Essays in Financial Economics

by

Kai Li

Department of Economics
Duke University

Date: _____
Approved:

_____
Ravi Bansal, Supervisor

_____
Hengjie Ai

_____
Tim Bollerslev

_____
Jia Li

Dissertation submitted in partial fulfillment of the requirements for the degree of
Doctor of Philosophy in the Department of Economics
in the Graduate School of Duke University
2013

Abstract

Essays in Financial Economics

by

Kai Li

Department of Economics
Duke University

Date: _____
Approved:

_____
Ravi Bansal, Supervisor

_____
Hengjie Ai

_____
Tim Bollerslev

_____
Jia Li

An abstract of a dissertation submitted in partial fulfillment of the requirements for
the degree of Doctor of Philosophy in the Department of Economics
in the Graduate School of Duke University
2013

Abstract

My dissertation, consisting of three related essays, aims to understand the role of macroeconomic risks in the stock and bond markets. In the first chapter, I build a financial intermediary sector with a leverage constraint à la Gertler and Kiyotaki (2010) into an endowment economy with an independently and identically distributed consumption growth process and recursive preferences. I use a global method to solve the model, and show that accounting for occasionally binding constraint is important for quantifying the asset pricing implications. Quantitatively, the model generates a procyclical and persistent variation of price-dividend ratio, and a high and countercyclical equity premium. As a distinct prediction from the model, in the credit crunch, high TED spread, due to a liquidity premium, coincides with low stock price and high stock market volatility, a pattern I confirm in the data.

In the second chapter, which is coauthored with Hengjie Ai and Mariano Croce, we model investment options as intangible capital in a production economy in which younger vintages of assets in place have lower exposure to aggregate productivity risk. In equilibrium, physical capital requires a substantially higher expected return than intangible capital. Quantitatively, our model rationalizes a significant share of the observed difference in the average return of book-to-market-sorted portfolios (value premium). Our economy also produces (1) a high premium of the aggregate

stock market over the risk-free interest rate, (2) a low and smooth risk-free interest rate, and (3) key features of the consumption and investment dynamics in the U.S. data.

In the third chapter, I study the joint determinants of stock and bond returns in Bansal and Yaron (2004) long-run risks model framework with regime shifts in consumption and inflation dynamics – in particular, the means, volatilities, and the correlation structure between consumption growth and inflation are regime-dependent. This general equilibrium framework can (1) generate time-varying and switching signs of stock and bond correlations, as well as switching signs of bond risk premium; (2) quantitatively reproduce various other salient empirical features in stock and bond markets, including time-varying equity and bond return premia, regime shifts in real and nominal yield curve, the violation of expectations hypothesis of bond returns. The model shows that term structure of interest rates and stock-bond correlation are intimately related to business cycles, while long-run risks play a more important role to account for high equity premium than business cycle risks.

I dedicate my dissertation to my wife, Weiwei Hu, and my son, Kevin Haoyang Li. They are my love and everything. I also have a special feeling of gratitude to my parents, Ruixing Li and Hua Gu, whose support, encouragement, and constant love have sustained me throughout my life.

Contents

List of Tables

xiii

List of Figures

Acknowledgements

1

Asset Pricing with a Financial Sector

1.1 Introduction

This paper studies the quantitative asset pricing implications of financial intermediary[1]. I embed a financial intermediary sector with an endogenous leverage constraint à la Gertler and Kiyotaki (2010) into an endowment economy. The model features a calibrated financial sector, recursive preferences, and an independently and identically distributed consumption growth process. The leverage constraint makes intermediary equity capital (net worth) to be an important state variable that affects asset prices and helps to understand a wide variety of dynamic asset pricing phenomena. Rather than a log-linear approximation method, I use a global method that allows for occasionally binding constraint to solve the model, and show the global

---

[1] In this paper, the financial intermediary sector is meant to capture the entire banking sector, including commercial banks, investment banks as well as hedge funds. Thus, I use "financial intermediary sector" and "banking sector", "financial intermediaries" and "banks", interchangeably. For the composition of aggregate financial intermediary sector, see Table A.1 in Appendix A.3.

method is critical for quantifying asset pricing implications.

Quantitatively, an i.i.d. consumption growth shock, calibrated to match the standard deviation of the aggregate consumption growth, is amplified and accumulated through the propagation mechanism of the leverage constraint, and has large and long-lasting effects on asset prices, which are absent in the model without frictions. In particular, the model produces a high equity premium (in log units) of 4.1%, a significant share (78%) of the equity premium observed in the data, a low interbank interest rate volatility of 0.58%, consistent with the data (0.55%), and a persistent and procyclical variation of price-dividend ratio, with first order autocorrelation of 65%, relatively lower than that in the data (89%). The equity premium is strongly countercyclical in the model, and predictable with the leverage ratio of aggregate financial intermediary sector, a pattern I confirm in the data. The model also produces an average stock market volatility of 16.5%, only slightly lower than a volatility of 19.8% in the data.

