V_t)$
True β	1	0.597	0.322	-0.764
		(0.031)	(0.023)	(0.012)
FF-β		1	0.269	-0.761
			(0.035)	(0.041)
$\log[B_t/V_t]$			1	-0.268
				(0.019)
$\log(V_t)$				1

Table 1.10 : Exact Regressions — Sensitivity Analysis

This table lists summary statistics for the coefficients and the t-statistics of Fama-MacBeth regressions using exact conditional β. The dependent variable is the realized stock return. Independent variables are market β, size measured as the log market value ($\log(V_t)$), and the log of book-to-market ratio ($\log(B_t/V_t)$). The column denoted (FF) gives the empirical results obtained by Fama and French (1992), Table III, using the actual returns of 2,267 firms over 318 months. The column denoted (Benchmark) reports the regression results for the benchmark model, the same as the last column in Table 1.8. The column denoted (High Variance) reports the results from the model with perfect correlated shocks within each firm but with the calibrated parameter values $\kappa = 0.51$ and $\sigma_\epsilon = 2.82$ such that $\sigma_f = 30\%$, which is higher than the benchmark case when $\sigma_f = 25\%$. The column denoted (Low Persistence) reports the results from the model with perfect correlated shocks within each firm but with the calibrated parameter values $\kappa = 0.40$ and that σ_f remains at the benchmark level of 25%. However, the persistence level is now lower. The regression coefficients are in percentage terms. The numbers in parenthesis are t-statistics.

	FF	Benchmark	High Variance	Low Persistence
$\log(V_t)$	-0.15	-0.138	-0.134	-0.133
	(-2.58)	(-2.583)	(-2.246)	(-2.669)
$\log[B_t/V_t]$	0.50	0.079	0.084	0.085
	(5.71)	(1.866)	(1.667)	(2.205)
$\log(V_t)$	-0.11	-0.126	-0.120	-0.120
	(-1.99)	(-2.474)	(-2.115)	(-2.502)
$\log[B_t/V_t]$	0.35	0.043	0.040	0.043
	(4.44)	(1.157)	(0.887)	(1.286)
β	-0.37	1.026	1.000	0.938
	(-1.21)	(2.477)	(2.032)	(2.561)
$\log(V_t)$	-0.17	0.029	0.027	0.024
	(-3.41)	(0.449)	(0.344)	(0.402)
β	–	0.892	0.891	0.831
	–	(2.933)	(2.604)	(2.992)
$\log[B_t/V_t]$	–	0.013	0.010	0.010
	–	(0.355)	(0.204)	(0.313)
β	0.15	0.913	0.914	0.846
	(0.46)	(3.007)	(2.682)	(3.086)

Table 1.11 : Fama-French Regressions — Sensitivity Analysis

This table lists summary statistics for the coefficients and the t-statistics of Fama-MacBeth regressions using Estimated Portfolio β. The dependent variable is the realized stock return. Independent variables are market beta β, size measured as the log market value ($\log(V_t)$), and the log of book-to-market ratio ($\log(B_t/V_t)$). The column denoted (FF) gives the empirical results obtained by Fama and French (1992), Table III, using the actual returns of 2,267 firms over 318 months. The column denoted (Benchmark) reports the regression results for the benchmark model, the same as the last column in Table 1.8. The column denoted (High Variance) reports the results from the model with perfect correlated shocks within each firm but with the calibrated parameter values $\kappa = 0.51$ and $\sigma_\epsilon = 2.82$ such that $\sigma_f = 30\%$, which is higher than the benchmark case when $\sigma_f = 25\%$. The column denoted (Low Persistence) reports the results from the model with perfect correlated shocks within each firm but with the calibrated parameter values $\kappa = 0.40$ and that σ_f remains at the benchmark level of 25%. However, the persistence level is now lower. The regression coefficients are in percentage terms. The numbers in parenthesis are t-statistics.

	FF	Benchmark	High Variance	Low Persistence
$\log(V_t)$	-0.15	-0.138	-0.134	-0.133
	(-2.58)	(-2.583)	(-2.246)	(-2.669)
$\log[B_t/V_t]$	0.50	0.079	0.084	0.085
	(5.71)	(1.866)	(1.667)	(2.205)
$\log(V_t)$	-0.11	-0.126	-0.120	-0.120
	(-1.99)	(-2.474)	(-2.115)	(-2.502)
$\log[B_t/V_t]$	0.35	0.043	0.040	0.043
	(4.44)	(1.157)	(0.887)	(1.286)
β	-0.37	0.087	0.018	0.080
	(-1.21)	(0.273)	(0.030)	(0.269)
$\log(V_t)$	-0.17	-0.126	-0.131	-0.123
	(-3.41)	(-2.112)	(-1.955)	(-2.203)
β	0.15	0.557	0.488	0.556
	(0.46)	(2.031)	(1.625)	(2.162)

Table 1.12 : Cross-Sectional Return Dispersion As a Predictor of Market Volatility

This table illustrates the intertemporal relation between market volatility and the lagged cross-sectional return dispersion (RD). The volatility is measured by the absolute value of the market excess return. Variations of the following model are estimated:

$$|R_t^e| = a + b_1 RD_{t-1} + b_2 \mathbf{1}_{\{R_{t-1}^e < 0\}} RD_{t-1} + c_1 |R_{t-1}^e| + c_2 \mathbf{1}_{\{R_{t-1}^e < 0\}} |R_{t-1}^e| + \epsilon_t$$

where $|R_t^e|$ ia the absolute value of the market excess return, RD_t is the cross-sectional standard deviation of the individual stock returns, $\mathbf{1}_{\{R_{t-1}^e < 0\}}$ is a dummy variable that equals one when the market excess return is negative and zero otherwise, and ϵ_t is the residual. All t-statistics are adjusted with respect to heteroskedasticity and autocorrelation using Newey-West procedure. For the F-test on joint restrictions, the p-values are in parentheses. Panel A is from Stivers (2000) who uses 400 firm returns from July 1962 to December 1995. Panel B is generated as the average coefficients and statistics across repeated simulations.

Coefficients	b_1	b_2	c_1	c_2	Joint $b_1 = b_2 = 0$	Joint $c_1 = c_2 = 0$	$R^2(\%)$
			Panel A: Results from Stivers (2000)				
Full Model	0.365	0.111	-0.157	0.221	10.08	2.69	10.45
	(3.61)	(1.40)	(-2.94)	(1.84)	(0.000)	(0.069)	
			Panel B: Simulation Results				
Full Model	1.198	-0.008	-0.083	0.172	6.206	2.038	4.10
	(3.09)	(-0.138)	(-1.203)	(1.487)	(0.038)	(0.282)	

Figure 1.1 : Some Key Variables in Competitive Equilibrium

Panel A shows $\bar{e}^{\star}$ or equivalently V^a/K in (1.26). Panel B shows the ratio of total market value to aggregate capital stock, V/K, and Panel C shows the ratio of aggregate value of assets-in-place to total market value, V^a/V. Panel D shows three aggragate level βs, β^a (solid line), $\widetilde{\beta}^a$ (dashed-dotted line), and β^o (dashed line), defined in (1.33).

Panel A: $\bar{e}^*$ or V^a/K

Panel B: V/K

Panel C: V^a/V

Panel D: β^a, $\widetilde{\beta}^a$, and β^o
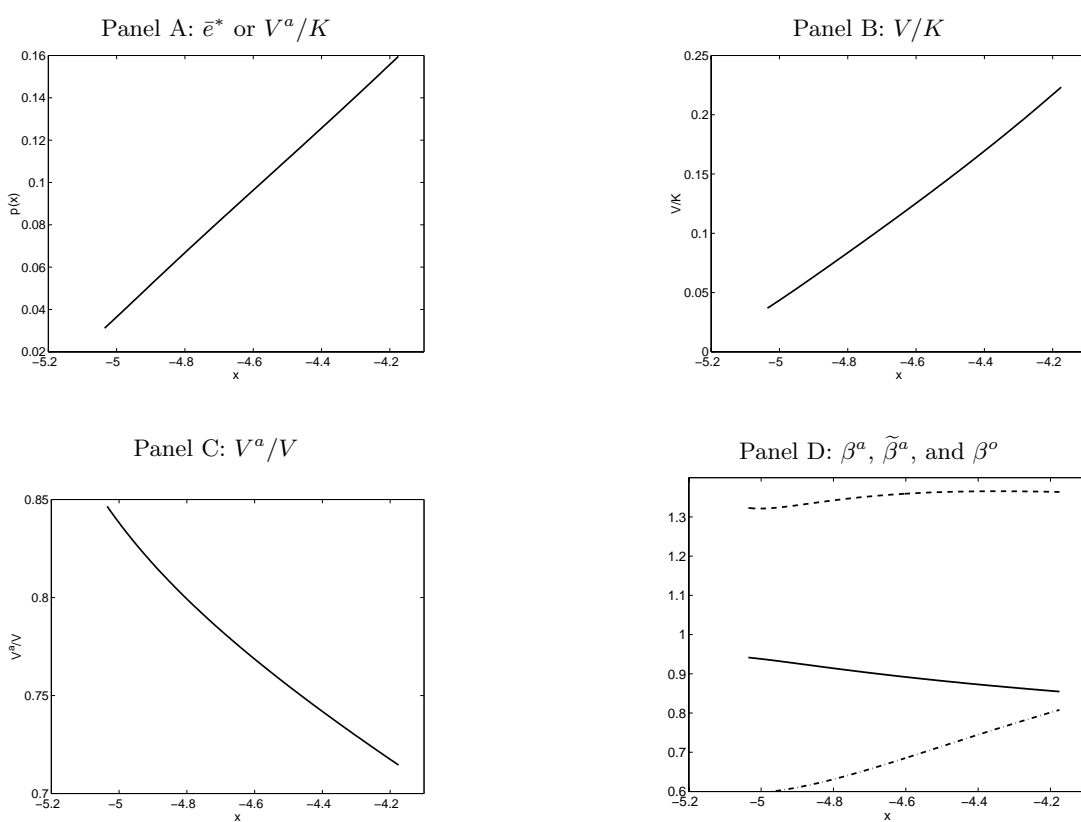
Figure 1.2 : Size and Book-to-Market in Cross-sectional Regressions

Panel A shows the histogram of t-statistic of univariate regressions of returns on size and Panel B shows the histogram of t-statistic of univariate regressions of returns on book-to-market across 100 simulations. Panel C reports the scatter plot of t-statistics on size and book-to-market and Panel D reports the scatter plot of t-statistics on size and Fama-French (FF) β in a joint regression of returns.

Panel A: t-statistic on Size

Panel B: t-statistic on Book-to-Market
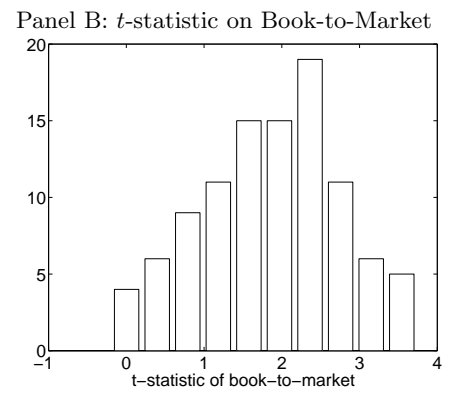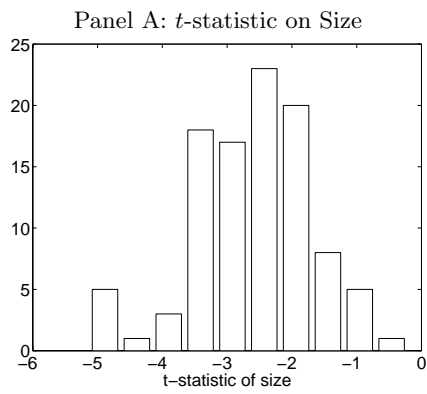
Panel C: Size and Book-to-Market: t-statistics
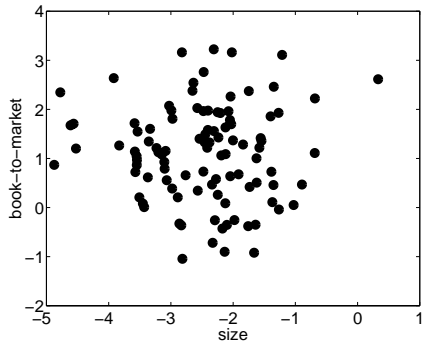
Panel D: Size and FF-β: t-statistics
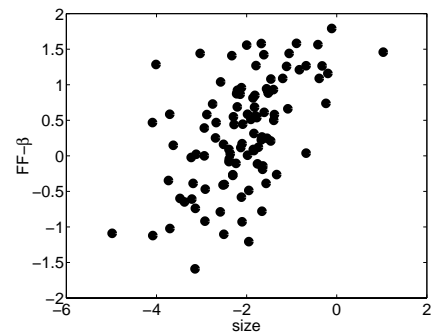
Figure 1.3 : Business Cycle Properties: I

This Figure illustrates the business cycle properties of some aggregate and cross-sectional variables. Panel A plots V^a/V (the solid line) and $\tilde{V}^a/V$ (the dashed line) as functions of x. Panel B plots log price-dividend ratio as a function of $\log(X)$. Panel C plots the size ($\log(V_f)$) dispersion as a function of $\log(V/D)$ and Panel D plots the dispersion of book-to-market ($\log(B_f/V_f)$) as a function of $\log(V/D)$.

Panel A: V^a/V (solid) and $\tilde{V}^a/V$ (dash)

Panel B: $\log(V/D)$

Panel C: Size Dispersion

Panel D: Book-to-Market Dispersion
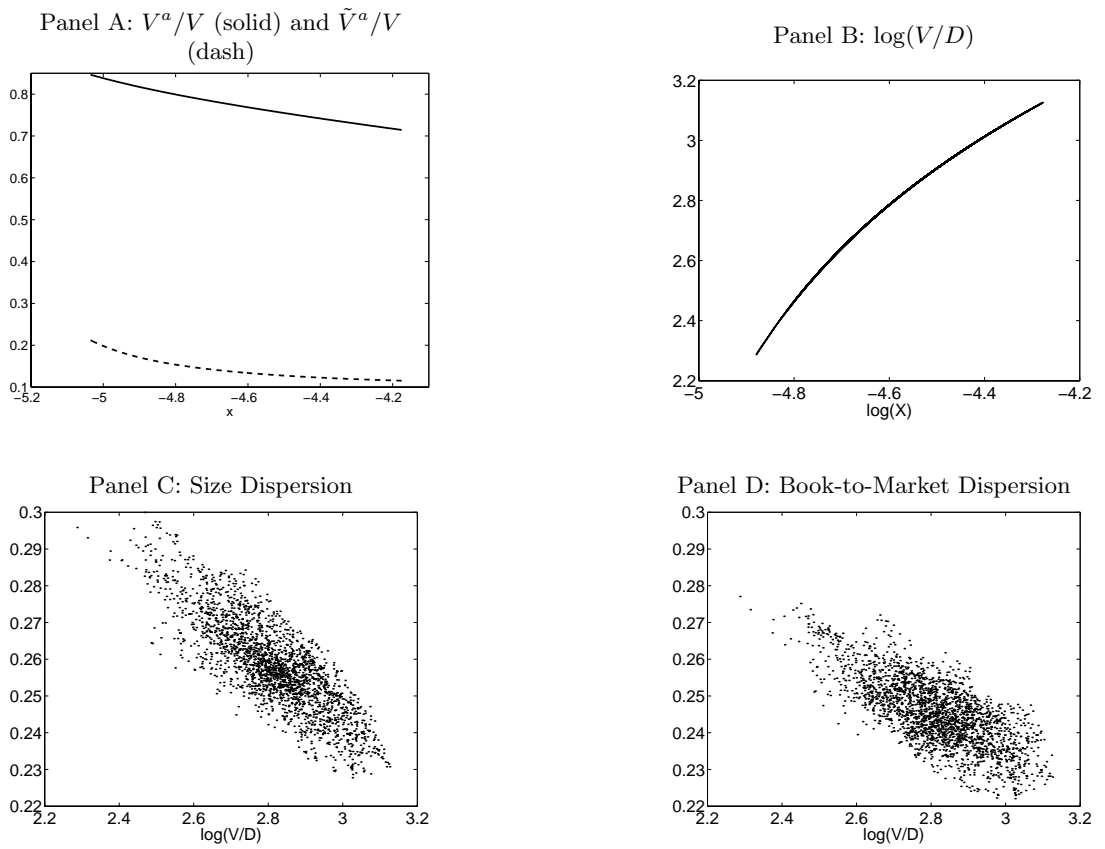
Figure 1.4 : Business Cycle Properties: II

Panel A: Market Volatility

Panel B: Beta Dispersion
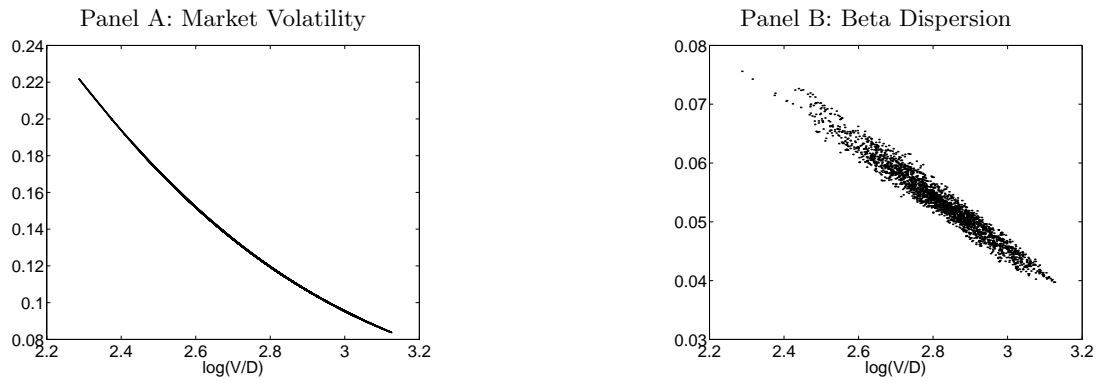
Figure 1.5 : Return Dispersion over Business Cycle

Panel A: Return Dispersion
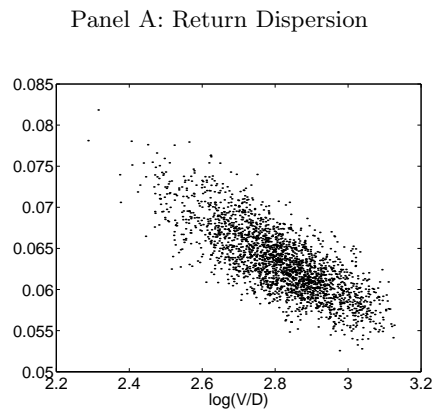
Chapter 2

Asset Pricing Implications of Firms' Financing Constraints

with Joao F. Gomes and Amir Yaron

2.1 Introduction

In this paper we ask whether financing constraints are quantitatively important in explaining a cross-section of expected returns. Specifically, we incorporate costly external finance into a production based asset pricing model and investigate whether financing frictions help in pricing the cross-section of expected returns.

Our analysis, as in Cochrane (1991, 1996), focuses on the link between asset returns and the returns on physical investment, implied by the optimal production and investment decisions of the firm. Our contribution is in augmenting this basic framework to explicitly consider the impact of financing frictions on the optimal decisions of the firm. To avoid specifying the underlying source of these frictions (e.g., asymmetric information, costly state verification or "lemon problems") we show that the typical assumptions about the nature of the financing frictions, as modelled in the existing literature, are captured by a simple "financing cost" function, equal to the product of the financing premium and the amount of external finance. Since both of these quantities are relatively easy to observe, this approach provides a tractable (and fairly general) framework to examine the role of financing frictions in pricing a cross-section of asset returns.

Our empirical analysis uses the Generalized Method of Moments (GMM) to explore the Euler equation restrictions imposed on expected returns by optimal investment behavior. Since this behavior is affected by the presence of the financing frictions, the returns to physical investment will depend on the financing variables. Thus, the ability of investment returns to price the cross section of expected returns will depend not only on "fundamentals" such as profits and investment, as in Cochrane (1996), but also on the financing variables.

As with any asset pricing model, financial frictions will be relevant for the pricing of expected returns only to the extent that they provide a common factor, in this context associated with financial distress as systematic (aggregate) risk. Thus, our focus on the importance of financing frictions through their effects on pricing expected returns seems a natural benchmark from the standpoint of asset pricing.

Our empirical findings suggest that the role played by financing frictions is fairly negligible, unless the premium on external funds is procyclical, a property not evident in the data and not satisfied by most models of costly external finance. Our results are also robust to several alternative formulations of our model, particularly the form of the financing cost function, the specific data used, and the set of returns used in our GMM implementations.

The intuition is simple. Absent financing frictions, firms would increase investment immediately in response to positive news about expected future productivity growth. This, in turn, generates a series of investment returns that lead the cycle, and creates a large correlation between current investment returns and future profits — a feature also documented by Fama and Gibbons (1982) for observed stock returns. In the presence of financing constraints, however, the countercyclical nature of the financing premium implies that the expected rise in future productivity is also associated with lower future expected financing costs. This induces firms to try to capitalize on the lower expected costs by delaying their investment response, which changes the implied dynamics of investment returns and lowers their correlation with the observed stock returns.

Our findings contribute to three strands of the literature in economics and finance. First, from an asset pricing perspective, they suggest that financing variables are not an important factor in pricing the cross-section of asset returns. Although our approach to incorporate financing frictions as a pricing factor is more structural, our results seem to complement those in Lamont, Polk, and Saá-Requejo (2000). Using an aggregate index of financing frictions as a common factor in a reduced form factor pricing model, they also document that the cyclical fluctuations in asset returns do not appear to be linked to financial frictions. Together, these results seem to support recent work that emphasizes the role of firm productivity and growth in generating the observed fluctuations in returns (e.g., Berk, Green, and Naik (1999), Gomes, Kogan, and Zhang (2002), and Zhang (2002)).

These results also have important implications for research on corporate finance regarding the role of financing constraints on firm investment.[1] Since the investment returns are based on the Euler equations implied by the firm's optimal investment behavior, our findings provide an alternative to standard tests of financing constraints focusing on the observed cash-flow sensitivities.

Finally, in the macroeconomic literature, several authors have argued that financing constraints improve the ability of macroeconomic models to replicate the behavior of typical macro aggregates.[2] Our findings suggest, however, that those models' ability to match financial data is severely strained unless the implied costs of external finance are procyclical, thus placing important restrictions on the nature of the financing frictions supported by the data.

In addition to Cochrane (1991, 1996) we also build on additional theoretical work by Restoy and Rockinger (1994) who generalize some of the results in Cochrane (1991) to an environment with investment constraints, and on Bond and Meghir (1994) who explicitly characterize the effects of financing frictions on the optimal investment decisions of the

---

[1] E.g., Fazzari, Hubbard, and Petersen (1988), Hayashi and Inoue (1991), Hoshi, Kashyap, and Scharfstein (1991), Blundell, Bond, Devereaux, and Schiantarelli (1994), Kashyap, Lamont, and Stein (1994), and Kaplan and Zingales (1997).

[2] See, among others, Bernanke and Gertler (1989), Bernanke, Gertler, and Gilchrist (2000), den Haan, Ramey, and Watson (1999), and Kiyotaki and Moore (1997).

firm.

Finally, our work is also related to research by Li, Vassalou and Xing (2001) who compare the performance of alternative investment growth factors in pricing the Fama and French (1993) size and book to market portfolios, and to work by Lettau and Ludvigson (2001) who re-examine the empirical link between aggregate investment and stock returns using information about the consumption to wealth ratio.

The remainder of this paper is organized as follows. Section 2.2 shows that much of the existing literature on firms' financing constraints can be characterized by specifying a simple dynamic problem to describe firm behavior. Section 2.2 also derives the expression for returns to physical investment, and its relation to stock and bond returns, which can be used to evaluate the asset pricing implications of the model. The next section describes our data sources and econometric methods, while Section 2.4 reports the results of our GMM tests and examines both the performance of the model and the role of financing constraints. The robustness of our results to the use of alternative data or modelling assumptions is examined in Section 2.5. Finally, Section 2.6 offers some concluding remarks.

2.2 Investment Based Asset Pricing with Costly External Finance

In this section we incorporate costly external finance in Cochrane's (1996) production based asset pricing framework. We do this by summarizing the common properties of alternative models of financing frictions with a very simple set of restrictions on the costs of external funds. We then show that this formulation leads to a fairly simple characterization of the optimal investment decisions of the firm and derive a set of easily testable asset pricing conditions that shed light on the role of financing frictions.

2.2.1 Modelling Financing Frictions

Theoretical foundations of financing frictions have been provided by several researchers over the years and we do not attempt to provide yet another rationalization for their existence. Rather, we seek to summarize the common ground across much of the existing literature with a representation of financing constraints that is both parsimonious and empirically useful.

While exact assumptions and modelling strategies can differ quite significantly across the various models, most share the key feature that external finance (equity or debt) is more "costly" than internal funds. It is this crucial property that we explore in our analysis below by assuming that the financial market imperfections will be entirely captured by the unit costs of raising external finance.

Consider first the case of equity finance. Suppose a firm issues an amount N in new shares and let W denote the reduction on the claim of existing shareholders associated with the issue of one dollar of new equity. Clearly, in a Modigliani-Miller world $W = 1$ since the total value of the firm is unaffected by financing decisions. If Modigliani-Miller fails to hold however, new equity lowers the total value of the firm, and $W > 1$. Now, new issues are costly to existing shareholders, not only because they reduce claims on future dividends, but because they also reduce value due to the presence of additional transaction or informational costs.[3]

Suppose now that the firm decides to use debt financing, B, and let the function R denote the future repayment costs of this debt.[4] If Modigliani-Miller holds these repayments will just equal the opportunity cost of internal funds, captured by the relevant discount factor for shareholders, M. In this case we will simply have that $\mathrm{E}[MR(\cdot)] = 1$, where $\mathrm{E}[\cdot]$ denotes the expectation over the relevant probability measure. Absent Modigliani-Miller, debt is more costly than internal funds and $\mathrm{E}[MR(\cdot)] > 1$, at

---

[3]E.g., Jensen and Meckling (1976), Myers and Majluf (1984), and Greenwald, Stiglitz, and Weiss (1984)

[4]If there is no possibility of default these costs will just equal the gross interest on the loan. If default is allowed, they may depend on the liquidation value of the firm.

least when $B > 0$.[5]

In addition, it is often assumed that the "financing costs" are increasing in the amount of external finance, so that $\partial W(\cdot)/\partial N$ and $\partial R(\cdot)/\partial B$ are positive. It also seems reasonable to assume that the costs depend on the amount of financing normalized by firm size, K, which allows for the possibility that large firms will face lower financing costs. Finally, these costs may also be state-dependent. In this case we would write $W(\cdot) = W(N/K, X)$, where X summarizes both firm-level and aggregate uncertainty, and similarly $R(\cdot) = R(B/K, X)$.

These additional properties are also common and fairly intuitive. We summarize them in Assumption 1.

Assumption 1 *Let X summarize all forms of uncertainty. The cost functions $W(\cdot)$ and $R(\cdot)$ satisfy:*

$$W(N/K, X) > 1, \quad W_1(\cdot) \equiv \partial W(\cdot)/\partial N \geq 0 \quad \text{for} \quad N > 0 \tag{1}$$

and

$$\mathrm{E}_t[MR(B/K, X)] \geq 1, \quad R_1(\cdot) \equiv \partial R(\cdot)/\partial B \geq 0 \quad \text{for} \quad B > 0 \tag{2}$$

This is an extremely weak assumption as it merely requires that external finance, whether debt or equity, is more expensive than internal funds, with non-decreasing unit costs.

Essentially, the existing corporate finance literature has focused so far on establishing the nature and properties of the functions $W(\cdot)$ and $R(\cdot)$, by focusing on optimal contracts in the presence of, for example, transaction costs, moral hazard, asymmetric information or costly-state verification. These alternative arguments provide different rationales, and sometimes different forms, for the functions $W(\cdot)$ and $R(\cdot)$, but most share the basic

---

[5]E.g., Myers (1977), Townsend (1979), Stiglitz and Weiss (1981), Diamond (1984), Gale and Hellwig (1985), and Bernanke and Gertler (1990)

properties captured by our assumption. By focusing on the common ground across much of this existing literature on financing frictions, we seek to provide a fairly general test of the role of these constraints for asset pricing purposes.[6]

2.2.2 Firm's Problem

Consider now the problem of a firm seeking to maximize the value to existing sharehold-ers, denoted $V(\cdot)$, in an environment where external finance is costly. This firm makes investment decisions by choosing the optimal amount of capital at the beginning of the next period, K_{t+1}. Investment spending, I_t, as well as dividends, D_t, can be financed with internal cash flows $\Pi(\cdot)$, new equity issues, N_t, or one period debt B_{t+1}.[7]

The problem of this firm can then be easily summarized by the following dynamic program:

$$V(K_t, B_t, X_t) = \max_{\substack{D_t, B_{t+1}, \\ K_{t+1}, N_t}} \{D_t - W(N_t/K_t, X_t)N_t + \mathrm{E}_t[M_{t,t+1}V(K_{t+1}, B_{t+1}, X_{t+1})]\} \quad (3)$$

$$\text{s.t.} \quad D_t = C_t + N_t + B_{t+1} - R(B_t/K_t, X_t)B_t \quad (4)$$

$$I_t = K_{t+1} - (1-\delta)K_t, \qquad \delta \geq 0 \quad (5)$$

$$C_t = \Pi(K_t, X_t) - I_t - \frac{a}{2}[I_t/K_t - \delta]^2 K_t \qquad a \geq 0 \quad (6)$$

$$D_t \geq \overline{D}, \quad N_t \geq 0$$

where $M_{t,t+1}$ is the stochastic discount factor (of the owners of the firm) from time t to $t+1$ and $\overline{D}$ is the firm's minimum, possibly zero, dividend payment. Note that firms can accumulate financial assets, in which case debt is negative. The exact nature of the cash flow function, $\Pi(\cdot)$, is assumed to exhibit constant returns scale, but its exact form is unimportant

---

[6]A recent strand of literature on financing frictions focus instead on "quantity" constraints of varying complexity (e.g. Kehoe and Levine (1993), Hart and Moore (1996), Kotcherlakotta (1996), Zhang (1997), Alvarez and Jermann (2000), Albuquerque and Hopenhayn (2001), Clementi and Hopenhayn (2001) and Cooley, Quadrini, and Marimon (2001)). In general these models do not satisfy our assumptions. Strictly speaking then, our characterization below applies only to models of "costly" external finance.

[7]One-period debt simplifies the algebra considerably but has no bearing on our results.

Equation (4) shows the resource constraint of the firm. It implies that dividends must equal internal funds, net of investment spending, C_t, plus new external funds, net of debt repayments. Equation (5) is the standard capital accumulation equation, relating current investment spending, I_t, to future capital, K_{t+1}. We assume that old capital depreciates at the rate δ. As in Cochrane (1991, 1996) also investment requires the payment of adjustment costs, and these are captured by the term $\frac{a}{2}\left[I_t/K_t - \delta\right]^2 K_t$.

Given our assumptions, it is immediate that the firm will only use external finance after internal cash flows are exhausted and no dividends are paid, above the required level $\overline{D}$. Conversely, dividends can only exceed this minimum requirement if no external funds are required to finance them. Hence, the model extends the familiar hierarchical financing derived by Myers (1984) in a static framework to a dynamic setting.

2.2.3 Asset Pricing Implications

To save notation it is convenient to combine the two constraints (4) and (5) by eliminating investment, and noting that $C_t = C(K_t, K_{t+1}, X_t)$. The asset pricing implications of the model can then be summarized by arranging the optimality conditions with respect to K_{t+1} and B_{t+1} to obtain:

$$
\mathrm{E}_t[M_{t,t+1}R_{t+1}^I] = \mathrm{E}_t\left[M_{t,t+1}\left(\frac{V_1(K_{t+1}, B_{t+1}, X_{t+1})}{-\mu_t C_2(K_t, K_{t+1}, X_t)}\right)\right] = 1 \tag{7}
$$

$$
\mathrm{E}_t[M_{t,t+1}R_{t+1}^B] = \mathrm{E}_t\left[M_{t,t+1}\left(\frac{V_2(K_{t+1}, B_{t+1}, X_{t+1})}{-\mu_t}\right)\right] = 1 \tag{8}
$$

where R_{t+1}^I and R_{t+1}^B denote the returns on physical investment and debt, respectively, and μ_t is the Lagrange multiplier on the combined constraint.

Equations (7) and (8), provide a simple summary of the role of financing constraints for the optimal behavior of firms. For empirical purposes however, this characterization is extremely difficult to implement, since it requires an explicit solution to the value function, $V(K_t, B_t, X_t)$ as well as the multiplier, μ_t. More importantly, this procedure also requires an explicit assumption about the nature of the cost functions, $W(\cdot)$ and $R(\cdot)$,

thus rendering our tests below dependent on these restrictions.

Instead, we pursue an alternative approach by exploiting the fact that the solution to the problem above, can be characterized by solving the following "frictionless" problem

$$\widetilde{V}(K_t, B_t, X_t) = \max_{K_{t+1}} \left\{ \widetilde{C}(K_t, K_{t+1}, X_t) + \mathrm{E}_t \left[M_{t,t+1} \widetilde{V}(K_{t+1}, B_{t+1}, X_{t+1}) \right] \right\}, \quad (9)$$

where

$$\widetilde{C}(K_t, K_{t+1}, X_t) = C(K_t, K_{t+1}, X_t) - b(X_t) \times E_t \quad (10)$$

and $\widetilde{V}(\cdot)$ denotes the total value of the firm for both stock and bond holders.

The linear term $b(X_t) \times E_t$ captures the role of the financing frictions. Here, $b(X_t) \geq 0$ is the (possibly stochastic) premium on external finance, and $E_t = B_{t+1} + N_t$ denotes the total amount of external finance used by the firm. Using the resource constraint E_t can be computed as:

$$E_t = B_{t+1} + N_t = R_t B_t + \overline{D} - C(K_t, K_{t+1}, X_t) \quad (11)$$

Proposition 6 formally establishes the equivalence between the formulation in (9) and the original problem in (3).[8]

Proposition 6 *Let the adjusted cash flow function $\widetilde{C}(\cdot)$ be given by (10). Then the two dynamic programs (3) and (9) produce the same solution.*

Proof We present the proof for the case of equity only finance. The proof for the case with debt is provided in the appendix. When firms issue only new equity, $B_t = B_{t+1} = 0$, and $E_t = N_t$. Replacing the resource constraints in (3) we obtain

$$V(K_t, X_t) = \max_{K_{t+1}, N_t} \left\{ C(K_t, K_{t+1}, X_t) - (W(\cdot) - 1)N_t + \mathrm{E}_t \left[M_{t,t+1} V(K_{t+1}, X_{t+1}) \right] \right\}$$

---

[8]Gilchrist and Himmelberg (1998) use a similar cost representation for the case of debt finance. Effectively Proposition 6 rationalizes their result for a much larger class of models.

Letting $b(X_t) = (W(\cdot) - 1)$ be the premium on external finance, it follows that

$$\widetilde{C}(K_t, K_{t+1}, X_t) = C(K_t, K_{t+1}, X_t) - (W(\cdot) - 1)N_t.$$

∎

While the proof for the case of debt financing requires a fairly elaborate verification of integrability conditions, the basic argument of the proof lies in the characterization of the multiplier. In some sense this proposition merely explores the fact that one can always rewrite a constrained problem as an unconstrained one with embedded multipliers. What is novel here is the precise characterization of the multiplier, μ_t, as a measure of the premium on external finance. By linking this "shadow-price" to an essentially observable variable, we are able to recast the problem in a way that lends itself to empirical analysis.

Our financing cost function provides a very simple, but quite general, characterization of the financing constraints. It implies that they can be effectively summarized by the product of two terms, one, E_t, which captures the amount of external finance raised, and the other, $b(\cdot)$, summarizing the premium on external funds.

In addition, the optimality conditions from the "frictionless" problem (9) imply that the return on investment equals:

$$R_{t+1}^I(i, \pi, b) = \frac{(1 + b(X_{t+1}))(\pi_{t+1} i_{t+1} + \frac{a}{2} i_{t+1}^2 + (1 + a i_{t+1})(1 - \delta))}{(1 + b(X_t))(1 + a i_t)} \qquad (12)$$

where $i \equiv (I/K)$ is the investment to capital ratio, and $\pi \equiv (\Pi/I)$ is the profit to investment ratio, and we have used the fact that the amount of external finance is a function of these two variables to eliminate E_t. To complete our description of investment returns all we need is a specification for the premium on external finance. While several measures can be used it seems natural to start by assuming that this cost is proportional in the default premium, DF_t:

$$b(X_t) = b \times DF_t \qquad b \geq 0.$$

and b is a parameter to be estimated in our empirical work.[9] Thus, investment returns are entirely driven by the two "fundamentals", i and π, as well as the cyclical properties of the financing premium. This implementation is very appealing from an empirical point of view, since it requires only a measure of the premium on external finance as well as data on profits, investment and financing variables.

Finally, our approach also provides a measure of the economic magnitude of the financing costs. Specifically, the ratio of these costs to investment spending provides a meaningful measure of the importance of the financing costs. Hence our alternative characterization provides not only a useful tool for empirical analysis but also a simple and straightforward measure of the magnitude of the financing costs.

2.3   Investment Based Factor Pricing Models

This section describes our empirical methodology in detail and it provides an overview of our data sources and the construction of the series of returns.

2.3.1   Asset Pricing Tests

The essence of our strategy is to use the information contained in the asset prices restrictions above to formally investigate the importance of financing constraints. As we have seen above, these restrictions are summarized by the Euler equations:

$$\mathrm{E}_t(M_{t,t+1}R^I_{n,t+1}) = \mathrm{E}_t(M_{t,t+1}R^B_{l,t+1}) = 1 \tag{13}$$

for investment returns, $R^I_{n,t+1}$, $n = 1,2,...J_I$, and bond returns $R^B_{l,t+1}$, $l = 1,2,...J_B$. In addition, Proposition 7 shows a similar restriction must also hold for stock returns $R^S_{j,t+1}$, $j = 1,2,..J_S$.

---

[9]The exact form of the financing costs is $b_1(\cdot)B_{t+1} + b_2(\cdot)N_t$. However, (12) holds exactly as long as both costs are proportional to the default premia. While this is might be an oversimplification in the case of equity, what will matter for our results are the cyclical properties of the financing premium, and these are likely to be similar for both sources of external finance.

Proposition 7 *Stock returns satisfy the following conditions*

$$\mathrm{E}_t(M_{t,t+1}R_{t+1}^S) = 1 \tag{14}$$

$$R_{t+1}^I = \omega_t R_{t+1}^S + (1 - \omega_t)R_{t+1}^B \tag{15}$$

where $(1 - \omega_t)$ is the leverage ratio.

Proof   See Appendix A ■

Although the proof is somewhat elaborate, equation (15) merely states that stock returns are a weighted average of investment and bond returns. Given (15) and (13) it is immediate to verify that stock returns must satisfy the moment condition (14).

Equations (13)-(15) offer two alternative ways to examine the asset pricing implications of financing frictions. While the identity (15) focuses on *ex-post* returns, the Euler equations (13) and (14) are entirely about *expected* returns. Thus, while firm specific risks may play an important role in examining the former, for the latter only systematic risk is relevant.

In Gomes, Yaron, and Zhang (2002) we investigate the importance of these idiosyncratic components using firm level data. Here, we concentrate on the role financing frictions play in pricing the cross-section of *expected* returns, by focusing only on aggregate factors. Specifically then, we use a pricing kernel that depends only on the returns to aggregate investment and a bond index:

$$M_{t,t+1} = l_0 + l_1 R_{t+1}^I + l_2 R_{t+1}^B, \tag{16}$$

a specialization that only requires individual returns to be approximately linear in aggregate returns.[10]   In the context of production based asset pricing this approach seems a

---

[10]From Harrison and Kreps (1979) and Hansen and Richard (1987) we know that one pricing kernel that satisfies (13) is $M_{t,t+1} = \sum_j l_j R_j^S + \sum_n l_n R_n^I + \sum_l l_l R_l^B$. Stock returns can be eliminated since (15) implies that only two of these returns are independent. For using aggregate investment return, we formally only need that $R_{d,t+1}^I \approx \gamma_d^0 + \gamma_d^1 R_{t+1}^I + \epsilon_{d,t+1}$ for portfolio d and the $\epsilon_{d,t+1}$ be *i.i.d.* This is only a statement about technologies and not about market completeness, and it appears reasonable provided

reasonable first step. Cross-sectional variations in firms's investment opportunities may be important in pricing asset returns only to the extent that they affect some aggregate systematic risk. Unlike the consumption-based literature on asset pricing, where the use of the cross-sectional distribution was motivated by the lack of success of aggregate consumption-based models (see Constantinides and Duffie (1996)), aggregate investment returns actually work very well in pricing the cross-section of returns (Cochrane (1996)); thus, the scope for firm heterogeneity affecting the systematic risk for financial distress seems fairly limited. [11]

As we can see from (12), information about the degree of financial frictions is contained in investment returns, which will then serve as a factor capturing the extent to which aggregate financial conditions are priced. In this sense, our formulation is essentially a structural version of an APT-type framework in which one of the factors proxies for an aggregate distress variable (and where different portfolios have varying loading on this factor), such as that taken in Fama and French (1993,1996) and Lamont, Polk, and Saá-Requejo (2000).

In essence then, our metric for evaluating whether financing frictions are important is whether they show as a systematic risk for the cross section of returns. This seems a natural benchmark from the standpoint of asset pricing.

2.3.2 Econometric Methodology

Our estimation strategy allows us to estimate factor loadings, $\mathbf{l}$, as well as the parameters, a and b, by utilizing M as specified in (16) in conjunction with moment conditions (13).

We follow Cochrane's (1996) estimation techniques for assessing the asset pricing implications of our model. Specifically, three alternative sets of moment conditions in implementing (13) are examined. First, we look at the relatively weak restrictions implied

that the level of portfolio disaggregation is not too fine, as will be the case.

[11]It is important to note, however, that, in principle, there is no problem in modifying our approach to include measures of cross-section variation across firms in the pricing kernel, by adding more disaggregated investment returns. For example, Li, Vassalou and Xing (2001) study the effects of cross-sectional variation by including investment growth in five separate sectors in their construction of the pricing kernel.

by the unconditional moments. We then focus on the conditional moments by scaling returns with instruments, and finally we look at time variation in the factor loadings, by scaling the factors.

For the unconditional factor pricing we apply standard GMM procedures to estimate the cost parameters, a and b, and loading factors, $\mathbf{l}$, by simply minimizing a weighted average of the sample moments (13). Letting $\sum_T$ denote the sample mean we can rewrite these moments, $\mathbf{g}_T$ as:

$$\mathbf{g}_T \equiv \mathbf{g}_T(a, b, \mathbf{l}) \equiv \sum_T [M\mathbf{R} - \mathbf{p}]$$

where $\mathbf{R} = [R^S, R^I(\mathbf{y}; a, b), R^B]$ is the menu of asset returns being priced, $\mathbf{p} = [\mathbf{1}, 1, 1]$ is a vector of prices, and $\mathbf{y} = (i, \pi, DF)$. One can then choose $(a, b, \mathbf{l})$ to minimize a weighted sum of squares of the pricing errors across assets:

$$J_T = \mathbf{g}_T' \mathbf{W} \mathbf{g}_T \tag{17}$$

A convenient feature of our setup is that given a and b, the criterion function above is linear in $\mathbf{l}$ — the factor loading coefficients. Standard χ^2 tests of over-identifying restrictions follow from this procedure. This also provides a natural framework to assess whether the loading factors or technology parameters are important for pricing assets. Note that the investment return appears both in the pricing kernel as well as part of the menu of assets being priced. As Cochrane (1996) notes, this consistency is required so investment returns do not have arbitrary properties.

It is straightforward to include the effects of conditioning information by scaling the returns and/or scaling the factors by instruments. The essence of this exercise lies in extracting the conditional implications of (13) since, for a time-varying conditional model, these implications may not be well captured by a corresponding set of unconditional moment restrictions as was noted by Hansen and Richard (1987).

To test conditional predictions of (13), we expand the set of returns to include

returns scaled by instruments to obtain the moment conditions:

$$\mathrm{E}\left[\mathbf{p}_t \otimes \mathbf{z}_t\right] = \mathrm{E}\left[M_{t,t+1}\left(\mathbf{R}_{t+1} \otimes \mathbf{z}_t\right)\right]$$

where $\mathbf{z}_t$ is some instrument in the information set at time t and $\otimes$ denotes the Kronecker product.

A more direct way to extract the potential non-linear restrictions embodied in (13) is to let the stochastic discount factor be a linear combination of factors with weights that vary over time. That is, the vector of factor loadings $\mathbf{l}$ is a function of instruments $\mathbf{z}$ that vary over time:[12] Therefore, to estimate and test a model in which factors are expected to price assets only conditionally, we simply expand the set of factors to include factors scaled by instruments. The stochastic discount factor utilized in estimating (13) is then,

$$M_{t,t+1} = \left[l_0 + l_1 R_{t+1}^I + l_2 R_{t+1}^B\right] \otimes z_t$$

2.3.3 Data

This section provides an overview of the data used in our study. A more detailed description is provided in Appendix 2.7. Our data for the economic aggregates comes from NIPA and the Flow of Funds Accounts. Information about financial assets is obtained from CRSP and Ibbotson. The construction of investment returns requires data on profits, investment and capital. Capital consumption data is used to compute the time series average of the depreciation rate and pin down the value of δ, the only technology parameter that we do not formally estimate. To avoid measurement problems due to chain weighting in the earlier periods our sample starts in the first quarter of 1954 and ends in the last quarter of 2000. Since versions of our model are generally used to describe the non-financial firms we construct series on investment, capital and profits of the Non-Financial Corporate Sector alone. For comparison purposes, we also report results for

---

[12]With sufficiently many powers of $z's$ the linearity of l can actually accommodate nonlinear relationships.

the aggregate economy. Investment data are quarterly averages, while asset returns are from the beginning to the end of the quarter. As a correction, we follow Cochrane (1996) and average monthly asset returns over the quarter and then shift them so they go from approximately the middle of the initial quarter to the middle of the next quarter.[13]

In order to implement the estimation procedure, we require a sufficient number of moment conditions. As described above, we limit ourselves to examining the model's implications for aggregate investment and bond returns. This means that we need to look at more than just the aggregate stock return. Thus, we focus on the ten size portfolios of NYSE stocks. Table 2.1 reports the summary statistics of these asset returns. In addition, we also provide results for the 25 Fama and French (1993) size and book-to-market portfolio returns. Bond data comes from Ibbotson's index of Long Term Corporate Bonds. The default premium is defined as the difference between the yields on AAA and BAA corporate bonds, both obtained from DRI.

Conditioning information comes from two sources: the term premium, defined as the yield on ten year notes minus that on three-month Treasury Bills, and the dividend-price ratio of the equally weighted NYSE portfolio. We follow Cochrane (1996) and limit the number of moment conditions and scaled factors in three ways: (1) we do not scale the Treasury-Bill return by the instruments since we are more interested in the time-variation of risk premium than that of risk-free rate. (2) Instruments themselves are not included as factors. (3) We use only deciles one, two, five, and ten in the conditional estimates.