The leverage constraint effectively introduces a wedge between interest rates on interbank and household loans. As a distinct implication from the model, the loan spread widens significantly in the credit crunch which features a large drop in intermediary net worth. This patten is consistent with the evidence that high TED spread[2] coincides with low price-dividend ratio and high stock market volatility, as shown in Fig. 1.1.

I emphasize the importance of using a global method that accounts for occasionally binding constraint to solve the model. In the benchmark calibration with

---

[2] TED spread is measured by the spread between 3-month LIBOR rate in U.S. dollars and 3-month U.S. government treasury bill rate.

a moderate risk aversion of 10 and a calibrated financial sector, the global solution suggests the constraint only binds for about 15% of the time. A third order local approximation method, imposing the assumption that the constraint is always binding around the steady state, greatly exaggerates the volatilities of asset prices and equity premium.

There are two main ingredients in the model. First, I build a stylized leverage constraint faced by financial intermediary into an otherwise standard endowment economy. As in Gertler and Kiyotaki (2010), a limited enforcement argument that financial intermediary can divert a fraction of bank assets and default on deposits provides a microfoundation for the leverage constraint. In particular, the debt financing capacity to an intermediary is proportional to the equity capital of the intermediary times a leverage multiple. In this setup, the intermediary net worth strongly affects asset prices through an adverse dynamic feedback effect: a negative consumption shock lowers the intermediary net worth, increases the probability that constraint becomes binding in the future, and therefore reduces the borrowing capacity of the intermediary sector today and in the future. Lower borrowing capacity results in lower demand for risky assets. In the equilibrium, the intermediary sector still holds all the risky assets. To clear the market, the asset price has to fall, and risk premium has to rise. The resulting fall in asset price further lowers the net worth. An initial small i.i.d. consumption shock is endogenously amplified through this propagation mechanism.

The leverage constraint also opens up an endogenous channel of countercyclical equity premium and stock market volatility, even though consumption growth is homoscedastic. The equilibrium asset prices are more sensitive to the fundamen-

3

tal shocks when the intermediary net worth is low. As the financial intermediary sector becomes more financially constrained, both the exposure of market return to consumption shock (i.e. return beta) and the market price of the consumption shock increase, and thus contribute to a higher equity premium. In the model, price-dividend ratio and leverage ratio of the aggregate intermediary sector predict long-horizon equity returns. Both the slope coefficients and R^2 line up with the data relatively well at all horizons. And the model also captures the volatility feedback effect; that is, a consumption shock, as a negative innovation to market return, is a positive innovation to return volatility.

As a distinct feature of the model, the leverage constraint introduces a wedge between interest rates on interbank and household loans. This spread, as a measure of the tightness of leverage constraint, is countercyclical and widens significantly in bad times when the intermediary sector is extremely financially constrained. I posit a retail interbank market where the banks can trade Arrow-Debreu securities (in zero net supply) that pay one unit of net worth given a certain state among themselves frictionlessly (i.e. the bank cannot default on them), assuming the banks have monitoring technology in evaluating and monitoring their borrowers. Under this asset market structure, the banks are unconstrained in choosing risky assets and interbank loans, though they are constrained agents to obtain debt from the household. The augmented stochastic discount factor suggested by the bank's portfolio choice problem price risky assets and interbank risk-free debt. It depends not only on household consumption, but also on intermediary equity capital. The banker dislikes assets with low return when aggregate consumption is low, and when his financial intermediary has low net worth. However, the interest rate on household loans is

4

priced by a different stochastic discount factor, which is suggested by the household optimization problem. In a credit crunch, modeled as a large drop in intermediary net worth so that the constraint binds, the banks are strongly liquidity constrained to lend net worth to others, and therefore the market clearing condition drives up the interbank interest rate.

The second ingredient of the model is that quantitatively I rely on recursive preferences (Kreps and Porteus, 1978; Epstein and Zin, 1989) which allow for a separation between the intertemporal elasticity of substitution (IES, hereafter) and risk aversion, and consequently permit both parameters to be simultaneously larger than 1. I calibrate the recursive preference with a moderate risk aversion of 10 and an IES of 1.5, consistent with Bansal and Yaron (2004). In this economy, when the IES is larger than 1, the level of interest rate on household loans is low, consistent with the data. Furthermore, a high IES (larger than 1) is also critical to produce high equity premium. In the CRRA utility case, as the risk aversion increases, the IES, which is the reciprocal of risk aversion, decreases simultaneously, and leads the average leverage ratio of the financial intermediary sector to decrease very rapidly. This significantly lowers the volatility of the stochastic discount factor, due to lower volatility of shadow price of net worth. In contrast, when the IES is larger than 1, the average leverage ratio of the financial sector only decreases slowly with the risk aversion, therefore, maintains a volatile stochastic discount factor, and thus a high equity premium.