2.4 Results

2.4.1 GMM Estimates

Table 2.2 reports iterated GMM estimates for the unconditional, conditional, and scaled models. First-stage estimates are very similar, particularly concerning the role of financing costs. In all cases we report the value of the parameters a and b as well as the estimated

---

[13]See also Lamont (2001) and Lettau and Ludvigson (2001) for a discussion of the important consequences of aligning investment and asset returns.

72

loadings l and corresponding t-statistics. Also included are the results of J tests on the model's overall ability to match the data, the corresponding p-values, and the root mean square (RMSE) of the pricing errors, α — mean return less predicted mean return.

Our model is quite successful at pricing the cross-section of returns. In spite of the inclusion of the last few years of stock market data, the model cannot be rejected using the overidentifying restriction tests, J_T. The root mean squared errors are all low (in particular when we use both investment and bond returns as pricing factors) — suggesting the statistical significance of the J tests is not due to an excessively large covariance matrix.[14] This is verified by Figure 2.1 that plots predicted versus actual mean excess returns from first stage estimation, and it clearly displays the goodness of fit of the model. In addition, the hypothesis that all factor loadings are zero is almost always rejected at the standard 5% significance level.

Although our model uses only a single aggregate investment return as a pricing factor (in addition to the corporate bond return) these results are generally comparable to Cochrane's (1996) findings. The reason for this empirical success is that our construction of investment returns, R^I, uses independent information on variations in the marginal productivity of capital, π_t, and investment, i_t. Cochrane (1996) on the other hand, abstracts from the variation in the marginal product of capital in constructing investment returns and hence uses two separate investment series (residential and non-residential) to construct two investment returns.[15]

Although our model requires the use of two pricing factors (R^I and R^B), our results are essentially the same whether or not we use bond returns as a pricing factor. The estimated loadings on the corporate bond returns are also statistically insignificant, suggesting that their role in pricing financial assets is fairly minor.

---

[14]RMSE (α) are actually cut in half if we truncate our sample in 1997.

[15]Economically, our estimate for a is also quite sensible, since it implies that adjustment costs are about 8-9% of total investment spending, which is comparable to Cochrane's (1996) estimate.

2.4.2 The Effect of Financing Constraints

The focus of our analysis, however, is the role of the financing cost parameter b. Here the message from all panels is very clear. In all cases the actual point estimate of b is exactly zero![16]

Why are the financing constraints not useful in pricing the cross-section of expected returns? Alternatively, why do they seem irrelevant for the construction of the stochastic discount factor? The answer lies in the countercyclical properties of the premium on external finance.

The Financing Premium

To gain some intuition on the role of the financing frictions on the pricing kernel, consider their impact on investment returns by decomposing (12) as:

$$R_{t+1}^I \approx \frac{1 + b(X_{t+1})}{1 + b(X_t)} \hat{R}^I \tag{18}$$

where $\hat{R}^I$ denotes investment returns with no financing costs. Loosely, this return summarizes the effects of the fundamentals, and is determined by the cyclical properties of both profits and physical investment. The role of the financing frictions is captured by $\frac{1+b(X_{t+1})}{1+b(X_t)}$.

Figure 2.2 displays the correlation structure between DF_{t+1}/DF_t, $\hat{R}_{t+1}^I$, R_{t+1}^I with a positive b, and R_{t+1}^S, with leads and lags of i_t (Panel A) and π_t (Panel B). The pattern is striking. In both cases, the pattern of $\hat{R}^I$ is very similar to that of observed R^S. Both returns *lead* future economic activity, while their contemporaneous correlation with fundamentals is somewhat low. As Cochrane (1991) notes, this is to be expected if firms adjust current investment in response to an anticipated rise in future productivity.

The behavior of the default premium, however, is quite different. Its negative correlation with future economic activity implies a series of investment returns that behaves

---

[16]Note that since costs can not be negative, values of a or b below zero are not admissible.

much less like the observed stock returns, thus straining the ability of R_{t+1}^I, inclusive of financing constraints, to be a useful pricing factor.

Alternatively, since a rise in expected future productivity (or profits) is associated with an expected decline in the financing premia (because of its counter-cyclical properties), there is an incentive for the firm to delay its investment response in the presence of financing constraints. From equation (12) we learn that this lowers investment returns. Given the observed pattern of stock returns in the data this leads to a lower correlation between investment and stock returns.

To summarize, productivity and financing costs provide two competing forces that determine the reaction of investment, and hence investment returns, to business cycle conditions. Productivity implies that firms should respond by investing immediately. On the other hand, since the future entails lower financing costs firms should delay investment. Figure 2.2 shows that consistency with asset return data requires the financing channel to be unimportant.

Figure 2.2 also suggests that these results are not likely to rely on timing issues such as those created by the existence of time to plan (or perhaps time to finance in this context), since there is no obvious phase shift between the premium and the return series.

What seems crucial is the countercyclical pattern of the premium on external finance, induced by the behavior of the default premium. However, since almost any realistic measure of the cost of external finance would exhibit this same countercyclical pattern, our conclusions should easily survive the use of alternative measures of the unit costs of external finance.

Limitations of Reduced Form Analysis

It is important to point out the benefits of imposing the theoretical restrictions, implied by our structural approach, in our estimation strategy. An alternative and common approach is simply to allow for some measure of financial distress (say F_{t+1}) to appear as a factor in an APT-like model. An example would be to model the pricing kernel as

$M_{t+1} = l_0 + l_1 R_{m,t+1} + l_2 F_{t+1}$, without any restriction on the sign and magnitude of l_2. The fact that financing frictions appear explicitly as costs, in our framework, requires that $b \geq 0$, since costs can not be negative. Ignoring this restriction by allowing $b < 0$ also reverses the cyclical properties of the financing costs, a feature that would enhance the correlation between the return on investment and profits. This in turn would lead one to conclude that financing frictions are relevant for pricing assets without realizing that it implies negative financing costs.

The Pricing Kernel

Financing frictions obviously change the dynamics of the pricing kernel. Table 2.3 shows a few statistical measures of the way these frictions influence the pricing kernel and pricing errors. It describes the effects of increasing the value of b in each set of moment conditions, while a is kept constant at its optimal level reported in columns 5–7 of Table 2.2.

As we can readily observe, the presence of financing constraints effectively lowers the market price of risk $\sigma(M)/E(M)$, as well as the (absolute) correlation between the pricing kernel and value-weighted returns for all three models, thus deteriorating the performance of the pricing kernel. Perhaps more direct evidence is given by examining the implied pricing errors. A simple way of doing this is to compute the beta representation:

$$R_i - R_f = \alpha_i + \beta_{1i}(R^I - R_f) + \beta_{2i}(R^B - R_f)$$

Given the assumed structure of the pricing kernel this representation exists, with $\alpha_i = 0$ (see discussion in Cochrane (2001)). Therefore, large values of α are evidence against the model. Table 2.3 reports the implied αs for the regressions on both decile 1 (small firms) and value-weighted returns. It displays a clear pattern of increasing α as we increase the magnitude of the financing costs. Indeed, while we cannot reject that $\alpha = 0$ when $b = 0$, this hypothesis is rejected for most of the other parameter configurations.

We also report the implications of financing costs for the raw moments of investment

returns and their correlation with market returns. While both the mean and the variance of investment returns are not changed by much as b increases (at least initially), the main implication of increasing financing constraints is to lower their correlation with asset returns. Since the overall performance of a factor model hinges on its covariance structure with returns, it is not surprising that financing costs are not important for the construction of the pricing kernel as documented in Table 2.2.[17]

2.5 Robustness

This section examines the robustness of our results by exploring several alternatives to our benchmark approach.

2.5.1 Small Firms Effects

Several studies on firm financing constraints emphasize that they are more likely to be detected when looking only at the behavior of small firms. Although our focus is on the implications for aggregate asset prices, an easy way to assess the model's implications for different firms is to test the moment conditions (13) for only small firms. We investigate this possibility in Table 2.4. Also included are the χ^2-statistics and corresponding p-values for the relevant Wald tests when our estimate of b is non-zero. As columns 2–4 show, even in this case we cannot find any evidence for a significant role of financing frictions. Even when b is slightly positive, the hypothesis that it is statistically zero can only be rejected at extremely high significance levels.

2.5.2 Fama-French Portfolios

Several authors interpret the cross-sectional variation in the Fama and French (1993) size and book-to-market portfolio returns as proxies for some measure of relative financial

---

[17]An alternative way of representing the impact of financing constraints is to compare their effect on the pricing kernels with the Hansen-Jagannathan (1991) bounds. Increasing b has the effect of moving the estimated kernels farther way from the bounds.

distress. Columns 5–7 in Table 2.4 report the results when our model is used to price the 25 Fama and French (1993) portfolio returns. However, the estimated value of b is zero, again suggesting that financing frictions do not play a crucial role in determining the cross-section of returns.

2.5.3 Different Macroeconomic Data

Table 2.5 shows the effects of using alternative data in the construction of the investment returns. Columns 2–4 report the results of using after tax profits in the construction of investment returns, while columns 5–7 report similar results when data on overall macroeconomic aggregates is used. It is easy to see that these alternative constructions have no impact on our main conclusions from Table 2.2.

2.5.4 Non-Linear Pricing Kernels

The use of a linear factor representation may be restrictive, and several alternative approaches modelling nonlinear pricing kernels have been recently advanced in the literature.[18] We explore this possibility by re-estimating the moment conditions using several nonlinear pricing kernels. Specifically, we consider examples where the pricing kernel is quadratic in either R^I alone or in both R^I and R^B. Again, as columns 2–7 in Table 2.6 show, none of these cases reveals any evidence for financing costs.

2.5.5 Alternative Cost Functions

While our financing cost function is derived from first principles, given our model's assumptions, we can also use our methodology to investigate the consequences of using alternative, less structural, functional forms. While these may not correspond exactly to the underlying constrained problem in (3), they may nevertheless provide a useful approximation for empirical purposes.

---

[18]E.g., Bansal and Vishwanathan (1993), Chapman (1997), and Brandt and Yaron (2001).

In this section we explore the implications of a simple alternative characterization of the cost function:

$$(b \times DF_t \times E_t) \times E_t = b \times DF_t \times E_t^2,$$

where the term $b \times DF_t \times E_t$ now captures the premium which multiplies external finance, E_t. Quadratic cost functions of this form correspond to some popular models of financing frictions, such as that in Stein (2001). Intuitively they correspond to the assumption that the premium on external finance, $b(\cdot)$, is linear in the amount of external finance raised.

Columns 8–10 in Table 2.6 confirm that this modification has a negligible impact on our results. Even when the actual point estimate of b is not exactly zero, the hypothesis that it differs from zero is easily rejected.

2.6 Conclusion

Despite its empirical success, the production based asset pricing model (Cochrane (1991, 1996)) has been, until recently, relatively neglected by researchers, in favor of either standard consumption based or APT-like asset pricing models. This is unfortunate since, by concentrating on optimal firm behavior, this approach provides a natural way of integrating new developments in the theory of corporate finance into an asset pricing framework.

In this paper we pursue this line of research by incorporating costly external finance in a production based asset pricing model and ask whether financing frictions help in pricing the cross-section of expected returns. To avoid specifying the underlying source of these frictions we show that the typical assumptions about the nature of the financing frictions are captured by a simple "financing cost" function, which provides a tractable framework to examine the role of financing frictions in pricing asset returns.

Our empirical findings suggest that the role played by financing frictions is fairly negligible, unless the premium on external funds is procyclical, a property not evident in the data and not satisfied by most models of costly external finance. This finding is robust to several alternative formulations of our model, particularly the form of the

financing cost function, the specific macroeconomic data used, and the set of returns used in our GMM implementations.

These findings question whether financing frictions are important for explaining the cross-section of returns and for determining investment behavior. Moreover, our results also cast doubt on whether financing constraints provide a realistic propagation mechanism in several macroeconomic models.

A few aspects of our empirical implementation suggest promising directions for future research. First, investment may have an important time-to-build component, and financing procedures may precede the actual investment spending by a quarter or more, leading firms to look at lagged measures of fundamentals when making their decisions. Although our results suggest that this explanation is unlikely to account for the rejection of financing frictions, only an explicit examination of the potential time aggregation implications can formally address this issue. Second, there remains the issue of the proper level of aggregation. Although financing constraints seem to play no role in determining the *portfolio* returns in this paper, they may still be fairly important at the individual firm level. Since our model holds for every firm however, it can also be used to investigate this issue, by looking directly at firm level implications.

Bibliography

[1] Abel, B. Andrew, and Janice C. Eberly, 2001, Q Theory Without Adjustment Cost and Cash Flow Effects Without Financing Constraints, Working Paper, The Wharton School at the University of Pennsylvania.

[2] Alvarez, Fernando, and Urban J. Jermann, 2000, Efficiency, Equilibrium, and Asset Pricing With Risk of Default, *Econometrica*, 68 (4), 775–797.

[3] Berk, Jonathan B., Richard C. Green and Vasant Naik, 1999, Optimal Investment, Growth Options and Security Returns, *Journal of Finance*, 54, 1513–1607.

[4] Bansal Ravi, and S. Viswanathan, 1993, No Arbitrage and Arbitrage Pricing: A New Approach, *Journal of Finance*, 48, 1231–1261.

[5] Bernanke, Ben and Mark Gertler, 1989, Agency Costs, Net Worth, and Business Fluctuations, *American Economic Review*, 79 (1), 14–31.

[6] Bernanke, Ben, Mark Gertler, and Simon Gilchrist, 2000, The Financial Accelerator in a Quantitative Business Cycle Framework, in *Handbook of Macroeconomics*, Edited by Michael Woodford and John Taylor, North Holland.

[7] Blundell, Richard W., Bond, Stephen R., Devereaux, Michael P. and Fabio Schiantarelli, 1992, Investment and Tobin's Q: Evidence from Company Panel Data, *Journal of Econometrics* 51, 233–257.

[8] Brandt, Michael W., and Amir Yaron, 2001, Time-Consistent No-Arbitrage Models of the Term Structure, Working Paper, The Wharton School at the University of Pennsylvania.

[9] Bond, Stephen and Costas Meghir, 1994, Dynamic Investment Models and the Firm's Financial Policy, *Review of Economic Studies*, 61, 197–222

[10] Carlstrom, Charles T., and Timothy S. Fuerst, 1997, Agency Costs, Net Worth, and Business Fluctuations: A Computable General Equilibrium Analysis, *American Economic Review*, 87 (5), 893–910.

[11] Chapman, David A., 1997, Approximating the Asset Pricing Kernel, *Journal of Finance*, 52, 1383–1410.

[12] Cochrane, John H., 1991, Production-Based Asset Pricing and the Link Between Stock Returns and Economic Fluctuations, *Journal of Finance*, 46 (1), 209–237.

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

[14] Cochrane, John H., 2001, *Asset Pricing*, Princeton University Press, Princeton, New Jersey.

[15] Cooley, Thomas F., Vincenzo Quadrini, and Ramon Marimon, 2001, Aggregate Consequences of Limited Contract Enforceability, Working Paper, Stern School of Business, New York University.

[16] Den Haan, Wouter, Gary Ramey and Joel Watson, 1999, Liquidity Flows and Fragility of Business Enterprises, Working Paper 7057, National Bureau of Economic Research.

[17] Diamond, Douglas W., 1984, Financial Intermediation and Delegated Monitoring, *Review of Economic Studies*, LI, 393–414.

[18] Fama, Eugene F., and Kenneth R. French, 1993, Common Risk Factors in the Returns on Stocks and Bonds, *Journal of Finance*, 33, 3–56.

[19] Fama, Eugene F., and Kenneth R. French, 1996, Multifactor Explanations of Asset Pricing Anomalies, *Journal of Finance*, 51, 55–84.

[20] Fazzari, Stephen, R. Glenn Hubbard, and Bruce Peterson, 1988, Financing constraint and Corporate Investment, *Brookings Papers on Economic Activity*, 1, 141–195.

[21] Gale, Douglas, and Martin Hellwig, 1985, Incentive-Compatible Debt Contracts: The One-Period Problem, *Review of Economic Studies*, LII, 647–663.

[22] Gilchrist, Simon and Charles Himmelberg, 1998, Investment: Fundamentals and Finance, in NBER Macroeconommics Annual, Ben Bernanke and Julio Rotemberg eds, MIT Press

[23] Gomes, Joao F., 2001, Financing Investment, *American Economic Review*, 90, 5, 1263-1285.

[24] Gomes, Joao F., Leonid Kogan, and Lu Zhang, 2001, Equilibrium Cross-Section of Returns, Working Paper, The Wharton School at the University of Pennsylvania.

[25] Gomes, Joao F., Amir Yaron and Lu Zhang, 2002, Asset Pricing Implications of Firm's Financing Constraints: Firm Level Evidence, mimeo, The Wharton School at the University of Pennsylvania.

[26] Greenwald, Bruce, Joseph Stiglitz and Andrew Weiss, 1984, Informational Imperfections in the Capital Market and Macroeconomic Fluctuations, *American Economic Review*, 74, 194-199.

[27] Hansen, Lars Peter, 1982, Large Sample Properties of Generalized Method of Moments Estimation, *Econometrica*, 50, 1029–1054.

[28] Hansen, Lars Peter, and S. Richard, 1987, The role of Conditioning Information in Deducing Testable Restrictions Implied by Dynamic Asset Pricing Models, *Econometrica*, 55, 587–613.

[29] Harrison, Michael and David Kreps, 1979, Martingales and Arbitrage in Multiperiod Securities Markets, *Journal of Economic Theory*, 20, 381–408.

[30] Hayashi, Fumio, and Inoue, T., 1991, The Relation Between Firm Growth and Q with Multiple Capital Goods: Theory and Evidence from Japanese Panel Data, *Econometrica,* 59, 731–754.

[31] Hoshi, T., Anil Kashyap and D. Scharfstein, 1991, Corporate Structure, Liquidity and Investment: Evidence from the Japanese Industrial Groups, *Quarterly Journal of Economics* 106, 33–60.

[32] Jermann, Urban J., 1998, Asset Pricing in Production Economies, *Journal of Monetary Economics*, 41, 257–275.

[33] Kaplan, Steve and Luigi Zingales, 1997, Do Financing Constraints Explain why Investment is Correlated with Cash-Flow, *Quarterly Journal of Economics*, 112 (1), 168–216.

[34] Kashyap, Anil, Lamont, Owen and Jeremy Stein, 1994, Credit Conditions and the Cyclical Behavior of Inventories, *Quarterly Journal of Economics*, 109 (3), 565–592.

[35] Kehoe, Timothy, and David K. Levine, 1993, Debt-Constrained Asset Markets, *Review of Economic Studies*, 60, 868–888.

[36] Kiyotaki, Nobuhiro, and John Moore, 1997, Credit Cycles, *Journal of Political Economy*, 105 (2), 211–248.

[37] Kocherlakota, Narayana R., 1996, Implications of Efficient Risk Sharing Without Commitment, *Review of Economic Studies*, 63, 595–609.

[38] Lamont, Owen, Investment Plans and Stock Returns, 2000, *Journal of Finance*, LV (6), pp. 2719–2745

[39] Lamont, Owen, Christopher Polk and Jesús Saá-Requejo, 2001, Financial Constraints and Stock Returns, forthcoming, *Review of Financial Studies*.

[40] Lettau, Martin, and Sidney Ludvigson, 2001, Time-Varying Risk Premia and the Cost of Capital: An Alternative Implication of the Q Theory of Investment, Working Paper, New York University.

[41] Li, Qing, Maria Vassalou, and Yuhang Xing, 2001, An Investment Growth Asset Pricing Model, Working Paper, Columbia University.

[42] Myers, Stewart C., 1977, Determinants of Corporate Borrowing, *Journal of Financial Economics*, 5, 147–175.

[43] Myers, Stewart C., 1984, The Capital Structure Puzzle, *Journal of Finance*, 39 (3), 575–592.

[44] Myers, Stewart C., and Majluf, N. 1984, Corporate Financing and Investment Decisions When Firms Have Information That Investors Do Not Have, *Journal of Financial Economics*, 13, 187–222.

[45] Newey, Whitney, and Kenneth West, 1987, Hypothesis Testing with Efficient Method of Moments Estimation, *International Economic Review*, 28, 777–787.

[46] Restoy, Fernando, and G. Michael Rockinger, 1994, On Stock Market Returns and Returns on Investment, *Journal of Finance*, 49 (2), 543–556.

[47] Stein, Jeremy, 2001, Agency, Information, and Corporate Investment, NBER Working Paper 8342.

[48] Stiglitz, Joseph and Andrew Weiss, 1981, Credit Rationing and Markets with Imperfect Information, *American Economic Review*, 71, 393–411.

[49] Townsend, Robert, 1979, Optimal Contracts and Competitive Markets with Costly State Verification, *Journal of Economic Theory*, 21, 265–293.

[50] Zhang Harold, 1997 Endogenous Borrowing Constraints with Incomplete Markets. *Journal of Finance*, 52, 2187-2205.

[51] Zhang, Lu, 2002, The Value Premium, Working Paper, The Wharton School at the University of Pennsylvania.

Proofs and Technical Results for Chapter Two

To prove Proposition 6 we need to establish the following proposition first.

Proposition 8 *When debt is positive, the multiplier μ_t satisfies the following conditions:*

$$\frac{\partial \mu_t}{\partial K_t} = \frac{\partial \mu_t}{\partial B_t} = 0$$

Proof The envelope conditions for respect to K_t and B_t imply:

$$
\begin{aligned}
V_{21}(K_t, B_t, X_t) &= -\frac{\partial \mu_t}{\partial K_t}\left[R(B_t/K_t) + R_1(B_t/K_t)(B_t/K_t)\right] \\
&\quad + \mu_t\left[R_1(B_t/K_t)(2B_t/K_t^2) + R_{11}(B_t/K_t)(B_t^2/K_t^3)\right] \quad\text{(A1)} \\
V_{22}(K_t, B_t, X_t) &= -\frac{\partial \mu_t}{\partial B_t}\left[R(B_t/K_t) + R_1(B_t/K_t)(B_t/K_t)\right] \\
&\quad - \mu_t\left[R_1(B_t/K_t)(2/K_t) + R_{11}(B_t/K_t)(B_t/K_t^2)\right]
\end{aligned}
$$

Now homogeneity of the value function implies that

$$
\begin{aligned}
0 &= V_{21}(K_t, B_t, X_t)K_t + V_{22}(K_t, B_t, X_t)B_t \\
&= -\left[R(B_t/K_t) + R_1(B_t/K_t)(B_t/K_t)\right]\left(\frac{\partial \mu_t}{\partial K_t}K_t + \frac{\partial \mu_t}{\partial B_t}B_t\right)
\end{aligned}
$$

thus confirming that μ_t is indeed homogeneous of degree zero in K_t and B_t.

Now since

$$
\begin{aligned}
V_{21}(K, B, X) &= V_{12}(K, B, X) = \frac{\partial \mu_t}{\partial B_t}\left[C_1(K_t, K_{t+1}, X_t) + R_1(B_t/K_t)(B_t/K_t)^2\right] \\
&\quad + \mu_t\left[R_1(B_t/K_t)(2B_t/K_t^2) + R_{11}(B_t/K_t)(B_t^2/K_t^3)\right] \quad\text{(A2)}
\end{aligned}
$$

Equating (A1) and (A2) and simplifying yields

$$-\frac{\partial \mu_t}{\partial K_t}\left[R(B_t/K_t) + R_1(B_t/K_t)(B_t/K_t)\right] = \frac{\partial \mu_t}{\partial B_t}\left[C_1(K_t, K_{t+1}, X_t) + R_1(B_t/K_t)(B_t/K_t)^2\right]$$

Thus,

$$\frac{\partial \mu_t}{\partial K_t}R(B_t/K_t) + \frac{\partial \mu_t}{\partial B_t}C_1(K_t, K_{t+1}, X_t) = \left(\frac{\partial \mu_t}{\partial K_t}K_t + \frac{\partial \mu_t}{\partial B_t}B_t\right)R_1(B_t/K_t)(B_t/K_t^2) = 0$$

Therefore, the derivatives of μ_t satisfy the following two conditions

$$
\begin{aligned}
\frac{\partial \mu_t}{\partial K_t}R(B_t/K_t) + \frac{\partial \mu_t}{\partial B_t}C_1(K_t, K_{t+1}, X_t) &= 0 \\
\left(\frac{\partial \mu_t}{\partial K_t}K_t + \frac{\partial \mu_t}{\partial B_t}B_t\right) &= 0
\end{aligned}
$$

But since $B_t > 0$

$$R(B_t/K_t)B_t + C_1(K_t, K_{t+1}, X_t)K_t > 0$$

and we must have that

$$\frac{\partial \mu_t}{\partial K_t} = \frac{\partial \mu_t}{\partial B_t} = 0$$

■

Proof of Proposition 6. In the case of debt financing only, investment returns can be written as:

$$R_{t+1}^I = \frac{\mu_{t+1}\left[C_1(K_{t+1}, K_{t+2}, X_{t+1}) + R_1(B_{t+1}/K_{t+1})(B_{t+1}/K_{t+1})^2\right]}{-\mu_t C_2(K_t, K_{t+1}, X_t)} \tag{A3}$$

Define the function:

$$G(K_t, K_{t+1}, X_t) = (\mu_t - 1)B_{t+1} \tag{A4}$$

it follows that

$$G_1(K_t, K_{t+1}, X_t) = -(\mu_t - 1)\left[C_1(K_t, K_{t+1}, X_t) + R_1(B_t/K_t)(B_t/K_t)^2\right] \tag{A5}$$

$$G_2(K_t, K_{t+1}, X_t) = -(\mu_t - 1)C_2(K_t, K_{t+1}, X_t) \tag{A6}$$

Integration of (A6) yields

$$G(K_t, K_{t+1}, X_t) = \int G_2(K_t, K_{t+1}, X_t)\, dK_{t+1} = -(\mu_t - 1)C(K_t, K_{t+1}, X_t) + f_1(K_t, X_t)$$

where $f_1(\cdot)$ is independent of K_{t+1}. Using Proposition 8 we know that the integral of (A5) equals

$$G(K_t, K_{t+1}, X_t) = -(\mu_t-1)C(K_t, K_{t+1}, X_t) - (\mu_t-1)\int R_1(B_t/K_t)(B_t/K_t)^2\, dK_t + f_2(K_{t+1}, X_t)$$

where $f_2(\cdot)$ is independent of K_t. Combining two equations above yields

$$G(K_t, K_{t+1}, X_t) = (\mu_t - 1)\left[R(B_t/K_t)B_t + \overline{D} - C(K_t, K_{t+1}, X_t)\right] = (\mu_t - 1)B_{t+1}$$

where the second equality follows from (4) and the fact that $B_t > 0 \Longrightarrow D_t = \overline{D}$. Equation (A3) now implies that:

$$R_{t+1}^I = \frac{C_1(K_{t+1}, K_{t+2}, X_{t+1}) - G_1(K_{t+1}, K_{t+2}, X_{t+1})}{-C_2(K_t, K_{t+1}, X_t) + G_2(K_t, K_{t+1}, X_t)} = \frac{\widetilde{C}_1(K_{t+1}, K_{t+2}, X_{t+1})}{-\widetilde{C}_2(K_t, K_{t+1}, X_t)}$$

■

To prove Proposition 7 we need to establish the following proposition first.

Proposition 9 *The value of the firm equals the sum of (cum-dividend) equity value and the value of outstanding debt:*

$$q_t K_t = V(K_t, B_t, X_t) + \mu_t B_t\left[R(B_t/K_t) + R_1(B_t/K_t)(B_t/K_t)\right] \tag{A7}$$

where $q_t = V_1(K_t, B_t, X_t)$ denotes the marginal q. Moreover, (A7) implies that marginal q equals Tobin's (average) q.

Proof For simplicity consider the case where $\overline{D} = 0$. Rewrite the value of the firm as

$$V(K_t, B_t, X_t) = \max_{\substack{D_t, B_{t+1}, \\ K_{t+1}, N_t}} \left\{ \begin{array}{c} (1 - \mu_t + \lambda_t^d)D_t + [\mu_t - W(N_t/K_t) + \lambda_t^n]N_t + \mu_t[C(K_t, K_{t+1}, X_t) \\ + B_{t+1} - R(B_t/K_t)B_t] + \mathrm{E}_t\left[M_{t,t+1}V(K_{t+1}, B_{t+1}, X_{t+1})\right] \end{array} \right\}$$

86

The complementarity-slackness conditions imply that the first term on the right-hand side is zero and the second equals $W_1 (N_t/K_t)(N_t/K_t)N_t$.

Next, homogeneity of the value function and the envelope conditions imply that:

$$E_t\left[M_{t,t+1}V(K_{t+1},B_{t+1},X_{t+1})\right] = -\mu_t C_2(K_t,K_{t+1},X_t)K_{t+1} - \mu_t B_{t+1}$$

while homogeneity of C yields

$$C_1(K_t,K_{t+1},X_t)K_t = C(K_t,K_{t+1},X_t) - C_2(K_t,K_{t+1},X_t)K_{t+1}$$

Hence the value function collapses to

$$V(K_t,B_t,X_t) = W_1(N_t/K_t)(N_t/K_t)N_t + \mu_t\left[C_1(K_t,K_{t+1},X_t)K_t - R(B_t/K_t)B_t\right]$$

Rearranging, and using the envelope condition, we have:

$$V(K_t,B_t,X_t) + \mu_t\left[R(B_t/K_t)B_t + R_1(B_t/K_t)(B_t/K_t)B_t\right] = V_1(K_t,B_t,X_t)K_t$$

∎

Proof of Proposition 7. By definition stock returns are given by

$$R_{t+1}^S = \frac{V^e(K_{t+1},B_{t+1},X_{t+1}) + [D_{t+1} - W(N_{t+1}/K_{t+1},X_{t+1})N_{t+1}]}{V^e(K_t,B_t,X_t)}, \qquad (A8)$$

where

$$V^e(K_t,B_t,X_t) \equiv V(K_t,B_t,X_t) - [D_t - W(N_t/K_t,X_t)N_t] \qquad (A9)$$

is the (current period) value of the firm to shareholders after new issues take place and dividends are paid.

Again consider the simple case where $\overline{D} = 0$. Starting from the definition of investment returns (12), we have

$$R^I = \frac{V_1(K_{t+1},B_{t+1},X_{t+1})}{-\mu_t C_2(K_t,K_{t+1},X_t)} = \frac{V_1(K_{t+1},B_{t+1},X_{t+1})}{\mu_t\left[C_1(K_t,K_{t+1},X_t)K_t - C(K_t,K_{t+1},X_t)\right]} \qquad (A10)$$

$$= \frac{V(K_{t+1},B_{t+1},X_{t+1}) + \mu_{t+1}B_{t+1}\left[R(B_{t+1}/K_{t+1}) + R_1(B_{t+1}/K_{t+1})(B_{t+1}/K_{t+1})\right]}{V(K_t,B_t,X_t) - \mu_t D_t + \mu_t B_{t+1} + N_t\left[\mu_t - W_1(N_t/K_t)(N_t/K_t)\right]} \qquad (A11)$$

where the second equality follows from homogeneity of C, and the third from the envelope condition and Proposition 9. Next, observe that the complementarity slackness conditions imply:

$$
\begin{aligned}
D_t(1-\mu_t) &= 0 \\
N_t[\mu_t - W_1(N_t/K_t)(N_t/K_t)] &= W(N_t/K_t)N_t
\end{aligned}
$$

Thus

$$R_{t+1}^I = \frac{V(K_{t+1},B_{t+1},X_{t+1}) + \mu_{t+1}B_{t+1}\left[R(B_{t+1}/K_{t+1}) + R_1(B_{t+1}/K_{t+1})(B_{t+1}/K_{t+1})\right]}{V(K_t,B_t,X_t) - D_t + \mu_t B_{t+1} + W(N_t/K_t)N_t}$$

Using the definitions of R_{t+1}^S, R_{t+1}^B it follows that:

$$R_{t+1}^I = (1 - \omega_t)R_{t+1}^S + \omega_t R_{t+1}^B$$

where the leverage ratio, ω_t, equals

$$\omega_t = \frac{\mu_t B_{t+1}}{V^e(K_t, B_t, X_t) + \mu_t B_{t+1}}. \tag{A12}$$

With this result established, it follows immediately that

$$1 = \mathrm{E}_t\left[M_{t,t+1}R_{t+1}^S(1 - \omega_t)\right] + \mathrm{E}_t\left[M_{t,t+1}R_{t+1}^B\omega_t\right] = (1 - \omega_t)\mathrm{E}_t\left[M_{t,t+1}R_{t+1}^S\right] + \omega_t$$

or, simply

$$\mathrm{E}_t\left[M_{t,t+1}R_{t+1}^S\right] = 1 \tag{A13}$$

∎

2.7   Data Construction

Macroeconomic data comes from NIPA, published by the BEA, and the Flow of Funds Accounts, available from the Federal Reserve System. These data are cross-referenced and mutually consistent, so they form, for practical purposes, a unique source of information. Most of our experiments use data for the Nonfinancial Corporate Sector. Specifically Table F102 is used to construct measures of profits before (item FA106060005) and after tax accruals (item FA106231005). To these measures we add both consumption capital (item FA106300015) and inventory valuation (item FA106020601) adjustments to obtain a better indicator of actual cash flows. Investment spending is gross investment (item 105090005). The capital stock comes from Table B102 (Item FL102010005). Since stock valuations include cash flows from operations abroad, we also include in our measures of profits the value of foreign earnings abroad (item FA266006003) and that of net foreign holdings to the capital stock (items FL103092005 minus FL103192005, from Table L230) and investment (the change in net holdings). Financial liabilities come also from Table B102. They are constructed by subtracting financial assets, including trade receivables, (Item FL104090005) from liabilities in credit market instruments (Item FL104104005) plus trade payables (Item FL103170005). Interest payments come from NIPA Table 1.16, line 35. All these are available at quarterly frequency and require no further adjustments. Series for the aggregate economy come from NIPA.

Financial data come from CRSP and Ibbotson. We use the ten size portfolios of NYSE stocks (CRSP series DECRET1 to DECRET10). Corporate bond data comes from Ibbotson's index of Long Term Corporate Bonds. The default premium is defined as the difference between the yields on AAA and Baa corporate bonds, from CRSP. Term premium, defined as the yield on 10 year notes minus that on three-month Treasury bills, and the dividend-price ratio of the equally weighted NYSE portfolio (constructed from CRSP EWRETD and EWRETX).[19]

---

[19]Dividend-price ratios are also normalized so that scaled and non-scaled returns are comparable.

Table 2.1 : Summary Statistics of the Assets Returns in GMM

This table reports the means, volatilities, Sharpe ratios, and first-order autocorrelations of excess returns of deciles 1–10, excess value-weighted market return (*vwret*), real t-bill rate (R^F), and corporate bond return (R^B). These returns are used in GMM estimation and tests. The sample period is from 1954:2Q to 2000:3Q. Means and volatilities are in annualized percent.

| | Decile Returns | | | | | | | | | | vwret | R^F | R^B |
	1	2	3	4	5	6	7	8	9	10			mean
mean	11.80	9.49	9.03	9.07	8.50	8.57	7.67	8.16	7.34	6.64	7.10	1.86	0.51
std	19.61	17.49	16.73	16.16	15.49	15.19	14.51	13.80	12.90	11.35	11.87	1.32	7.23
Sharpe	0.60	0.54	0.53	0.55	0.54	0.56	0.52	0.58	0.56	0.57	0.58	0.00	0.09
$\rho(1)$	0.26	0.29	0.29	0.31	0.29	0.28	0.32	0.27	0.27	0.36	0.33	0.67	0.29

Table 2.2 : GMM Estimates and Tests — The Benchmark

This table reports GMM estimates and tests of the benchmark model with linear G function where $b_t = b \times \mathrm{DF}_t$ and DF_t is the default premium. Investment return series are constructed from flow of funds accounts using nonfinancial profits before tax. T-statistics are reported in parentheses to the right of parameter estimates. Finally, we also report the root mean square pricing error α — mean return less predicted mean return — in percentage per quarter, where pricing errors are calculated as $\alpha_j = 100 \times \mathrm{E}[MR_j - p_j]/\mathrm{E}[M]$, the χ^2 statistic and corresponding p-value for the J_T test on over-identification, and p-values of the Wald test on the null hypothesis that $a = 0$. We conduct GMM estimates and tests for the unconditional model, unscaled and scaled conditional model, for both one-factor and two-factor specifications of the pricing kernel. The unconditional model uses as moment conditions the excess returns of 10 CRSP size decile portfolio and one investment return and the real Treasury-bill return (12 moment conditions). The unscaled and scaled conditional models use the deciles 1, 2, 5, 10, and investment returns, scaled by instruments, and the real Treasury-bill return (16 moment conditions). Instruments are the constant, term premium (tp), and equally weighted dividend-price ratio (dp). So the scaled factor model in the one-factor case features pricing kernel $M = l_0 + l_1 R^I + l_2 (R^I \cdot tp) + l_3 (R^I \cdot dp)$ and in the two-factor case $M = l_0 + l_1 R^I + l_2 R^B + l_3 (R^I \cdot tp) + l_4 (R^I \cdot dp) + l_5 (R^B \cdot tp) + l_6 (R^B \cdot dp)$.

	One Factor Model						Two Factor Model					
	Unconditional		Conditional		Scaled Factor		Unconditional		Conditional		Scaled Factor	
						Parameters						
a	9.63	(3.98)	8.42	(6.16)	8.90	(4.93)	13.65	(2.30)	9.71	(3.88)	8.79	(4.41)
b	0.00		0.00		0.00		0.00		0.00		0.00	
						Loadings						
l_0	60.50	(1.83)	108.79	(4.29)	96.13	(3.32)	55.02	(1.64)	103.01	(3.41)	89.68	(2.52)
l_1	-58.65	(-1.81)	-106.16	(-4.26)	-93.52	(-3.27)	-41.94	(-1.30)	-90.43	(-2.77)	-109.65	(-3.11)
l_2					-0.19	(-1.13)	-11.40	(-1.92)	-10.11	(-1.52)	23.10	(1.73)
l_3					0.08	(0.48)					9.81	(3.29)
l_4											6.37	(1.01)
l_5											-10.28	(-3.36)
l_6											-6.40	(-0.98)
					Root Mean Squared Pricing Error							
α	1.15		2.13		0.45		0.30		0.67		0.33	
						J_T Test						
χ^2	14.01		18.62		16.55		8.27		15.60		8.20	
p	0.08		0.10		0.08		0.31		0.16		0.32	

Table 2.3 : Properties of Pricing Kernels, Jensen's α, and Investment Returns

This table reports, for each combination of parameters a and b, properties of the pricing kernel, including market price of risk ($\sigma[M]/\mathrm{E}[M]$), the contemporaneous correlation between pricing kernel and real market return (ρ_{M,R^S}), Jensen's α and its corresponding t-statistic (t_α), summary statistics of investment return, including mean, volatility (σ_{R^I}), first-order autocorrelation ($\rho(1)$), and correlation with the real value-weighted market return (ρ_{R^I,R^S}). Jensen's α is defined from the following regression: $R^p - R^f = \alpha + \beta_1(R^I - R^f) + \beta_2(R^B - R^f)$ where R^p is either real value-weighted market return (R^{vw}) or real decile one return (R^1), R^f is real interest rate proxied by real treasury-bill rate, R^I is investment return, and R^B is real corporate bond return. In each case the cost parameters a's are held fixed at their GMM estimates. The assets returns used in the unconditional estimates are the 10 CRSP size decile portfolio, one investment excess return, one corporate bond excess return, and the real treasury-bill return. The assets returns used in the conditional estimates, in both unscaled and scaled model, are the deciles 1, 2, 5, 10 returns, and investment and corporate bond excess returns, scaled by instruments, plus the real Treasury-Bill return. Instruments are the constant, term premium, and equally weighted dividend-price ratio. θ_2 is the share of financing cost in investment.

b	θ_2	Pricing Kernel		Jensen's α				Investment Return			
		$\frac{\sigma[M]}{\mathrm{E}[M]}$	ρ_{M,R^S}	α^{vw}	t_α^{vw}	α^{d1}	t_α^{d1}	mean	σ_{R^I}	$\rho(1)$	ρ_{R^I,R^S}
					Unconditional Model						
0.00	0.00	0.76	-0.51	0.54	1.21	1.25	1.59	5.54	3.13	0.00	0.35
0.25	0.04	0.67	-0.20	1.74	3.65	3.37	4.12	5.56	3.35	-0.04	-0.10
0.50	0.09	0.41	-0.03	2.16	4.88	3.98	5.29	5.58	4.85	0.11	-0.29
					Conditional Model						
0.00	0.00	1.07	-0.47	0.25	0.51	0.81	0.95	5.82	2.44	0.06	0.35
0.25	0.05	0.69	-0.24	1.22	2.05	4.15	4.66	5.85	3.01	0.06	-0.19
0.50	0.10	0.72	-0.19	2.12	2.35	5.32	5.74	5.87	4.74	0.18	-0.35
					Scaled Factor Model						
0.00	0.00	1.31	-0.31	0.17	0.33	0.70	0.80	5.89	2.27	0.08	0.35
0.25	0.05	1.04	-0.01	2.15	4.32	4.11	4.84	5.92	2.94	0.09	-0.21
0.50	0.10	0.78	0.10	2.40	5.44	4.37	5.85	5.94	4.73	0.20	-0.36

Table 2.4 : GMM Estimates and Tests — Alternative Moment Conditions

This table reports results of GMM estimates and tests of the benchmark model with alternative sets of moment conditions. Under alternative one, unconditional model uses the excess returns of CRSP size deciles 1, 2, and 3 portfolios and one investment excess return, and the real Treasury-bill return (5 moment conditions). The conditional estimates, in nonscaled and scaled model, use the deciles 1 and 2 and investment excess returns, scaled by instruments, and the real Treasury-bill return (10 moment conditions). Under alternative two, the unconditional model uses the excess returns of portfolios 11, 13, 15, 21, 23, 25, 41, 43, 45, 51, 53, 55 of the Fama and French (1993) 25 portfolios, one investment excess return, and real Treasury-bill return (14 moment conditions). The FF portfolios are numbered such that the first digit denotes the size group and the second digit denotes the book-to-market group, both of which are in ascending order. For example, portfolio 15 denotes the one formed from the intersection of smallest size and highest book-to-market ratio. The conditional estimates, in nonscaled and scaled model, use excess returns of FF portfolio 11, 15, 33, 51, and 55, scaled by instruments, and the real Treasury-bill return (19 moment conditions). For simplicity, only results for the two factor specification of the pricing kernel are presented.

	Small Deciles						FF Portfolios					
	Unconditional		Conditional		Scaled Factor		Unconditional		Conditional		Scaled Factor	
	Parameters											
a	1.13	(0.12)	8.60	(3.07)	8.16	(2.45)	22.61	(2.34)	18.04	(2.95)	10.37	(4.01)
b	0.00		0.015	(1.24)	0.00		0.00		0.00		0.00	
	Root Mean Squared Pricing Error											
α	0.10		0.69		0.07		0.78		1.08		0.42	
	J_T Test											
χ^2	-		9.66		2.46		43.91		49.80		28.85	
p	-		0.09		0.12		0.00		0.00		0.00	
	Wald Test ($a=0$)											
$\chi^2_{(1)}$	0.77											
p	0.38											
	Wald Test ($b=0$)											
$\chi^2_{(1)}$			0.93									
p			0.33									

# Table 2.5 : GMM Estimates and Tests — Alternative Measures of Profits

This table reports GMM estimates and tests of the benchmark model with a linear G (as in Table 2.2) using alternative sources of data. Specifically, we consider two alternatives for profit series: nonfinancial profits after tax and aggregate (both financial and nonfinancial) profits. T-statistics are reported in parentheses to the right of parameter estimates. Finally, we also report the root mean square pricing error α — mean return less predicted mean return — in percentage per quarter, where pricing errors are calculated as $\alpha_j = 100 \times \mathrm{E}[MR_j - p_j]/\mathrm{E}[M]$, the χ^2 statistic and corresponding p-value for the J_T test on over-identification, and p-values of Wald test on the null hypothesis that $a = 0$. We conduct GMM estimates and tests for the unconditional model, the unscaled and scaled conditional models, for both one-factor and two-factor specifications of the pricing kernel. The unconditional model uses as moment conditions the excess returns of 10 CRSP size decile portfolio and one investment return, and the real Treasury-bill return (12 moment conditions). The unscaled and scaled conditional model use the deciles 1, 2, 5, 10, and investment returns, scaled by instruments, and the real Treasury-bill return (16 moment conditions). Instruments are the constant, term premium (tp), and equally weighted dividend-price ratio (dp). For brevity, only results for two factor specifications of the pricing kernel are presented.

| | Nonfinancial After Tax | | | | | | Aggregate Profits | | | | | |
	Unconditional		Conditional		Scaled Factor		Unconditional		Conditional		Scaled Factor	
	Parameters											
a	4.16	(2.32)	4.60	(2.85)	3.67	(3.70)	7.36	(1.12)	13.61	(3.16)	6.35	(1.61)
b	0.00		0.00		0.01	(0.70)	0.00		0.00		0.00	
	Root Mean Square Pricing Error											
α	0.27		0.67		0.26		0.22		0.54		0.29	
	J_T Test											
χ^2	4.67		14.60		10.46		9.56		17.72		9.04	
p	0.70		0.20		0.16		0.21		0.09		0.25	
	Wald Test ($a=0$)											
$\chi^2_{(1)}$							5.89				11.17	
p							0.02				0.00	
	Wald Test ($b=0$)											
$\chi^2_{(1)}$					0.35							
p					0.55							

Table 2.6 : GMM Estimates and Tests — Alternative Specifications

This table reports GMM estimates and tests of two alternative pricing kernels M and one alternative G function. Alternative one of M assumes that $M=l_0+l_1R^I+l_2(R^I)^2$, where R^I is the investment return. Alternative two uses that $M=l_0+l_1R^I+l_2R^B+l_3(R^I)^2+l_4(R^B)^2$, where R^B denotes real corporate bond return. The alternative G is that $G=b_t(RB_t+I_t+H_t-\Pi_t)^2/K_t$, where RB_t denotes outstanding debt including interest and $b_t=b\times DF_t$ with DF_t being the default premium. The investment return series is constructed from flows of funds accounts using nonfinancial profits before tax. The unconditional model uses as moment conditions the excess returns of 10 CRSP size decile portfolio and one investment return and real Treasury-bill return (12 moment conditions). The unscaled and scaled conditional models use the deciles 1, 2, 5, 10, and investment returns, scaled by instruments, and the real Treasury-bill return (16 moment conditions). Instruments are the constant, term premium (tp), and equally weighted dividend-price ratio (dp).

|	$M=l_0+l_1R^I+l_2(R^I)^2$			$M=l_0+l_1R^I+l_2R^B+l_3(R^I)^2+l_4(R^B)^2$			$G=b_t(RB_t+I_t+H_t-\Pi_t)^2/K_t$		
	Unconditional	Conditional	Scaled Factor	Unconditional	Conditional	Scaled Factor	Unconditional	Conditional	Scaled Factor
					Parameters				
a	9.35 (4.35)	9.60 (4.87)	9.89 (4.87)	11.02 (2.49)	16.12 (1.88)	17.72 (0.55)	15.13 (1.73)	9.29 (4.28)	8.97 (4.02)
b	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.025 (0.63)	0.04 (0.68)
				Root Mean Squared Pricing Error					
α	0.29	1.01	0.44	0.18	0.45	0.11	0.34	0.70	0.33
					J_T Test				
χ^2	10.99	21.28	10.47	6.42	8.72	4.59	8.24	15.09	7.95
p	0.14	0.09	0.16	0.27	0.46	0.03	0.31	0.18	0.34
					Wald Test ($a=0$)				
$\chi^2_{(1)}$					52.61	1.37	18.33		
p					0.00	0.24	0.00		
					Wald Test ($b=0$)				
$\chi^2_{(1)}$								0.70	1.57
p								0.40	0.20

Figure 2.1 : Predicted Versus Actual Mean Excess Returns

This figure plots the mean excess returns against predicted mean excess return, both of which are in % per quarter, for conditional model (Panel A), conditional model (Panel B), and scaled factor model (Panel C). All plots are from first-stage GMM estimates.

Panel A: Unconditional Estimates
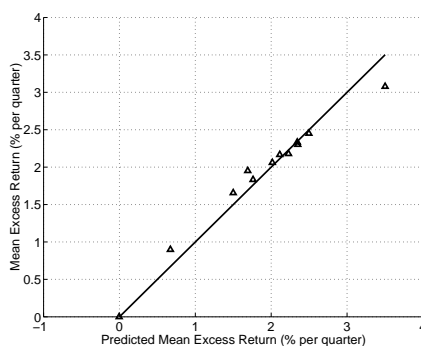
Panel B: Conditional Estimates
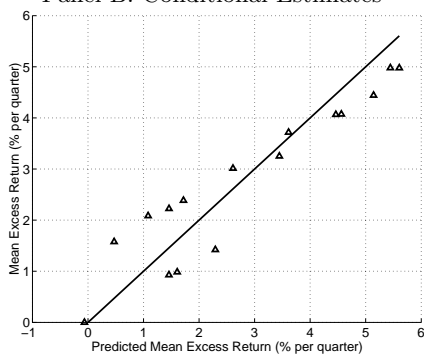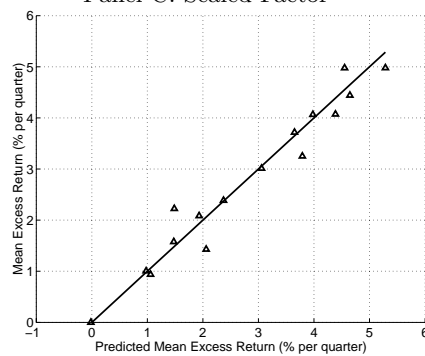

Panel C: Scaled Factor

95

Figure 2.2 : Correlation Structure

This figure presents the correlations of investment return R^I, real value-weighted market return R^S, the growth rate of default premium DF_{t+1}/DF_t with I/K and Π/K with various leads and lags. Panel A plots the correlation structure of the above series with I/K and Panel B plots that with Π/K.

Panel A: Correlation With I/K Panel B: Correlation With Π/K
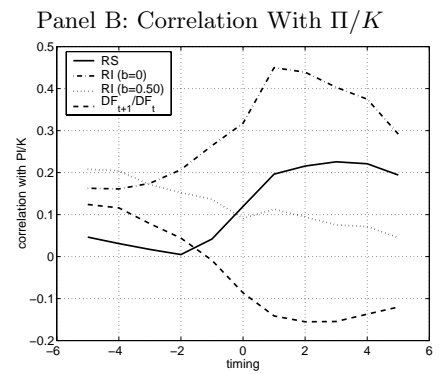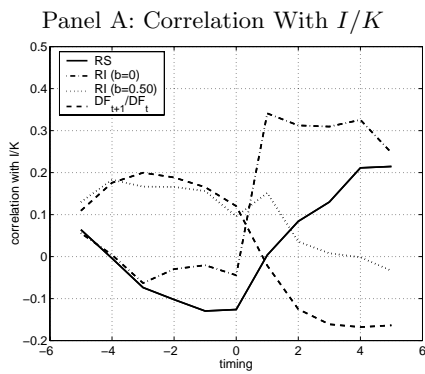
Chapter 3

The Value Premium

> *The point here is simple: although the returns to the B/M strategy are impressive, B/M is not a "clean" variable uniquely associated with economically interpretable characteristics of the firms.*
>
> *Lakonishok, Shleifer, and Vishny (1994)*

3.1 Introduction

Why does aggregate book-to-market predict expected market returns in the time series? Why do value (high book-to-market) stocks earn higher returns than growth (low book-to-market) stocks in the cross-section in spite of the fact that their unconditional risk dispersion is low? Why are value stocks more risky than growth stocks in bad times when risk or risk premium is high? In this paper, I propose a novel economic mechanism underlying the celebrated value premium. The new insight emphasizes the asymmetry of the capital adjustment cost across business cycles, i.e., it is much more difficult for firms to adjust their capital stocks downwards than upwards.[1]

---

[1]Abel and Eberly (1994, 1996) study the optimal investment dynamics of firms in the presence of asymmetry or costly reversibility in a partial equilibrium setting. Ramey and Shapiro (2001) provide direct empirical evidence on the costs of reversing investments using equipment-level data from aerospace plants that closed during the 1990s. A large portion of the literature on capital investment is devoted to examining the implications of a special case of asymmetric adjustment cost, i.e., irreversible investment, which says the cost of adjusting capital downwards via negative investment is infinity. Dixit and Pindyck (1994) survey the literature on irreversible investment and Kogan (2001a, 2001b) examine the implications of irreversibility on time-varying volatility of returns.

Asymmetric adjustment cost has important asset pricing implications. First of all, it gives rise to the asymmetric conditional risk dispersion between value and growth, which is the empirical phenomenon documented in Lettau and Ludvigson (2001) that value stocks are more risky than growth stocks in bad times and less risky in good times, but to a much lesser extent.

To see how this works, note that the asymmetry in adjustment cost says that, in bad times, value firms find it more difficult than growth firms in severing capital stocks, since these firms are typically burdened with more unproductive capital. As a result, the dividend streams and returns of value stocks will fluctuate more with economic downturns. On the other hand, in good times, growth firms face the tougher challenge in expanding their capital stocks to take advantage of favorable economic conditions. This upward adjustment of capital is less urgent for value firms since their previously unproductive capital stocks now become productive. Since expanding capital is relatively easy, the dividend streams and returns of growth firms do not have to covary much with economic booms. The net effect is the asymmetric conditional risk dispersion between value and growth.

Second, asymmetric adjustment cost is also consistent with a low unconditional risk dispersion between value and growth. The key is that bad times or economic recessions characterized by negative investment happen less often than good times or economic booms. In addition, when bad times do happen they tend to last for relatively short periods of time. The upshot is that high positive risk dispersion between value and growth in bad times is offset by a series of low negative risk dispersion in good times, leading to a low unconditional risk dispersion.

Finally, combined with a countercyclical market price of risk, the asymmetric conditional risk dispersion between value and growth implies a high unconditional value premium.

To evaluate the quantitative asset pricing implications of asymmetry in adjusting capital, I construct an industry equilibrium model in the tradition of Hopenhayn (1992a)

augmented with capital accumulation and aggregate uncertainty. I also explicitly model the endogenous entry and exit decisions of firms in order to quantify the impact of survival bias, which has been proposed in the literature as one of the driving forces of the value premium.

My main findings can be summarized as follows. First, the model is able to generate a value premium close to that observed in the data under reasonable parameterizations. Comparative static analysis shows that, given a countercyclical market price of risk, the asymmetric adjustment cost is indeed the main driving force of the value premium. Second, based on the model with endogenous entry and exit, the impact of survival bias is shown to be quantitatively negligible.

Apart from that the economic mechanism underlying the value premium is novel, the model constructed below has a distinctive feature that separates this paper from early literature. Specifically, the cross-sectional distribution of firms is modeled endogenously and it in turn feeds back to the aggregate quantities of the economy. This feature is essential since the firms make dynamic investment decisions and the cross-sectional distribution provides useful information for the firms to predict future evolution of output price.

More from the modeling perspective, most, if not all, of the extant industry equilibrium models abstract from aggregate uncertainty.[2] However, aggregate uncertainty is indispensable for asset pricing purposes. An important contribution of this paper is thus to extend the traditional industry equilibrium framework to incorporate aggregate uncertainty.

The itinerary for the rest of the trip is as follows. Section 3.2 connects this paper to the recent theoretical literature which aims at understanding the cross-sectional variations of expected returns through explicit, structural modeling. The industry equilibrium model is constructed in Section 3.3 and its quantitative implications are explored in Section

---

[2]Prominent examples include Hopenhayn (1992a, 1992b), Cooley and Quadrini (2000), and Gomes (2001).

3.4. Section 3.5 offers some concluding remarks. Appendix 3.5 provides the empirical motivation of this paper by reviewing briefly the recent literature on both the time series and cross-sectional predictability of returns associated with book-to-market.

3.2 Related Literature

Berk (1995) shows that if expected returns rise then prices must be driven down, since future dividends are discounted at a higher rate. The financial ratios such as book-to-market and earnings-price ratios can forecast expected returns, since price appears as the denominator of these ratios. However, with this argument alone, book-to-market does not seem to have a separate role in explaining returns, besides what is implied by its inverse relation with market capitalization. Since book-to-market is a more powerful predictor of returns than size in the data, there must exist some other channel through which book-to-market affects returns independently.

In a risk-based paradigm, the value premium can be considered as a product of two terms, risk dispersion between value and growth and the market risk premium. A high unconditional value premium can coexist with a low unconditional risk dispersion; the reason is that market risk premium is countercyclical (due to time-varying risk and/or market price of risk) and value stocks are more risky than growth stocks in bad times. The conditional nature of the value premium is highlighted in Campbell and Cochrane (2000) and Lettau and Ludvigson (2001).

The economic source underlying the countercyclical market price of risk is by now well understood. A prominent example is the time-varying risk aversion in an external habit model of Campbell and Cochrane (1999). In contrast, the economic understanding of the asymmetric risk dispersion between value and growth across business cycles is still in its infancy.

Gomes, Kogan, and Zhang (2001) link expected returns to size and book-to-market in a dynamic general equilibrium production economy. Size and book-to-market can pre-

dict returns because they are correlated with the firm's systematic risk. Specifically, size in their model determines the risk of firms associated with their future growth opportunities and book-to-market serves as a proxy for the risk of existing projects. Berk, Green, and Naik (1999) construct a partial equilibrium model to explain some of the cross-sectional variations of returns based on similar idea of time-varying risks. However, the value effect generated through the risk of assets-in-place is relatively weak.

A popular interpretation of the value factor is that it is the proxy for a state variable associated with relative financial distress. Since value stocks are typically in distress, if a credit or liquidity crunch comes along, these stocks will do very badly.[3] Taking the idea of financial distress seriously, Gomes, Yaron, and Zhang (2001) incorporate financing constraints into a partial equilibrium model in the tradition of Cochrane (1991, 1996), and investigate the quantitative implications of these constraints on the cross-section of returns. However, these authors find that financial constraints actually worsen the performance of an investment based factor pricing model in pricing a cross-section of returns.

Partly because of the difficulty in understanding the value premium within a risk-based paradigm, a very popular line of explanation attributes it to market mispricing due to investors' overreaction to firm performance or overconfidence in their own abilities. Prominent examples include DeBondt and Thaler (1987), Lakonishok, Shleifer, and Vishny (1994), Daniel and Titman (1997), Barberis, Shleifer, and Vishny (1998), and Daniel, Hirshleifer, and Subrahmanyam (1998, 2001). Barberis and Huang (2001) study the firm-level stock returns by modifying investor's preferences to reflect loss aversion and mental accounting. Another line of explanation, advocated by Lo and MacKinlay (1990) and MacKinlay (1995), argue that the size and book-to-market anomalies may be due to statistical problems such as data-snooping bias.

In summary, the current economic understanding of the asymmetric conditional risk dispersion between value and growth across business cycles is surprisingly fragmentary.

---

[3]See, among others, Chan and Chen (1991), Fama and French (1993, 1996), and Cochrane (2001).

The central goal of this paper is thus to identify and articulate the potential economic mechanism driving the asymmetric conditional risk dispersion between value and growth. Moreover, I will also evaluate the quantitative importance of the identified mechanism in explaining the empirically observed value premium.

3.3 The Model

I construct an industry equilibrium model in the tradition of Hopenhayn (1992a) augmented with capital accumulation and aggregate uncertainty. I adopt the industry framework since the value premium is effectively an *intra*-industry phenomenon, as documented in Cohen, Polk, and Vuolteenaho (2000).

3.3.1 The Environment

The industry is composed by a continuum of competitive firms which produce a homogeneous product. Firms behave competitively, taking the price in the goods market as given.

Technology

Production requires one input, capital, k, and is subject to both an aggregate shock and an idiosyncratic shock. Specifically, the production function is given by:

$$y_{jt} = e^{x_t + z_{jt}} k_{jt}^\alpha \qquad 0 < \alpha < 1 \tag{1}$$

where y_{jt} and k_{jt} are the output and capital stock of firm j at period t, respectively. The next two assumptions concern the nature of productivity shocks:

Assumption 1 The aggregate productivity shock has a stationary and monotone Markov transition function, denoted $Q_x(x_{t+1}|x_t)$, as follows: $x_{t+1} = \overline{x}(1 - \rho_x) + \rho_x x_t + \sigma_x \varepsilon_{t+1}^x$ where ε_{t+1}^x is an IID standard normal shock.

Assumption 2 (i) The idiosyncratic productivity shocks, denoted z_{jt}, are uncorrelated across firms, indexed by j, and have a common stationary and monotone Markov transition function, denoted $Q_z(z_{jt+1}|z_{jt})$, as follows: $z_{jt+1} = \rho_z z_{jt} + \sigma_z \varepsilon_{jt+1}^z$ where ε_{jt+1}^z is IID standard normal shock and ε_{jt+1}^z and ε_{it+1}^z are uncorrelated of each other for any pair (i, j) with $i \neq j$. Moreover, ε_{t+1}^x is independent of ε_{jt+1}^z for all j. (ii) The initial level of idiosyncratic productivity parameters of new entrants are identically zero, the long run mean of the invariant distribution induced by the transition function Q_z.

Both aggregate shock and idiosyncratic shock are necessary to generate a nontrivial cross-section of returns. Aggregate uncertainty is clearly needed; otherwise all firms in the economy would earn exactly the risk-free rate. In addition, idiosyncratic shocks are necessary; otherwise all firms would not only have the same decision rules but also make the same choices.

Pricing Kernel

I follow Berk, Green, and Naik (1999) and parameterize directly the pricing kernel without modeling explicitly the consumer's problem.[4] In particular, I specify:

$$\log M_{t+1} = \log \beta + \gamma (x_t - x_{t+1}) \tag{2}$$

$$\gamma = \gamma_0 + \gamma_1(x_t - \overline{x}) \qquad \gamma_1 < 0 \tag{3}$$

where M_{t+1} denotes the stochastic discount factor from time t to $t+1$ and β and γ_0 are positive parameters. It follows that the real interest rate across the same time interval is given by $r_{t+1} = 1/\mathrm{E}_t[M_{t+1}]$, which is known at the beginning of period t.

(2) can be motivated as follows. Suppose there is a consumer side of the economy featuring one representative agent with isoelastic utility and the relative risk aversion coefficient $\widetilde{\gamma}$. It follows that the log pricing kernel can be written as $\log M_{t+1} = \log \beta + \widetilde{\gamma}(c_t - c_{t+1})$ where c_t denotes log aggregate consumption. Since I do not solve the

---

[4]Parameterizing the pricing kernel has been popular in recent asset pricing literature. Besides Berk, Green, and Naik (1999) cited above, see also Bansal and Vishwanathan (1993) and Chapman (1997).

consumer's problem which would be necessary in a general equilibrium model, I can link c_t to aggregate state variables in a simplistic, reduced-form way. This is done by letting $c_t \approx a + bx_t$ with $b > 0$.[5,6] Then (2) follows immediately by defining γ to be $\widetilde{\gamma} b$.

It is well-known that isoelastic utility has many limitations, one of which is that it implies a constant market price of risk in the current setting.[7] To generate a counter-cyclical market Sharpe ratio, I thus assume that the slope coefficient of log pricing kernel, γ, to be decreasing in the demeaned aggregate productivity $x_t - \overline{x}$. The economic insight captured by (3) is the time-varying risk aversion implied by the external habit model of Campbell and Cochrane (1999).

To see that the pricing kernel (2) implies a countercyclical market price of risk, note by Hansen and Jagannathan (1991) bound, the market price of risk can be written as:

$$\frac{\sigma_t(M_{t+1})}{\mathrm{E}_t(M_{t+1})} = \frac{\sqrt{e^{[\gamma_0 + \gamma_1(x_t - \overline{x})]^2 \sigma_x^2} \left[e^{[\gamma_0 + \gamma_1(x_t - \overline{x})]^2 \sigma_x^2} - 1 \right]}}{e^{\frac{1}{2}[\gamma_0 + \gamma_1(x_t - \overline{x})]^2 \sigma_x^2}} \tag{4}$$

where $\mathrm{E}_t(M_{t+1})$ and $\sigma_t(M_{t+1})$ denote the conditional expectation and volatility of the pricing kernel, respectively. (4) follows by combining (2) with the specification of x_t process in Assumption 1 and using the properties of log-normal distribution. Figure 3.2 plots the market price of risk against aggregate shock x_t using (4) under benchmark parameterization. It is obvious that the market price of risk is countercyclical. As will be seen later, a countercyclical market price of risk is necessary for generating a high value premium while maintaining a low unconditional risk dispersion between value and growth.

---

[5]Since there exist a large number of firms, the law of large number implies that firm-specific shocks do not affect the aggregate consumption. Moreover, the stationarity of the economy implies that the level of aggregate capital stock affects consumption only indirectly through aggregate shock, given the initial level of aggregate capital.

[6]I have also tried quadratic functional form for log consumption c_t in terms of aggregate productivity x_t and obtained quantitatively similar results.

[7]The market price of risk can be written as the product of risk aversion coefficient and volatility of consumption growth. In a general equilibrium production economy such as that in Gomes, Kogan, and Zhang (2001), the volatility of consumption growth is endogenously determined and countercyclical. In contrast, in the current setting, consumption growth is exogenously specified and homoscedastic.

Industry Demand

The inverse industry demand function is denoted by $P(Y_t)$, where P_t is the output price and Y_t is the total output in the economy at time t. I parameterize $P(\cdot)$ to be the following isoelastic form:

$$P(Y_t) = Y_t^{-\eta} \tag{5}$$

where $0 < \eta < 1$ and $1/\eta$ can be interpreted as the price elasticity of demand.

3.3.2 The Firms

Timing of Events

I solve two versions of the model, with and without entry and exit. The timing of events in the case without entry and exit is standard. Upon observing shocks at the beginning of period t, the firms make optimal investment decisions. The timing of events in the model with endogenous entry and exit is only slightly more involved. From time t to time $t+1$, the action sequence of the firms is as follows:

- At the beginning of time t, the firms that have made their decisions *last* period to exit the economy at the beginning of *current* period leave the economy.

- Aggregate shock, x_t, is revealed to the incumbent firms as well as potential entrants.

- Potential entrants make their decisions whether or not to enter the economy. Incumbent firms and new entrants form the current pool of active firms.

- Idiosyncratic shocks are revealed to incumbent firms. New entrants have initial capital stocks of $\underline{k}$ and idiosyncratic shocks of zero. The cross-sectional distribution of firms at time t, denoted μ_t, is measured at this instant.

- Active firms make their optimal investment decision and exit decision on whether or not to exit at the beginning of *next* period simultaneously after observing current period shocks.

Active firms

I now summarize the decisions of active firms. The profit function for an individual firm that has capital stock k_t and idiosyncratic productivity z_t facing aggregate shock x_t and log output price $p_t \equiv \log P_t$ is:[8]

$$\pi(k_t, z_t; x_t, p_t) = e^{x_t + z_t + p_t} k_t^\alpha \tag{6}$$

Let $v(k_t, z_t; x_t, p_t)$ denote the market value of the firm. In the case without entry and exit, I can use Bellman's Principle of Optimality to state the firm's dynamic problem as:

$$v(k_t, z_t; x_t, p_t) = \max_{k_{t+1}, i_t} \{\pi(k_t, z_t; x_t, p_t) - i_t - c(i_t, k_t) - f +$$
$$\iint M_{t+1} v(k_{t+1}, z_{t+1}; x_{t+1}, p_{t+1}) \, Q_z\left(dz_{t+1}|z_t\right) Q_x\left(dx_{t+1}|x_t\right)\} \tag{7}$$

subject to the capital accumulation rule:

$$k_{t+1} = i_t + (1 - \delta)k_t \tag{8}$$

The first four terms in the right-hand side of (7) reflect the current dividends, i.e., profit minus investment expenditure i, adjustment cost c, and the fixed cost of production $f \geq 0$. The fixed cost of production f must be paid every period by all the firms in production. A positive fixed cost is equivalent to the existence of fixed outside opportunity costs for some scarce resources, e.g., managerial labor, used by the firms.

The adjustment cost is assumed to be asymmetric and quadratic:

$$c(i_t, k_t) = \frac{g_t}{2} \left(\frac{i_t}{k_t}\right)^2 k_t \tag{9}$$

where

$$g_t = g^+ \times \mathbf{1}_{\{i_t \geq 0\}} + g^- \times \mathbf{1}_{\{i_t < 0\}}$$

---

[8]I suppress the firm index j for notational simplicity.

and where $\mathbf{1}_{\{\cdot\}}$ is the indicator function that equals one if the event described in $\{\cdot\}$ is true and equals zero otherwise. Moreover, $g^- \geq g^+ > 0$. Figure 1 provides a graphical illustration of the asymmetric specification in (9).

Some special cases quickly come to mind. If $g^- = g^+ > 0$ then (9) reduces to the standard symmetric specification of quadratic adjustment cost. If instead $g^- = \infty$ and $g^+ = 0$, we have the case of investment irreversibility, i.e., the cost of negative investment is infinity. Here I follow the spirit of Abel and Eberly (1994, 1996) to capture the asymmetry of adjusting capital. That is, it is more difficult for firms to adjust their capital stock downwards than upwards.

Exit

In the case with endogenous entry and exit, the Bellman equation for the value function becomes:

$$
\begin{aligned}
v(k_t, z_t; x_t, p_t) \quad = \quad & \max_{k_{t+1}, i_t} \left\{ \pi(k_t, z_t; x_t, p_t) - i_t - c(i_t, k_t) - f + \max \left[r_{t+1}^{-1} (1-\xi) k_{t+1}, \right. \right. \\
& \left. \left. \iint M_{t+1} v(k_{t+1}, z_{t+1}; x_{t+1}, p_{t+1}) Q_z (dz_{t+1}|z_t) Q_x (dx_{t+1}|x_t) \right] \right\} (10)
\end{aligned}
$$

The last term in the right-hand side with the inner max operator is the expected continuation value that allows for exit decision. Firms exit when their expected continuation values are below the values of corresponding outside options, which equal the discounted (at risk-free rate r_{t+1}) values of assets-in-place k_{t+1} net of a bankruptcy cost:

$$
\iint M_{t+1} v(k_{t+1}, z_{t+1}; x_{t+1}, p_{t+1}) Q_z (dz_{t+1}|z_t) Q_x (dx_{t+1}|x_t) < r_{t+1}^{-1} (1-\xi) k_{t+1} \qquad (11)
$$

The bankruptcy cost is assumed to be proportional to the discounted value of assets-in-place with the factor being $\xi \in (0, 1)$. I introduce the proportional bankruptcy cost to capture the idea that it is more costly for a large firm, measured by its capital stock or book value, to exit the economy than a small firm, so that firms with small capital stocks

are more likely to exit the economy.

Note that, although exit decision is completely determined by the current period state variables $(k_t, z_t; x_t, p_t)$, exit actually takes place at the beginning of *next* period before the shocks are observed.

It is clear from (11) that the mass of exiting firms is countercyclical. The reason is that the expected continuation value rises as aggregate productivity x_t goes up, so the exit criterion (11) is less likely to hold.

Proposition 1 There exists a unique value function $v(k, z; x, p)$ that satisfies (10). Moreover, the value function is continuous and increasing in k, z, x, and p.

Entry

In every period there is also a continuum of potential entrants who decide whether or not to enter the industry. Entrants make their decisions at the beginning of the period *after* aggregate shock is revealed.

New firms will enter the market until expected discounted profits net of the entry cost is zero. Formally,

$$v(\underline{k}, 0; x_t, p_t) - \underline{k} \leq \kappa \qquad (12)$$

where $\underline{k} > 0$ is the minimum amount of capital required for any firm to operate,[9] and where $\kappa \geq 0$ denotes the entry cost apart from what must be spent on the initial capital stock. The free-entry condition (12) holds with equality if and only if entry is positive.[10]

The exact level of entry at the beginning of period t, denoted B_t, will be determined by the market clearing condition (5). For now it suffices to note that the model generates procyclical entry mass. To see this, first notice that when output price p_t goes up, the

---

[9]This restriction can be justified by institutional considerations. There is also a technical reason for restricting $\underline{k} > 0$, i.e., when $k = 0$, the adjustment cost function (9) is undefined.

[10]To see this, suppose (12) holds with strict inequality, then entry mass must be zero since entry profits are less than entry costs. Now suppose (12) holds with entry profits strictly greater than entry costs, then entrants will keep entering the industry, driving up total supply, dampening output price; until the net profits of doing so equal zero, or (12) holds with equality in equilibrium.

firm value v rises since the current period profits increase. Now suppose productivity x_t increases in good times, the value function v goes up as expected; to keep the left-hand side of the free-entry condition (12) pegged at a constant entry cost κ, I need to move down p_t somewhat to keep v in check. By the market clearing condition (5), lower price implies higher output, which in turn leads to more entry mass.[11] In sum, higher aggregate productivity x_t leads to more entry mass. The exact opposite is true in bad times.

3.3.3 Aggregation and Heterogeneity

Having described the optimization behavior of the firms, I am now ready to characterize the aggregate behavior of the industry.

The output price will be determined competitively in equilibrium to equate industry demand and supply in the goods market. It is obvious that industry output, and hence price, depends upon the firm distribution.

Let μ_t denote the measure over the capital stocks and idiosyncratic shocks of the firms at time t. Let $s(k_t, z_t; x_t, p_t)$, $i(k_t, z_t; x_t, p_t)$, and $y(k_t, z_t; x_t)$ denote, respectively, the optimal exit decision (that takes value zero if exit and one if stay), investment decision, and output for the firm with capital k_t and idiosyncratic productivity z_t facing log output price p_t and aggregate productivity x_t. Define Θ to be any measurable set in the product space of k and z and let $H(\mu_t, x_t, x_{t+1})$ be the law of motion for the firm distribution, then $H(\cdot, \cdot, \cdot)$ can be stated formally as:

$$
\begin{aligned}
\mu_{t+1}\left(\Theta ; x_{t+1}\right) &= H\left(\mu_t, x_t, x_{t+1}\right) \\
&\equiv \begin{cases} T\left(\Theta,\left(k_t, z_t\right) ; x_t\right) \mu\left(k_t, z_t ; x_t\right) & \text{for } k_t > \underline{k} \\ T\left(\Theta,\left(\underline{k}, z_t\right) ; x_t\right) \mu\left(\underline{k}, z_t ; x_t\right) + B_t \times \mathbf{1}_{\{(\underline{k}, 0) \in \Theta\}} & \text{for } k_t = \underline{k} \end{cases}
\end{aligned} \tag{13}
$$

---

[11]Although x_t moving up results in higher output, the fact that price elasticity of demand is larger than unity says that output has to increase more than what would be achieved by higher productivity. This extra increment of output is provided by more entry mass.

where

$$T\left(\Theta,(k_t,z_t)\,;x_t\right) = \iint \mathbf{1}_{\{(i+(1-\delta)k_t,z_{t+1})\in\Theta\}} s(k_t,z_t;x_t,p_t)\, Q_z(dz_{t+1}|z_t)Q_x(dx_{t+1}|x_t) \quad (14)$$

Although the exact condition is somewhat technical, the underlying intuition is quite straightforward. (13) says that next period distribution is determined from combining the distribution of surviving firms with that of the new entrants, which have minimum capital and long run average idiosyncratic productivity at the time of entry. The law of motion for the individual states for the surviving firms is obtained by combining their optimal decision rules concerning exit and capital accumulation, as formalized in (14).

With the definition of the firm distribution, I can now write the total industry output as:

$$Y_t \equiv \int y(k_t,z_t;x_t)\,\mu_t\,(dk,dz) \quad (15)$$

The entry level B_t can then be determined by combining the market clearing condition (5) and the definition of aggregate output (15).[12]

3.3.4 Recursive Competitive Equilibrium

With the model complete I am now ready to state the conditions required to characterize the stationary competitive equilibrium for this economy.

Definition 1 A recursive competitive equilibrium is characterized by: (i) a market log output price p_t^*; (ii) a set of decision rules, exit $s^*(k_t,z_t;x_t,p_t^*)$ and investment $i^*(k_t,z_t;x_t,p_t^*)$, as well as a value function $v^*(k_t,z_t;x_t,p_t^*)$ for each firm; and (iii) a law of motion of firm distribution H^*, such that:

- optimality: $s^*(k_t,z_t;x_t,p_t^*), i^*(k_t,z_t;x_t,p_t^*)$, and $v^*(k_t,z_t;x_t,p_t^*)$ solve (10) subject to (8) and (9) for each firm;

- the free-entry condition (12) holds;

---

[12]See Appendix 3.5 for the detail description of this procedure.

- consistency: the aggregate output Y_t is consistent with production of all firms in the industry, i.e., (15) holds. The law of motion of firm distribution H is consistent with the firms's optimal decisions, i.e., (13) and (14) hold.

- goods market clearing:

$$e^{p_t^*} = Y_t^{-\eta} \tag{16}$$

Proposition 2 A unique stationary equilibrium with positive entry and exit exists.

3.3.5 Computational Strategy

Since the analytical solution to the model is impossible to obtain, I develop a numerical algorithm that is capable of approximating the competitive equilibrium up to any arbitrarily small error.

The primary obstacle to solve the model stems from the endogeneity of the log output price p_t, an aggregate state variable, which depends upon the cross-sectional distribution of firms, an infinite-dimensional object. To deal with this difficulty, I follow the "approximate aggregation" idea of Krusell and Smith (1998).[13] Specifically, I assume that the firms are imperfect in their perceptions of how price evolves over time. I then progressively increase the sophistication of these perceptions until the errors that the firms make become negligible.

The solution algorithm amounts to the following iterative procedure. (i) Guess a parametric law of motion of log output price and its initial parameter values. Specifically, I assume that log output price follows a log-linear functional form:

$$p_{t+1} = a_1 + a_2 \times p_t + a_3 \times (x_t - \overline{x}) \tag{17}$$

(ii) Solve the individual firm's problem (10). I compute an approximation to the value function on a grid of points in the discrete state space. Piecewise linear interpolation

---

[13]This idea has proved powerful in analyzing economies featuring heterogeneous consumers and aggregate shock. See, for example, Storesletten, Telmer, and Yaron (2001).

technique is used to compute the value function at points not on the grid. (iii) Use the optimal decision rule to simulate the behavior of N firms (with N a large enough number) over a large number, T, time periods. In particular, I include 5,000 firms and 12,000 periods of monthly frequency and discard the first 2,000 periods of data. Typically, the initial firm distribution is one in which all firms hold the same level of capital stock and idiosyncratic shocks are drawn independently from the unconditional, normal distribution of z process with mean zero and volatility $\sigma_z/\sqrt{1-\rho_z}$. The final results are not sensitive to changes in the initial firm distribution. (iv) Use the stationary region of the simulated data to estimate the coefficients in (17). (v) If the estimation gives the coefficients that are very close to those obtained from last iteration and the goodness of fit is satisfactory then stop. Otherwise update these coefficients and go back to step (ii). If the coefficients have converged but the goodness of fit is not satisfactory enough, then try a more sophisticated specification to replace (17). Appendix 3.5 contains a detail description of the solution algorithm.

3.4 Findings

This section explores the quantitative implications of the model. In Section 3.4.1, I calibrate as many parameters as possible by using extant empirical studies and matching some of the basic moments in the model to the US data. Section 3.4.2 examines the quality of approximate aggregation. I then explore the model's empirical implications both on the time series in Section 3.4.3 and on the cross-section in Section 3.4.4. Section 3.4.5 investigates in detail the economic sources driving the value premium generated in the model.

3.4.1 Calibration

To be consistent with the empirical literature on the cross-section of returns, I calibrate the model at monthly frequency.

Stochastic Discount Factor

I use the data on asset returns to calibrate the parameters in pricing kernel (2). Specifically, I set $\gamma_0 = 53$ and $\beta = 0.995$. These parameter values imply an annual average market Sharpe ratio of 0.43, average equity premium of 6.70% with volatility 15.58%, and an average risk free rate of 1.5% with volatility 2.9%. These moments are close to those reported by Campbell and Cochrane (1999) using the postwar sample. The parameter γ_1 determines the time-variation of conditional market Sharpe ratio. I set $\gamma_1 = -1250$ such that market price of risk at business cycle peak is, on average, 0.9 less than that at trough.[14]

Technology

The capital share α is set to be 30%, similar to that in Kydland and Prescott (1982). The monthly rate of depreciation, δ, is set to be 0.01, which implies an annual rate of 12%, the empirical estimate obtained by Cooper and Haltiwanger (2000).

The persistence of aggregate productivity process, ρ_x, is calibrated to be 0.983 and its conditional volatility, σ_x, 0.0023. These monthly numbers correspond, respectively, to 0.95 and 0.007 in the quarterly frequency, which are those used by Cooley and Prescott (1995). In addition, the long run average level of the aggregate productivity, $\overline{x}$, determines the long run average level of capital stock. I calibrate $\overline{x}$ such that the long run average capital stock is normalized to be around one.

For idiosyncratic productivity process z_t, its firm-level persistence, ρ_z, affects directly that at the portfolio level. I experiment with different values of ρ_z and settle with 0.975 such that the time window, during which growth firms are on average more profitable than value firms, is around five years before and after portfolio formation, similar to that documented by Fama and French (1995) (See Section 3.4.5 below for details). Given ρ_z, the conditional volatility, σ_z, determines the cross-sectional dispersion of firm-

---

[14]This level of Sharpe ratio dispersion is the highest that the specification of pricing kernel (2)–(3) is able to generate. Note that this is lower than the annual dispersion of 1.38 estimated by Whitelaw (1997).

level variables. Noting the important role of investment and disinvestment in generating the value premium in the model, I calibrate σ_z to be 0.05 such that the average rates of investment and disinvestment are 0.147 and 0.021, respectively, similar to those reported by Abel and Eberly (2001).

The fixed cost of production f is calibrated as 0.04 to match the average aggregate book-to-market, which is 0.56 in the model and is close to the average book-to-market of 0.67 reported by Pontiff and Schall (1998).

Asymmetric Adjustment Cost

The key parameter that determines the amount of value premium in the model is the degree of asymmetry in capital adjustment, measured by g^-/g^+. Ramey and Shapiro (2001) find that the estimated average market value of equipment is 28 cents per dollar of replacement cost using a sample of equipment-level data from aerospace plants that closed during the 1990s. I thus calibrate the ratio of g^-/g^+ to be four.[15]

The absolute levels of g^+ and g^- determine the share of adjustment cost in total investment expenditure. I experiment with different values and choose g^+ and g^- to be 10 and 40, respectively, so that the implied share of adjustment cost is around 8%.

Table 3.1 summarizes all the benchmark parameter values used in solving and simulating the model. Table 3.2 reports the set of key moments generated using these parameters. Moreover, the investment and dividend shares in total industry revenue are 0.19 and 0.08, both of which are fairly reasonable.

---

[15]This level of asymmetry is likely to be a conservative one. The reason is that, although capital resale price being much lower than purchase cost (due to specificity and "lemon" problem) is the main source of asymmetry, it is unlikely the only one. For example, investments in new workers may be partly irreversible because of high costs of hiring, training, and firing.

3.4.2 Quality of Approximate Aggregation

With benchmark parameterization and a log-linear functional form for the output price, I obtain the following approximate equilibrium:

$$p_{t+1} = 0.068 + 0.976 \times p_t - 0.134 \times (x_t - \overline{x}) \qquad R^2 = 0.9998 \qquad \widehat{\sigma} = 0.0011$$

There are two measures of fit: R^2 and the standard deviation of the regression error, $\widehat{\sigma}$. In terms of these two measures, the goodness-of-fit is extremely good. Thus, an individual firm perceiving the approximate law of motion for output price makes extremely small mistakes compared to using the precise law of motion. In this sense, the firms are extremely close to optimal behavior, which is precisely what competitive equilibrium dictates.

3.4.3 Time Series

Kothari and Shanken (1997) and Pontiff and Schall (1998) find that the beginning-of-period book-to-market is a significant predictor of the end-of-period market return in an univariate regression. Table 3.3 reports the results of the same regression performed on the simulated sample. These statistics are obtained by averaging regression results from 100 samples of 840 monthly periods, which is comparable to the sample size used by Pontiff and Schall (1998). Book value in the model is defined as the capital stock since the price of capital is normalized to be one. Table 3.3 shows that the aggregate book-to-market is a significant predictor of one-period-ahead return both in monthly and annual frequency. Moreover, both the slopes and adjusted R^2s are comparable to those found empirically.

That high book-to-market predicts high expected returns in the time series is probably not surprising. The reasons are as follows: (i) marginal q and future discount rates are negatively correlated; (ii) marginal q and average q are positively correlated;[16] and

---

[16]Hayashi (1982) shows that if the firm is a price-taker with constant returns to scale in both production and adjustment cost technology, then marginal q equals average q. Moreover, Abel and Blanchard (1986) construct series of marginal q under various assumptions on demand and technology and find that

(iii) suppose for simplicity that the price of capital goods is constant; then book value of capital is also the replacement cost, meaning that average q is just the inverse of book-to-market. To sum up, high book-to-market implies low average q, low marginal q, and hence high future returns.

3.4.4 Cross-Section

Table 3.4 and 3.5 report average returns for portfolios constructed by one-dimensional sort of stocks on size and book-to-market, respectively.[17] Panel B of Table 3.4 shows that, when portfolios are formed on market capitalization, they exhibit a negative relation between size and average returns, similar to that observed in Fama and French (1992). Next, Table 3.5 presents average returns for portfolios formed based on ranked values of book-to-market ratios. Again consistent with the data, Panel B indicates a positive relation between book-to-market ratios and average returns.[18]

Table 3.6 presents the time series properties of market excess return, HML, and SMB. These portfolio returns are constructed following the same method in Fama and French (1993). Specifically, in June of each year, all firms are sorted independently to two size groups and three book-to-market groups. Big stocks (B) are above the median market equity and small stocks (S) are below. Similarly, low book-to-market stocks (L) are below the 30th percentile of book-to-market, medium book-to-market stocks (M) are in the middle 40 percent, and high book-to-market stocks (H) are in the top 30 percent. Six value-weighted portfolios, LS, MS, HS, LB, MB, and HB, are then formed as the intersections of the size and book-to-market groups. For example, LS is the value-weighted

---

regression results of investment on marginal q are very similar to those on average q.

[17]Following Fama and French (1992), I form portfolios at the end of June each year and the value-weighted returns are calculated for the next 12 months. In each of these sorts, I form 12 portfolios. The middle eight portfolios correspond to the middle eight deciles of the corresponding characteristics, with four extreme portfolios (1A, 1B, 10A, and 10B) splitting the bottom and the top deciles in half. I repeat the entire simulation 100 times and report the average results of the sorting procedure across the simulations.

[18]The difference in average returns between Panel A and Panel B in both Table 3.4 and Table 3.5 is due to the inflation rate, as Fama and French (1992) use nominal returns in the data while I model real returns.

return on the portfolio of stocks that are below medium in size and in the bottom 30 percent of book-to-market. HML is the difference between the returns on a portfolio of high book-to-market stocks and a portfolio of low book-to-market stocks, constructed to be neutral with respect to size: $\text{HML} \equiv (\text{HS} + \text{HB})/2 - (\text{LS} + \text{LB})/2$. Similarly, SMB is the difference between the returns on a portfolio of small stocks and a portfolio of big stocks, constructed to be neutral with respect to book-to-market: $\text{SMB} \equiv (\text{LS} + \text{MS} + \text{HS})/3 - (\text{LB} + \text{MB} + \text{HB})/3$. The statistics reported in Table 3.6 are obtained by averaging results from 100 artificial samples, each with on average 3,500 firms for 820 months, similar to the sample size used in Davis, Fama, and French (2000).

The second row of Table 3.6 shows that, under benchmark parameterization, the model generates a reliable value premium which is quantitatively similar to that found in the data. Moreover, the correlation between HML and SMB returns in the simulation is only -0.11, close to -0.08 reported by Davis, Fama, and French (2000). This correlation implies that the value premium is quite independent of the size premium in the model.

Another important statistic is the unconditional correlation between HML and market return, which is 0.044 in the simulation. This is consistent with the observation in Fama and French (1992, 1993) that the unconditional risk dispersion between value and growth is essentially flat.

3.4.5 Intuition

Given that the model is capable of generating reasonable amount of the value premium, the interesting question is of course what drives these results. I provide some intuition in this subsection.

Time-Varying Risk

Intuitively, time-varying risk premium can be generated either by time-varying market price of risk or the amount of risk in the economy. In contrast to an endowment economy in which the former channel is usually invoked with time-varying risk aversion, the capital

accumulation channel is able to deliver the time-varying amount of risk endogenously in a production economy.

To see this, consider a standard production economy featuring a representative firm, aggregate uncertainty, and adjustment cost of capital. In contrast to an endowment economy, the presence of capital accumulation enables the firm to mitigate the shock of aggregate uncertainty so as to generate a relatively smooth dividend stream.[19] For example, when productivity goes up, all the additional cash flow will translate into dividend one-by-one in an endowment economy. In the production economy, however, the firm will invest to increase its capital stock since the productivity process is persistent. Thus part of the incremental cash flow will be allocated as investment, and the resulting dividend process will not covary with business cycle as much as it would in the endowment economy. The point is that, all else being equal, the market return in a production economy is less risky than that in an endowment economy.

Capital adjustment cost, by its very nature, is the offsetting force of the above smoothing channel via capital accumulation. In addition, the higher the adjustment cost the firm faces, the more risky its dividend stream is. The endowment economy can be considered as the limiting case of the production economy when the adjustment cost is infinity.

Value Factor in Earnings

Having linked risk to adjustment cost, I now show that firms with high book-to-market signals sustained low earnings on book equity, and those with low book-to-market indicates persistent high earnings. These earning patterns of value and growth firms implied by the model is consistent with what was documented in Fama and French (1995).

Panel A of Figure 3.3 shows the average values of profitability, measured as return on book equity, for 11 years around portfolio formation.[20] Book-to-market equity is as-

---

[19]This mechanism is the main reason why it is more difficult to generate high enough equity premium in a standard real business cycle model than in an endowment economy. See Rouwenhorst (1995), Jermann (1998), Lettau (2000), and Tallarini (2001).

[20]Profitability (or return on book equity) is measured by $[\triangle BE_t + D_t]/BE_{t-1}$, where BE denotes book

sociated with persistent differences in profitability. Growth firms are on average more profitable than value firms for more than five years before and five years after portfolio formation. Moreover, the profitability of growth firms improves prior to portfolio formation and deteriorates thereafter. This pattern appears because growth firms invest more and their book equity grows faster than earnings. The opposite is true for value firms.

Similar to profitability, book-to-market ratio at portfolio level is also quite persistent, as shown in Panel B of Figure 3.3. The distressed value firms in the high book-to-market portfolios tend to have high book-to-market for five years before and at least five years after portfolio formation. In contrast, growth firms have sustained high profitability and they are priced persistently at a premium above their book values.

Asymmetric Adjustment Cost

The relative earnings patterns of value versus growth firms uncovered above have significant impact on these firms' optimal investment behavior across business cycles.

Since value firms are typically burdened with more unproductive capital stock, in bad times they face more challenge than growth firms in adjusting capital stock downwards. Hence value stocks covary more with economic downturns, i.e., value stocks are more risky than growth stocks in bad times.

The exact opposite is true for growth stocks in good times. In good times, growth firms face more challenge of adjusting their capital stocks upwards to take advantage of favorable economic conditions. This challenge of adjusting capital stock upwards is less urgent for value firms, since their previously unproductive capital stocks now turn productive. Thus, growth stocks are more risky than value stocks in good times.

Since on average investment happens more frequently than disinvestment and the magnitude of investment is usually higher than that of disinvestment, it seems that, un-

---

value of equity and D is the dividend payout. Thus profitability equals the ratio of common equity income for the fiscal year ending in calender year t to the book value of equity for year $t-1$. The profitability of a portfolio is defined as the sum of $[\triangle BE_{jt} + D_{jt}]$ for all firm j in a portfolio divided by the sum of BE_{jt-1}, thus it is the return on book equity by merging all firms in the portfolio. For each portfolio formation year t, the ratios of $[\triangle BE_{t+i}+D_{t+i}]/BE_{t+i-1}$ are calculated for year $t+i$ where $i=-5,\ldots,5$. The ratio for year $t+i$ is then averaged across portfolio formation years.

conditionally, growth firms would be more risky than value firms, if the capital adjustment cost were symmetric. This is contradictory to the finding in Fama and French (1992) that the unconditional risk dispersion between value and growth is effectively zero.

Here is precisely where the asymmetric specification of adjustment cost plays its key role. Effectively, it implies that in good times capital stock can be easily adjusted upwards so that the returns of growth firms do not fluctuate much with economic conditions. In contrast, in bad times the hurdles of adjusting capital downwards are more daunting and more relevant to value firms than to growth firms. As a result, value stocks are more risky than growth stocks in bad times, and less risky in good times but to a much lesser extent. The net effect is a flat unconditional risk dispersion between value and growth.

To investigate the patterns of risk dispersion between value and growth stocks, both unconditional and conditional, Table 3.7 reports the β's with respect to the stochastic discount factor M_{t+1} for four portfolios: HS (high and small), HB (high and big), LS (low and small), and LB (low and big). Risk is measured by the slope coefficient from the following regression: $R_{pt+1} = \alpha_p + \beta_p M_{t+1} + e_{t+1}$, where the subscript p denotes specific portfolio. This regression is performed on three different samples: full sample, good times (defined as periods when $x_t \geq \overline{x} + \sigma_x/\sqrt{(1 - \rho_x^2)}$), and bad times (defined as periods when $x_t \leq \overline{x} - 1.5\sigma_x/\sqrt{(1 - \rho_x^2)}$).[21] Table 3.7 reports that, as the model has predicted, value firms are more risky than growth firms in bad times, and less risky in good times but to a much lesser extent. Moreover, the unconditional risk dispersion between value and growth is effectively zero.

Comparative Statics

If the asymmetric conditional risk dispersion is the base for a positive value premium, then the countercyclicality of market price of risk, captured by γ_1, can be understood as the propagation mechanism. To investigate the quantitative importance of these two

---

[21]The asymmetry in the definition of good versus bad times corresponds to the fact that disinvestment only happens when times are really bad. Or equivalently, the long run average rate of investment is δ, the depreciation rate, instead of zero.

terms, I conduct the following two computational experiments.

First, to understand the role of asymmetric adjustment cost, I set $g^- = g^+ = 10$ while keeping all the other benchmark parameterizations. The value premium generated in this case is around 0.55% per annum, much lower than that in the benchmark case. This is expected since the risk dispersion in bad times is much lower due to the symmetry in adjustment cost. Moreover, the unconditional correlation between HML and market portfolio is -0.19. This negative unconditional risk dispersion happens since the long run average investment rate equals the positive depreciation rate, implying that good times happen more often than bad times and growth firms are more risky than value firms unconditionally.

Second, to examine the role of countercyclical market price of risk, I set γ_1 to be zero and then recalibrate the model to see how much asymmetry in capital adjustment cost is needed in order to generate the observed unconditional value premium. The degree of asymmetry in this case turns out to be quite high; g^-/g^+ has to be around 20. Moreover, a side effect shows up in the unconditional correlation between HML and market return, which is 0.17 and too high compared to that reported in Fama and French (1993). This happens because the market price of risk is constant through time and the countercyclicality of market risk premium stems only from time-varying risk. The consequence is that the asymmetry in conditional risk dispersion between value and growth has to be so high that it leads an implausibly wide unconditional risk dispersion as well.

3.4.6 Quantifying the Survival Bias

Survival bias has been proposed as a potential driving force of the value premium. For example, Kothari, Shanken, and Sloan (1995) argue that the data source for book equity from Compustat contains a disproportionate number of high book-to-market firms that survive distress, so their average returns are overstated. Specifically, firms experiencing unfavorable economic conditions have a high propensity to delay the filing of their financial statements. Some of these firms eventually are delisted from the exchanges because of

failure to comply with disclosure requirements. The accounting information for these firms is thus less likely to obtain and to be included in the Compustat. For those firms that delay filing of financial statement due to distress but subsequently improve their performance, they then file their previously delayed statements and Compustat then incorporates their data.[22]

Within the model with entry and exit, I have the luxury to observe the data for firms that would otherwise fail to satisfy some sample selection criterion typically imposed in empirical studies. Thus, I can quantify the magnitude of survival bias by comparing the value premia generated from the artificial samples with and without selection. The sample without bias includes all the firms that are currently in the economy, as well as those that exited during the year after portfolio formation. In the other sample, I include only those that are currently in the economy and exclude those exited during the year after portfolio formation.

The intuition that the survival bias has a positive impact on value premium within the model appears straightforward. The inner max operator in the Bellman equation (10) indicates that exiting firms typically have higher values of book-to-market. To be precise, (11) requires that exiting firms have ratios of next period capital to expected continuation value higher than a threshold value $r_{t+1}/(1-\xi)$. It follows that the book-to-market ratios of exiting firms at the beginning of time t must also be high because these ratios are highly persistent. In addition, (10) implies that exiting firms typically have lower returns since their expected continuation values are lower than the corresponding values of outside options. In sum, exiting firms are typically those with high book-to-market ratios and low returns. If an empirical researcher excludes some of these firms from analysis by following

---

[22]Another source of the survival bias noted by Kothari, Shanken, and Sloan (1995) is the practice followed by Compustat during its early expansion phase of broadening its coverage in late 1970s. Typically, firms were newly introduced into the file with up to five years of past data. This procedure of back-filling data introduced a bias because the firms that were newly added tended to be those with better performance. In contrast, firms that had been large but declined due to poor performance or no longer surviving at the time would not have been introduced to the file. I do not focus on this specific source of survival bias because Fama and French (1992, 1993) document significant value effect in both the 1963 to 1976 and 1977 to 1990 subperiods, and Davis (1994) offers pre-Compustat evidence as well.

some specific sample selection rule, then the high book-to-market group of the sample will tilt towards higher returns.[23]

To solve the model with entry and exit, I need to calibrate the proportional bankruptcy cost parameter ξ. The fraction of entry and exit in the model is pinned down by this parameter. I choose ξ to be 15% so that the implied annual turnover rate in terms of the number of firms is around 7%.[24] Moreover, ξ being 15% seems consistent with extant empirical evidence on bankruptcy costs.[25]

Table 3.8 reports the results on value premia generated in the samples with and without bias. The effect of survival bias does increase the value premium but only by 0.03% per month. This finding lends support to the empirical result in Chan, Jegadeesh, and Lakonishok (1995) that the survival bias on Compustat may not be a severe problem, as far as the relative performance of value versus growth stocks is concerned. However, the survival bias shrinks the volatility of HML portfolio somewhat thus delivering higher t-statistics.

3.5   Conclusion

High book-to-market predicts high expected returns in the time series and value stocks earn higher expected returns than growth stocks in the cross-section. The economic mechanism proposed here is that the capital adjustment cost is asymmetric. Since the capital stocks of value firms have been unproductive in the past and will remain so for

---

[23]As for the impact of entry on the value premium, the consensus in the empirical literature seems to be that the effect is, by and large, neutral. I capture this observation in the model as follows. By the free-entry condition (12), all entrants have the same book-to-market ratio $\underline{k}/v(\underline{k}, 0, x_t, p_t) = \underline{k}/(\underline{k} + \kappa)$ at the time of entry. The entry cost κ is calibrated such that this ratio is around the average level of book-to-market in the sample. Thus the impact of entry on the value premium is minimized.

[24]Dunne, Roberts, and Samuelson (1989a, 1989b) document that around 40% of the firms in manufacturing disappear over five year periods and are replaced by new ones. Using international panels, Cable and Schwalbach (1991) show that average annual entry rates in terms of number of firms is about 6.5% and that average exit rates are very similar to entry rates.

[25]Altman (1984) estimates the average bankruptcy costs to be 12.4% of the firm value three years prior to the petition date, and 16.7% at the petition date. Recently, Andrade and Kaplan (1998) estimate direct and indirect financial distress costs to be between 10–20% of firm value. Since exiting firms have higher than average book-to-market ratios in the model, bankruptcy costs around 15% of assets value correspond to 10–20% of firm value, depending on the book-to-market ratios of the specific exiting firms.

certain periods in the future, these firms face more difficulty in downsizing capital in recessions. The upshot is that value stocks covary more with economic downturns and are hence more risky than growth stocks in bad times. An industry equilibrium model indicates that this mechanism, combined with a countercyclical market price of risk, goes a long way in generating a value premium that is quantitatively close to that observed in the data. Moreover, based on the model with endogenous entry and exit, the impact of survival bias is shown to be quantitatively negligible.

To the best of my knowledge, my paper is the first in the finance literature that models the cross-sectional distribution of firms endogenously. This dynamic framework opens the door for equilibrium analysis for more complex issues concerning the firms, such as optimal dividend policy, equilibrium determinants of financial leverage, and the effects of agency costs on investment and stock return, which are traditionally corporate finance questions. The equilibrium analysis of financial contracting with corporate finance applications appears to a promising avenue for future research.

Bibliography

[1] Abel, Andrew B., and Olivier J. Blanchard, 1986, The Present Value of Profits and Cyclical Movements in Investment, *Econometrica*, 54 (2), 249–273.

[2] Abel, Andrew B., and Janice C. Eberly, 2001, Investment and q With Fixed Costs: An Empirical Analysis, Working Paper, The Wharton School at the University of Pennsylvania.

[3] Abel, Andrew B., and Janice C. Eberly, 1996, Optimal Investment with Costly Reversibility, *Review of Economic Studies*, 63, 581–593.

[4] Abel, Andrew B., and Janice C. Eberly, 1994, A Unified Model of Investment Under Uncertainty, *American Economic Review*, 84 (5), 1369–1384.

[5] Altman, Edward I., 1984, A Further Empirical Investigation of the Bankruptcy Cost Question, *Journal of Finance*, 39, 1067–1089.

[6] Andrade, Gregor, and Steven N. Kaplan, 1998, How Costly is Financial (Not Economic) Distress? Evidence from Highly Leveraged Transactions That Became Distressed, *Journal of Finance*, 53, 1443–1494.

[7] Avramov, Doron, 2002, Stock Return Predictability and Model Uncertainty, Forthcoming, *Journal of Financial Economics*.

[8] Bansal Ravi, and S. Viswanathan, 1993, No Arbitrage and Arbitrage Pricing: A New Approach, *Journal of Finance*, 48, 1231–1261.

[9] Barberis, Nicholas, Andrei Shleifer, and Robert Vishny, 1998, A Model of Investor Sentiment, *Journal of Financial Economics*, 49, 307–345.

[10] Barberis, Nicholas, and Ming Huang, 2001, Mental Accounting, Loss Aversion, and Individual Stock Returns, *Journal of Finance*, forthcoming.

[11] Berk, Jonathan B., 1995, A Critique of Size Related Anomalies, *Review of Financial Studies*, 8, 275–286.

[12] Berk, Jonathan B, Richard C. Green and Vasant Naik, 1999, Optimal Investment, Growth Options and Security Returns, *Journal of Finance*, 54, 1153–1607.

[13] Cable, John, and Joachim Schwalbach, 1991, International Comparisons of Entry and Exit, in *Entry and Market Contestability*, Edited by Geroski and Schwalbach, Oxford, Blackwell.

[14] Campbell, John Y., and John H. Cochrane, 1999, By Force of Habit: A Consumption-Based Explanation of Aggregate Stock Market Behavior, *Journal of Political Economy*, 107, 205–251.

[15] Campbell, John Y., and John H. Cochrane, 2000, Explaining the Poor Performance of Consumption-based Asset Pricing Models, *Journal of Finance*, LV (6), 2863–2878.

[16] Campbell, John Y., and Andrew W. Lo, and A. Craig MacKinlay, 1997, *The Econometrics of Financial Markets*, Princeton University Press, Princeton, New Jersey.

[17] Chan, Louis K. C., and Nai-fu Chen, 1991, Structural and Return Characteristics of Small and Large Firms, *Journal of Finance*, 46, 1467–1484.

[18] Chan, Louis K. C., Narasimhan Jegadeesh, Josef Lakonishok, 1995, Evaluating The Performance of Value Versus Glamour Stocks: The Impact of Selection Bias, *Journal of Financial Economics*, 38, 269–296.

[19] Chapman, David A., 1997, Approximating the Asset Pricing Kernel, *Journal of Finance*, 52, 1383–1410.

[20] Cochrane, John H., 1991, Production-Based Asset Pricing and the Link Between Stock Returns and Economic Fluctuations, *Journal of Finance,* XLVI (1), 209–237.

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

[22] Cochrane, John H., 2001, *Asset Pricing*, Princeton University Press, Princeton, New Jersey.

[23] Cohen, Randolph B., Christopher Polk, and Tuomo Vuolteenaho, 2000, The Value Spread, Working Paper, Kellogg Graduate School of Management, Northwestern University.

[24] Cooley, Thomas F., and Vincenzo Quadrini, 2000, Financial Markets and Firm Dynamics, Working Paper, Stern School of Management, New York University.

[25] Cooley, Thomas F., and Edward C. Prescott, 1995, Economic Growth and Business Cycles, In *Frontiers of Business Cycle Research,* Edited by Thomas F. Cooley, Princeton University Press.

[26] Cooper, Russell W., and John C. Haltiwanger, 2000, On the Nature of Capital Adjustment Costs, NBER Working Paper 7925.

[27] Daniel, Kent, David Hirshleifer, and Avanidhar Subrahmanyam, 1998, Investor psychology and security market under- and over-reaction, *Journal of Finance* 53, 1839–1886.

[28] Daniel, Kent, David Hirshleifer, and Avanidhar Subrahmanyam, 2001, Overconfidence, arbitrage, and equilibrium asset pricing, *Journal of Finance* 56, 921–965.

[29] Daniel, Kent, and Sheridan Titman, 1997, Evidence on the Characteristics of Cross Sectional Variation in Stock Returns, *Journal of Finance,* 52, 1–33.

[30] Davis, James L., 1994, The Cross-Section of Realized Stock Returns: The Pre-Compustat Evidence, *Journal of Finance,* 49, 1579–1593.

[31] Davis, James L., Eugene F. Fama, and Kenneth R. French, 2000, Characteristics, Covariances, and Average Returns: 1929—1997, *Journal of Finance* 55 (1), 389–406.

[32] DeBondt, Werner F.M., and Thaler, Richard H., 1987, Further Evidence on Investor Overreaction and Stock Market Seasonality, *Journal of Finance,* 42, 557–581.

[33] Dixit, Avinash K., and Robert S. Pindyck, 1994, *Investment Under Uncertainty*, Princeton University Press, Princeton, New Jersey.

[34] Dunne, T., M. Roberts, and L. Samuelson, 1989a, The Growth and Failure of US Manufacturing Plants, *Quarterly Journal in Economics*, 104, 671–698.

[35] Dunne, T., M. Roberts, and L. Samuelson, 1989b, Plant Turnover and Gross Employment Flows in the US Manufacturing Sector, *Journal of Labor Economics*, 7, 48–71.

[36] Fama, Eugene F., and Kenneth R. French, 1992, The Cross-Section of Expected Stock Returns, *Journal of Finance*, XLVII (2), 427–465.

[37] Fama, Eugene F., and Kenneth R. French, 1993, Common Risk Factors in the Returns on Stocks and Bonds, *Journal of Financial Economics*, 33, 3–56.

[38] Fama, Eugene F., and Kenneth R. French, 1995, Size and Book-to-Market Factors in Earnings and Returns, *Journal of Finance*, 50, 131–155.

[39] Fama, Eugene F., and Kenneth R. French, 1996, Multifactor Explanations of Asset Pricing Anomalies, *Journal of Finance*, 51, 55–84.

[40] Fama, Eugene F., and Kenneth R. French, 1999, Value Versus Growth: The International Evidence, *Journal of Finance*, 53 (6), 1975–1999.

[41] Ferson, Wayne E., and Campbell R. Harvey, 1999, Conditioning Variables And Cross-Section of Stock Returns, *Journal of Finance*, 54, 1325–1360.

[42] Kydland, Finn E., and Edward C. Prescott, 1982, Time to Build and Aggregate Fluctuations, *Econometrica*, 50, 1345–1370.

[43] Gomes, Joao F., 2001, Financing Investment, forthcoming, *American Economic Review*.

[44] Gomes, Joao F., Leonid Kogan, and Lu Zhang, 2001, Equilibrium Cross-Section of Returns, Working Paper, The Wharton School at the University of Pennsylvania.

[45] Gomes, Joao F., Amir Yaron, and Lu Zhang, 2001, Asset Pricing Implications of Firm's Financing Constraints, Working Paper, The Wharton School at the University of Pennsylvania.

[46] Hansen, Lars Peter, and Ravi Jagannathan, 1991, Restrictions on Intertemporal Marginal Rates of Substitution Implied by Asset Returns, *Journal of Political Economy*, 99, 225–262.

[47] Hayashi, Fumio, 1982, Tobin's Marginal q and Average q: A Neoclassical Interpretation, *Econometrica*, 50 (1), 213–224.

[48] Hopenhayn, Hugo A., 1992a, Entry, Exit, and Firm Dynamics In Long Run Equilibrium, *Econometrica*, 60 (5), 1127–1150.

[49] Hopenhayn, Hugo A., 1992b, Exit, Selection, and the Value of Firms, *Journal of Economic Dynamics and Control,* 16, 621–653.

[50] Jagannathan, Ravi, and Zhenyu Wang, 1996, The Conditional CAPM and the Cross-Section of Expected Returns, *Journal of Finance*, 51, 3–54.

[51] Jermann, Urban J., 1998, Asset Pricing in Production Economies, *Journal of Monetary Economics*, 41, 257–275.

[52] Kogan, Leonid, 2000a, An Equilibrium Model of Irreversible Investment, Working Paper, the Wharton School, University of Pennsylvania.

[53] Kogan, Leonid, 2000b, Asset Prices and Irreversible Real Investment, Working Paper, the Wharton School, University of Pennsylvania.

[54] Kothari, S. P. and Jay Shanken, 1997, Book-to-Market, Dividend Yield, and Expected Market Returns: A Time-Series Analysis, *Journal of Financial Economics*, 44, 169–203.

[55] Kothari, S. P., Jay Shanken, and Richard G. Sloan, 1995, Another Look at the Cross-Section of Expected Stock Returns, *Journal of Finance*, 50, 185–224.

[56] Krusell, Per, and Anthony A. Smith, Jr., 1998, Income and Wealth Heterogeneity in the Macroeconomy, *Journal of Political Economy*, 106 (5), 867–896.

[57] Lakonishok, Josef, Andrei Shleifer, and Robert W. Vishny, 1994, Contrarian Investment, Extrapolation, and Risk, *Journal of Finance*, XLIX (5), 1541–1578.

[58] Lettau, Martin, 2000, Inspecting the Mechanism: Closed-Form Solutions for Asset Prices in Real Business Cycle Models, Working Paper, Federal Reserve Bank of New York.

[59] Lettau, Martin, and Sydney Ludvigson, 2001, Resurrecting the (C)CAPM: A Cross-Sectional Test When Risk Premia Are Time-Varying, forthcoming, *Journal of Political Economy*.

[60] Lewellen, Jonathan, 1999, The Time-Series Relations Among Expected Return, Risk, and Book-to-Market, *Journal of Financial Economics*, 54, 5–43.

[61] Liew, Jimmy, and Maria Vassalou, 2000, Can Book-to-Market, Size, and Momentum Be Risk Factors That Predict Economic Growth? *Journal of Financial Economics*, 57, 221–245.

[62] Lo, Andrew W., and A. Craig MacKinlay, 1990, Data-Snooping Biases in Tests of Financial Asset Pricing Models, *Review of Financial Studies*, 3, 431–468.

[63] MacKinlay, A. Craig, 1995, Multifactor Models Do Not Explain Deviations From The CAPM, *Journal of Financial Economics*, 38, 3–28.

[64] McGrattan, Ellen R., 1999, Application of Weighted Residual Methods to Dynamic Economic Models, in *Computational Methods for the Study of Dynamic Economies*, Edited by Ramon Marimon and Andrew Scott, Oxford University Press.

[65] Pontiff, Jeffrey and Lawrence Schall, 1999, Book to Market as a Predictor of Market Returns, *Journal of Financial Economics*, 49, 141–160.

[66] Ramey, Valerie A. and Matthew D. Shapiro, 2001, Displaced Capital: A Study of Aerospace Plant Closings, *Journal of Political Economy*, 109 (5), 958–992.

[67] Rosenberg, Barr, Kenneth Reid, and Ronald Lanstein, 1985, Persuasive Evidence of Market Inefficiency, *Journal of Portfolio Management,* 11, 9–17.

[68] Rouwenhorst, G., 1995, "Asset Pricing Implications of Equilibrium Business Cycle Models," in Frontiers of Business Cycle Research, T. Cooley, Ed., Princeton University Press, Princeton, NJ.

[69] Storesletten, Kjetil, Chris I. Telmer. and Amir Yaron, 2001, Asset Pricing with Idiosyncratic Risk and Overlapping Generations, Working Paper, The Wharton School at the University of Pennsylvania.

[70] Stokey, Nancy L., and Robert E. Lucas, Jr. with Edward Prescott, 1989, Recursive Methods in Economic Dynamics, Harvard University Press.

[71] Tallarini, Thomas D. Jr., 2000, Risk-Sensitive Real Business Cycles, *Journal of Monetary Economics*, 45, 507–532.

[72] Tauchen, George, and Robert Hussey, 1991, Quadrature Based Methods for Obtaining Approximate Solutions to Nonlinear Asset Pricing Models, *Econometrica*, 59, 371–396.

[73] Whitelaw, Robert F., 1997, Time-Varying Sharpe Ratios and Market Timing, Working Paper, Stern School of Business at New York University.

Stylized Facts

In this section, I briefly review the stylized facts on the time series and cross-sectional predictability of returns related to book-to-market. This section serves both as the empirical motivation and the quantitative target of the theoretical exercise conducted in this paper.

Time Series Evidence

Kothari and Shanken (1997) and Pontiff and Schall (1998) investigate the time series relation between book-to-market and expected market returns. Univariate regression of market return on beginning-of-period book-to-market shows that the slope is significantly positive and R^2 is around 1% in monthly frequency and 16% in annual frequency.

Lewellen (1999) investigates the time-series relations among expected returns, risk, and book-to-market at the portfolio level, and finds that book-to-market predicts economically and statistically significant time-variation in expected stock returns. Moreover, after controlling for risk, book-to-market provides no incremental information about expected returns.

Liew and Vassalou (1999) represent a recent attempt to link value and small-firm returns to macroeconomic variables. These authors find that HML and SMB contain information above and beyond that in the market return for forecasting GDP growth.

Bayesian analysis of predictability and model uncertainty is provided by Avramov (2002).

Cross-Section Evidence

That value stocks earn higher returns than growth stocks in the cross-section has been known at least since Rosenberg, Reid, and Lanstein (1985). Recently, Fama and French (1992, 1993, 1996) document covariation in returns related to size and book-to-market beyond the covariation explained by the market return; they further propose a three-factor model that uses the market portfolio and mimicking portfolios for factors related to size and book-to-market to describe returns. Out-of-sample evidence of the value premium is provided by Davis (1994), and Davis, Fama, and French (1999). In addition, Fama and French (1999) offer international evidence using data from the U.S. and 12 major Europe, Australia, and the Far East (EAFE) countries.

To understand the economic forces behind size and book-to-market factors, Fama and French (1995) show that book-to-market is related to persistent properties of earnings. High (low) book-to-market signals sustained low (high) earnings on book equity. Value stocks are less profitable than growth stocks for four years before and at least five years after ranking dates. Moreover, the earnings of growth stocks continue to grow relative to market earnings but less rapidly. The profitability, measured by earnings to book equity ratios, of growth stocks falls after portfolio formation because book equity grows faster than earnings.

Using log consumption-wealth ratio as an instrument, Lettau and Ludvigson (2001) show that a conditional CCAPM performs far better than unconditional specifications and about as well as the Fama and French three-factor model on portfolios sorted by size and book-to-market. These authors argue that value stocks earn higher average returns than growth stocks because they are more highly correlated with consumption growth in bad times when risk premia are high.[26]

Summary

To sum up, the recent empirical work has documented the following set of stylized facts concerning the value premium:

- Book-to-market are positively correlated with expected return and risk at the aggregate level (Kothari and Sloan [1998] and Pontiff and Schall [1999]) and at the portfolio level (Lewellen [1999]).

- There exists reliable value premium, defined as the average HML return, for the sample from July 1929 to June 1997, around 5.50% per annum (Davis, Fama, and French [2000]).

- Value stocks are more risky than growth stocks in bad times, and growth stocks are more risky than value stocks in good times but to a much lesser extent. (Lettau and Ludvigson [2001]).

- Prior to portfolio formation, the earnings of growth (value) stocks grow faster (slower) than book equity, causing profitability to increase (decrease), but the opposite is true in the years after portfolio formation (Fama and French [1995]).

In this paper I seek to understand the economics underlying the above set of stylized facts.

Proofs

I prove Propositions 1 and 2 in this section. Suppressing the time subscripts and using primes to denote next period variables, I can rewrite the value function (10) as

$$v(k, z; x, p) = \pi(k, z; x, p) - f + \max\{\max_i \left[-i - c(i, k) + r^{-1} (1 - \xi) k' \right],$$

$$\max_i [-i - c(i, k) + \iint Mv(k', z'; x', p') Q_z \left(dz' | z \right) Q_x \left(dx' | x \right)]\} \tag{B1}$$

and define the operator $(Tv)(k, z; x, p)$ to be the right-hand size of (B1).

Proposition 10 *The operator T satisfies that $T : C(\mathcal{K} \times \mathcal{Z} \times \mathcal{X}) \to C(\mathcal{K} \times \mathcal{Z} \times \mathcal{X})$, where $C(\mathcal{K} \times \mathcal{Z} \times \mathcal{X})$ denotes the space of bounded and continuous functions on the product space of $\mathcal{K} \times \mathcal{Z} \times \mathcal{X}$.*

---

[26]The role of conditioning information is also illustrated in Jagannathan and Wang (1996) and Ferson and Harvey (1999).

Proof Consider first the function

$$H^1(k, z; x, p) = \max_{-(1-\delta)k \leq i \leq \overline{k}} \{-i - c(i, k) + r^{-1}(1 - \xi)[i + (1 - \delta)k]\} \qquad \text{(B2)}$$

The Theorem of Maximum guarantees that the optimal policy correspondence is well-defined and upper hemi-continuous. Moreover, H^1 is continuous.

Next suppose $v(k, z; x, p) \in C(\mathcal{K} \times \mathcal{Z} \times \mathcal{X})$. Since $Q_x(dx'|x)$ and $Q_z(dz'|z)$ satisfy the Feller property it follows from Lemma 9.5 in Stokey and Lucas (1989) that

$$\iint Mv(k', z'; x', p') \, Q_z \left(dz'|z \right) Q_x \left(dx'|x \right) \in C(\mathcal{K} \times \mathcal{Z} \times \mathcal{X})$$

The Theorem of Maximum now implies that function

$$H^2(k, z; x, p) = \max_{-(1-\delta)k \leq i \leq \overline{k}} \{-i - c(i, k) + \iint Mv(k', z'; x', p') \, Q_z \left(dz'|z \right) Q_x \left(dx'|x \right)\}$$
$$\text{(B3)}$$

is continuous and its optimal policy correspondence is well-defined and upper hemi-continuous.

The proposition now follows by rewriting T as

$$(Tv)(k, z; x, p) = \pi(k, z; x, p) - f + \max\{H^1(k, z; x, p), H^2(k, z; x, p)\}$$

and that all the functions in the right-hand side are continuous. $\blacksquare$

Proposition 11 *The operator T is a contraction in $C(\mathcal{K} \times \mathcal{Z} \times \mathcal{X})$.*

Proof The proof proceeds by verifying the Blackwell's sufficient conditions: monotonicity and discounting. To see monotonicity, let $v^1(k, z; x, p) \geq v^2(k, z; x, p)$ and let

$$H^{2j}(k, z; x, p) = \max_{-(1-\delta)k \leq i \leq \overline{k}} \{-i - c(i, k) + \iint Mv^j(k', z'; x', p') \, Q_z \left(dz'|z \right) Q_x \left(dx'|x \right)\}$$
$$\text{(B4)}$$

(B3) then says that $H^{21}(k, z; x, p) \geq H^{22}(k, z; x, p)$. Therefore,

$$
\begin{aligned}
(Tv^1)(k, z; x, p) &= \pi(k, z; x, p) - f + \max\{H^1(k, z; x, p), H^{21}(k, z; x, p)\} \\
&\geq \pi(k, z; x, p) - f + \max\{H^1(k, z; x, p), H^{22}(k, z; x, p)\} = (Tv^2)(k, z; x, p)
\end{aligned}
$$

To verify discounting, $\forall a \geq 0$, let

$$
\begin{aligned}
H^2(k, z; x, p; a) &\equiv \max_{-(1-\delta)k \leq i \leq \overline{k}} \{-i - c(i, k) + \iint M[v(k', z'; x', p') + a] \, Q_z \left(dz'|z \right) Q_x \left(dx'|x \right)\} \\
&\leq H^2(k, z; x, p) + \iint Ma \, Q_z \left(dz'|z \right) Q_x \left(dx'|x \right) \leq H^2(k, z; x, p) + \widetilde{\beta}a \qquad \text{(B5)}
\end{aligned}
$$

133

for some $\iint M a \, Q_z \, (dz'|z) \, Q_x \, (dx'|x) \leq \widetilde{\beta} \leq 1$. It then follows that

$$
\begin{aligned}
(Tv + a)(k, z; x, p) &= \pi(k, z; x, p) - f + \max\{H^1(k, z; x, p), H^2(k, z; x, p; a)\} \\
&\leq (Tv)(k, z; x, p) + \widetilde{\beta} a
\end{aligned}
$$

∎

Proof of Proposition 1.   The uniqueness and existence of the value function result from the Contraction Mapping Theorem and Lemma 10 and Lemma 11. The continuity and monotonicity of v in k and p follow from Lemma 9.5 and Theorem 3.2 in Stokey and Lucas (1989). The continuity and monotonicity of v in x and z follow from the continuity and monotonicity of π in x and z and the monotonicity of the Markov transition functions Q_x and Q_z. ∎

Proposition 12 *(Hopenhayn [1990, 1992]) Let the aggregate demand be $P(Y)$. An industry equilibrium exists if: (a) $\mathcal{X}$ and $\mathcal{Z}$ are compact metric spaces; (b) $Q_x(x_{t+1}|x_t)$ and $Q_z(z_{t+1}|z_t)$ are continuous transition functions; (c) technology has decreasing returns to scale and the technology set has a closed graph; (d) $P(Y)$ is weakly decreasing in Y and is measurable with respect to the information filtration generated by x and z; (e) $P(Y)$ is uniformly bounded above and β-integrable; (f) for any initial measure μ_0 there exists at least one feasible allocation; and (g) $\lim_{B\to\infty} ||\pi|| = \infty$ where π denotes profits. If in addition (h) the profitable function is separable in the form $\pi(k, z; x, p) = h_1(x, z)h_2(k, p)$ for some functions h_1 and h_2, then the industry equilibrium above is unique and stationary and exhibits positive entry and exit.*

Proof of Proposition 2.   It is straightforward to verify the conditions (a)–(g) in Proposition 12 hold in the model. Thus an industry equilibrium exists. Moreover the profit function given by (6) obviously satisfies condition (h) in Proposition 12. Thus, the industry equilibrium is also unique. ∎

Computation

Solving the Industry Equilibrium

In this section, I provide some details about the algorithm used to solve the industry equilibrium with entry and exit.

1. Initialize the economy, including all the approximating coefficients. The value function iteration starts now.

2. For given coefficients a's from last iteration, solve the individual firm's problem via the standard value function iteration technique. Some details are provided below.

   Armed with optimal decision rules, prepare to simulate the economy over a long time period.

(a) Initialize the (discretized) firm distribution through a dynamic matrix with two columns, capital stocks and idiosyncratic productivity parameters. Initial capital stocks are equal to one and idiosyncratic shocks are drawn from the steady state distribution of the z process. I use 5,000 firms to start with.

(b) Solve for the expected continuation value on the x grid once the firm's optimization problem is solved.

(c) Solve for log output price p on the x grid from the free-entry condition using a bisection procedure. Simulate a long time series (12,000 periods in monthly frequency) of $\{x_t\}$ and find its corresponding price series $\{p_t^*\}$ using a linear interpolation scheme.

(d) Solve for the aggregate series $\{Y_t^*\}$ associated with this log price series implied by the market clearing condition.

Note $\{p_t^*\}$ and $\{Y_t^*\}$ are the respective realizations of log price and output at time t if there exists positive entry in the same period. Now simulate the economy. The detail simulation steps within one time period are as follows:

(a) Delete the firms contained in the exit index (obtained last period) from the dynamic matrix.

(b) Entry:

- Compute the total output Y_t^c of incumbent firms.
- Compare Y_t^c with Y_t^*. If $Y_t^c > Y_t^*$ then the incumbent firms are over-producing already and entry is set to be zero.
- If $Y_t^c < Y_t^*$ then there is room for entry. Pin down entry mass B_t from the definition of aggregate output: $Y_t^* = Y_t^c + B_t e^{x_t} \underline{k}^\alpha$. Augment the firm matrix with new entrants that have capital stock $\underline{k}$ and idiosyncratic shock at zero.

(c) Exit: Locate the indices of firms that will exit the economy at the beginning of next period.

- Use a linear interpolation scheme to obtain the realized value of expected continuation value in the cross-section.
- Obtain the value of outside option as $r_{t+1}^{-1} (1 - \xi) k_{t+1}$.
- Compare expected continuation value with outside option. If the former is less than the latter for a specific firm then this firm enters the set of exiting firms.

(d) For active firms (incumbents and new entrants), update k_{t+1} using optimal decision rule and linear interpolation. Update z_{t+1} by drawing from the law of motion of that exogenous process.

(e) Check the goodness-of-fit of the approximate law of motion for p_t. If the coefficients have converged but the goodness-of-fit is not satisfactory enough, try an alternative specification for the approximate law of motion.

3. Update approximating coefficients.

4. If coefficients a's are sufficiently close to their values from last iteration, stop. Otherwise go back to step 2 for the next iteration.

Solving the Value Function

I provide some details on the value function iteration routine for solving the individual firm's problem below.

Given the firm's endogenous exit decision, the value function (10) cannot be solved with policy function iteration since the policy functions are not continuous. Moreover, since value function may have kinks in the area of the state space where the firm exits, a low-order polynomial cannot be used to approximate the value function either. I therefore rely on the more robust, standard value function iteration on a discrete state space.

I specify a grid with 100 points for the capital stock with an upper bound $\overline{k}$ (big enough to be non-binding at all times). I construct the grid for capital stock recursively following the method in McGrattan (1999), i.e, $k_i = k_{i-1} + c_{k1} \exp(c_{k2}(i-2))$, where $i = 1, \ldots, 100$ is the index of grid points and c_{k1} and c_{k2} are two constants chosen to provide desired number of grid points and $\overline{k}$ given a pre-specified lower bound $\underline{k}$. The advantage of this recursive construction is that more grid points are assigned around $\underline{k}$ where the value function has its most curvature.

The continuous state variables x and z have to be discretized as well. Since both productivity processes are highly persistent in monthly frequency, I use the method described in Rouwenhorst (1995) to discretize these processes, instead of that in Tauchen and Hussey (1991) which does not work well when persistence is higher than 0.90. I use 11 grid points for x process and 15 points for z process. In all cases the results are robust to finer grids.

Table 3.1 : Benchmark Parameterization

This table lists the benchmark parameter values used to solve and simulate the model. Whenever possible, the fourth column (Source) provides the source of empirical estimates for some parameters. The remaining parameters are calibrated to match a set of key moments in the model to the US data. See Section 3.4.1 for details.

Parameters	Benchmark Values	Description	Source
Stochastic Discount Factor			
β	0.995	Intercept	
γ_0	53	Slope (Fixed Component)	
γ_1	-1250	Slope (Time-varying Component)	
Technology			
α	0.30	Capital Share	Kydland and Prescott (1982)
δ	0.01	Depreciation Rate	Cooper and Haltiwanger (2001)
f	0.04	Fixed Cost of Production	
η	0.50	Inverse Price Elasticity of Demand	
Aggregate Shock			
ρ_x	$0.95^{1/3}$	Persistence	Cooley and Prescott (1995)
σ_x	0.007/3	Conditional Volatility	Cooley and Prescott (1995)
Idiosyncratic Shock			
ρ_z	0.975	Persistence	
σ_z	0.05	Conditional Volatility	
Asymmetric Adjustment Cost			
g^-/g^+	4	Degree of Asymmetry	Ramey and Shapiro (2001)
g^+	10	Cost Parameter when $i > 0$	

137

Table 3.2 : Key Moments Under Benchmark Parameterization

This table reports a set of key moments generated under the benchmark parameters in Table 3.1. These moments are obtained by simulating the economy for 300,000 periods at monthly frequency. The first four return moments and the average rates of investment and disinvestment are annualized. The data source for market excess return and real interest rate is Campbell, Lo, and MacKinlay (1997). The data source for aggregate book-to-market is Pontiff and Schall (1999). Finally, the data source for average rates of investment and disinvestment is Abel and Eberly (2001).

Moments	Model	Data
Market Excess Return		
Mean	6.71%	6%
Volatility	15.60%	18%
Real Interest Rate		
Mean	1.62%	1.8%
Volatility	2.86%	3%
Aggregate Book-to-Market		
Mean	0.56	0.67
Volatility	0.20	0.23
Ratios		
Adjustment Cost/Investment Ratio	8.17%	10%
Average Rate of Investment	0.147	0.15
Average Rate of Disinvestment	0.026	0.02

Table 3.3 : Aggregate Book-to-Market As a Predictor of Market Returns

This table reports the time series regressions of value-weighted industry return on aggregate book-to-market:

$$R_{t+1}^{vw} = a + b \times (b/m)_t + \epsilon_{t+1}$$

where R_{t+1}^{vw} denotes value-weighted industry return from time t to time $t+1$ and $(b/m)_t$ is aggregate book-to-market at the beginning of time t. The regression is conducted at both monthly and annual frequencies. The first row is from Table 2 of Pontiff and Schall (1999). The average slopes and the adjusted R^2s are obtained by averaging results across 100 simulations. The slopes are in percentage.

	monthly		annual	
	slope	adjusted R^2	slope	adjusted R^2
Data	3.02	0.010	42.18	0.160
Model	3.64	0.020	45.03	0.195

Table 3.4 : Properties of Portfolios Formed on Size

At the end of June of each year t, 12 portfolios are formed on the basis of ranked values of size. Portfolios 2–9 cover corresponding deciles of the ranking variables. The bottom and top 2 portfolios (1A, 1B, 10A, and 10B) split the bottom and top deciles in half. The breakpoints for the size portfolios are based on ranked values of size. Panel A is from Fama and French (1992) Table II, Panel A. Panel B is constructed from the simulated panel. The average returns are the time-series averages of the monthly equal-weighted portfolio returns, in percent. $\log(V_f)$ and $\log(B_f/V_f)$ are the time-series averages of the monthly average values of these variables in each portfolio. β is the time-series average of the monthly portfolio post-ranking βs.

Portfolios	1A	1B	2	3	4	5	6	7	8	9	10A	10B
					Panel A: Historical Data							
Return	1.64	1.16	1.29	1.24	1.25	1.29	1.17	1.07	1.10	0.95	0.88	0.90
β	1.44	1.44	1.39	1.34	1.33	1.24	1.22	1.16	1.08	1.02	0.95	0.90
$\log(V_f)$	1.98	3.18	3.63	4.10	4.50	4.89	5.30	5.73	6.24	6.82	7.39	8.44
$\log(B_f/V_f)$	-0.01	-0.21	-0.23	-0.26	-0.32	-0.36	-0.36	-0.44	-0.40	-0.42	-0.51	-0.65
					Panel B: Simulated Panel							
Return	1.13	1.09	1.07	1.02	0.99	0.95	0.92	0.88	0.85	0.82	0.77	0.73
β	1.11	1.10	1.02	1.00	0.98	0.97	0.95	0.94	0.93	0.91	0.87	0.87
$\log(V_f)$	4.59	0.76	0.84	0.88	4.91	4.94	4.97	5.00	5.04	5.09	5.14	5.23
$\log(B_f/V_f)$	-0.48	-0.60	-0.64	-0.67	-0.69	-0.71	-0.73	-0.75	-0.77	-0.80	-0.83	-0.87

Table 3.5 : Properties of Portfolios Formed on Book-to-Market

At the end of June of each year t, 12 portfolios are formed on the basis of ranked values of book-to-market, measured by $\log(B_f/V_f)$. The pre-ranking β's use 5 years of monthly returns ending in June of t. Portfolios 2–9 cover deciles of the ranking variables. The bottom and top 2 portfolios (1A, 1B, 10A, and 10B) split the bottom and top deciles in half. The breakpoints for the book-to-market portfolios are based on ranked values of book-to-market equity. Panel A is from Fama and French (1992) Table IV, Panel A. Panel B is from the simulated panel. The average returns are the time-series averages of the monthly equal-weighted portfolio returns in percent. $\log(V_f)$ and $\log(B_f/V_f)$ are the time-series averages of the monthly average values of these variables in each portfolio. β is the time-series average of the monthly portfolio post-ranking βs.

Portfolios	1A	1B	2	3	4	5	6	7	8	9	10A	10B
					Panel A: Historical Data							
Return	0.30	0.67	0.87	0.97	1.04	1.17	1.30	1.44	1.50	1.59	1.92	1.83
β	1.36	1.34	1.32	1.30	1.28	1.27	1.27	1.27	1.27	1.29	1.33	1.35
$\log(V_f)$	4.53	4.67	4.69	4.56	4.47	4.38	4.23	4.06	3.85	3.51	3.06	2.65
$\log(B_f/V_f)$	-2.22	-1.51	-1.09	-0.75	-0.51	-0.32	-0.14	0.03	0.21	0.42	0.66	1.02
					Panel B: Simulated Panel							
Return	0.77	0.79	0.81	0.83	0.86	0.88	0.90	0.94	0.99	1.02	1.07	1.14
β	0.98	0.95	0.94	0.97	0.99	1.00	1.01	1.03	1.02	1.03	1.05	1.04
$\log(V_f)$	5.15	5.07	5.02	5.96	5.92	4.88	4.84	4.79	4.74	4.68	4.60	4.47
$\log(B_f/V_f)$	-0.86	-0.81	-0.77	-0.73	-0.70	-0.67	-0.64	-0.61	-0.58	-0.53	-0.48	-0.39

Table 3.6 : Summary Statistics of HML and SMB

This table reports the monthly summary statistics of market excess return, HML, and SMB in data (Davis, Fama, and French [2000]) and the benchmark model. The average returns are in monthly percentage.

	$R_m - R_f$			HML			SMB		
	mean	std	t-stat	mean	std	t-stat	mean	std	t-stat
Data	0.67	5.75	3.34	0.46	3.11	4.24	0.20	3.26	1.78
Benchmark	0.56	4.50	3.56	0.41	2.61	4.49	0.35	4.65	2.15

Table 3.7 : Risk and Asymmetric Adjustment Cost

This table reports the β's with respect to the stochastic discount factor for four portfolios: HS (high and small), HB (high and big), LS (low and small), and LB (low and big) in the benchmark model. The risk is measured by the slope coefficient from the regression: $R_{pt+1} = \alpha_p + \beta_p M_{t+1} + e_{t+1}$ using three samples: full sample, good times ($x_t \geq \overline{x} + \sigma_x / \sqrt{(1 - \rho_x^2)}$), and bad times ($x_t \leq \overline{x} - 1.5\sigma_x / \sqrt{(1 - \rho_x^2)}$). The slope coefficients reported in the table are in percent.

Portfolios	Full	Good	Bad
Value			
HS	-0.57	0.02	-2.84
HB	-0.49	0.05	-2.13
Growth			
LS	-0.52	-0.03	-2.31
LB	-0.64	-0.02	-1.82

Table 3.8 : The Magnitude of Survival Bias

This table reports the monthly summary statistics of market excess return, HML, and SMB from two artificial samples generated from the model with entry and exit. The sample without bias includes all the firms that are currently in the economy, as well as those that exited during the year after portfolio formation. The sample with bias includes only those firms that are currently in the economy and exclude those exited during the year after portfolio formation. The average returns are in percentage.

	$R_m - R_f$			HML			SMB		
	mean	std	t-stat	mean	std	t-stat	mean	std	t-stat
Without Bias	0.59	5.32	3.17	0.42	3.44	3.49	0.37	4.52	2.34
With Bias	0.61	4.51	3.87	0.45	3.14	4.10	0.39	4.13	2.70

Figure 1: Asymmetric Adjustment Cost

This figure illustrates the nature of asymmetric adjustment cost. The x-axis is investment-capital ratio i/k and the y-axis is the amount of adjustment cost. The figure also shows the relative locations of value and growth firms when their investment rates are higher or lower than zero. The specification of adjustment cost is:

$$c(i_t, k_t) = \frac{g_t}{2} \left(\frac{i_t}{k_t} \right)^2 k_t$$

where $g_t = g^+ \times \mathbf{1}_{\{i_t \geq 0\}} + g^- \times \mathbf{1}_{\{i_t < 0\}}$ and where $\mathbf{1}_{\{\cdot\}}$ is the indicator function that equals one if the event described in $\{\cdot\}$ is true and equals zero otherwise. Moreover, $g^- \geq g^+ > 0$.
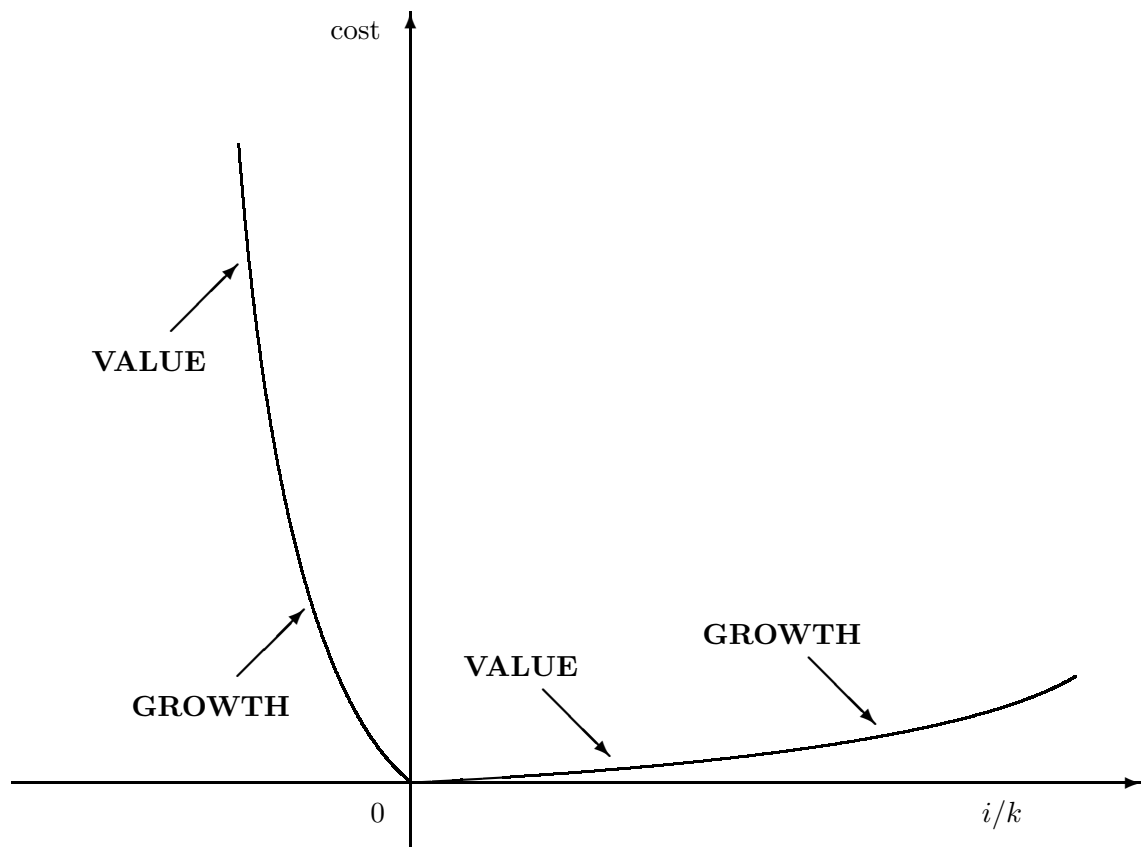
Figure 3.2 : Countercyclical Market Price of Risk

This figure plots the market Sharpe ratio, defined in (4), against the aggregate productivity shock.
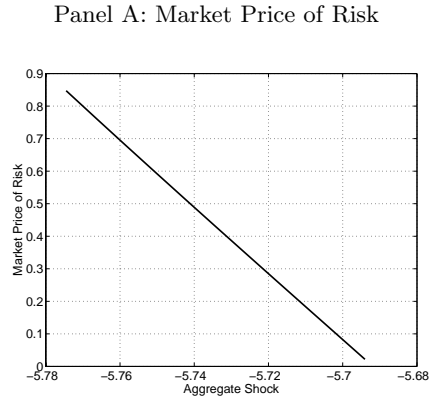
Panel A: Market Price of Risk

Figure 3.3 : Value Factor in Earnings

This figure documents the value factor in earnings. Panel A shows the 11-year evolution of earnings on book-equity $\triangle BE_{t+i} + D_{t+i}/BE_{t+i-1}$ for book-to-market portfolios. Time 0 in the horizontal axis is the portfolio formation year. Panel B shows the 11-year evolution of book-to-market BE_{t+i-1}/ME_{t+i-1} for book-to-market portfolios. Low (high) B/M indicates the portfolio containing firms in the bottom (top) 30 percent of the values of book-to-market ratios.

Panel A: Return on Equity

Panel B: Book-to-Market