Computationally, I use a recursive method to construct a global solution which accounts for occasional binding constraint. The theoretical underpinnings of the recursive method are developed in a companion paper (Ai, Bansal and Li, 2012),

while this paper focuses on the economics and quantitative analysis of the model. In the paper, I emphasize the importance of allowing for occasional binding constraint on quantifying the asset pricing implications. In the macroeconomics literature, equilibrium is often derived by log-linearizing around the steady state and assuming the constraint is always binding, for instance, Gertler and Kiyotaki (2010), Gertler and Karadi (2011), Gertler, Kiyotaki and Queralto (2011), among others. As a result, this method does not allow me to study the model nonlinearity and off-steady-state dynamics, which are the key to internally generate the time-varying equity premium and stock volatility. Furthermore, even a higher order local approximation method imposing the assumption that the constraint is always binding around the steady state is still problematic. In the parameter configuration with which the probability of a binding constraint is low, for instance, the benchmark calibration, a third order local approximation method that forces the constraint to be always binding around the steady state greatly exaggerates the volatilities of asset prices and equity premium. I use Den Haan and Marcet simulation accuracy test (1994) to confirm the advantage of the global method over a local approximation method.

My analysis contributes to several strands of literature. First, existing consumption based asset pricing models have been successful in specifying preferences and cash flow dynamics to explain a high and countercyclical equity premium in an endowment economy (Campbell and Cochrane, 1999; Bansal and Yaron, 2004; Barro, 2006). However, these models allow no roles for financial intermediary, but assume that a representative household is marginal in pricing all the assets, therefore, they cannot speak to the close relationship between financial intermediary equity capital and aggregate stock market. They also shed no light on interest rate spread between

interbank and household loans. In this paper, I show the single channel of a leverage constraint not only links asset prices to intermediary net worth, but also provides an additional important channel to understand a wide variety of asset market phenomena, even with an i.i.d. consumption growth process. The success of the model does not rely on a very high effective risk aversion as in habit model, or on consumption risks beyond the business cycle frequency, for instance, long-run risks or rare diasters, which are hard to detect empirically in the data.

Second, this paper is directly related to Maggiori (2012) and He and Krishnamurthy (2012b) on financial intermediary and asset pricing. As a continuous time adaptation of Gertler and Kiyotaki (2010) type of leverage constraint into an endowment economy, Maggiori (2012) is a special case of the model in this paper, in which the constraint never binds in the equilibrium. Thus, it has neither implications for interest rate spread, nor implications of occasionally binding constraint on asset pricing. In He and Krishnamurthy (2012b), the financial intermediary faces an equity financing constraint, rather than a debt financing constraint. The specialist who manages the intermediary has a separate utility function different from that of representative household, and is the unconstrained marginal investors who prices all the assets, therefore, there is no household and interbank loan spread in their framework. There are several important differences between my paper with He and Krishnamurthy (2012b). First, in my model, the stochastic discount factor depends both on the aggregate consumption growth and the marginal value of net worth, the variations of which are also driven by the aggregate consumption growth shock. However, in their model, the marginal investor's consumption process is not based on the aggregate consumption, rather it is endogenously determined by his portfolio

7

choice problem. Their model predicts significantly negative risk-free interest rate in the crisis, which suggest that model implied consumption volatility of marginal investor in the crisis state is very high, and thus induces a large precautionary saving effect to lower the risk-free rate. Second, in my model, the amplification effect is quantitatively large around the steady state where the constraint is not binding, due to the fact that the concern about potential future losses in net worth depresses the stock market today. However, He and Krishamurthy (2012b) framework is only to capture the risk premium behavior in crises, but features no amplification effect in the unconstrained region.

Third, the paper also relates to the theoretical literature on intermediary frictions. There are two broad classes of theories: leverage-constraints theories and equity risk-capital constraints. Both theories start with the assumption that intermediaries are constrained in raising more equity. They share two common predictions: First, intermediary equity (or net worth) is the key state variable that affects asset prices. Second, the effect of intermediary equity on asset prices is nonlinear, with a larger effect when the intermediary equity is low. The leverage-constraints models include Geanakoplos and Fostel (2008), Adrian and Shin (2010) and Brunnermeier and Pedersen (2009), Danielsson et al. (2011), Geanakoplos (2012), and Adrian and Boyarchenko (2012). Gertler and Kiyotaki (2010) type of frictions lies in the first category. He and Krishnamurthy (2012a) and Brunnemeier and Sannikov (2012) are examples of equity risk-capital models. The goal of this paper is different from the theoretical literature to propose alternative microfoundations for financial frictions, rather I focus on the quantitative asset pricing implications of a stylized type of leverage constraint as in Gertler and Kiyotaki (2010), which has been widely

8

studied in the macroeconomic and policy related literature.

More broadly, this paper is related to the literature in macroeconomics studying the effects of financial frictions on aggregate activity, including Kiyotaki and Moore (1997), Calstrom and Fuerst (1997) and Bernanke, Gertler and Gilchrsit (1999), among others. These papers focus on the credit frictions faced by non-financial borrowers. Gertler and Kiyotaki (2010) introduces a leverage constraint between household and financial intermediary, also see Gertler and Karadi (2011), Gertler, Kiyotaki and Queralto (2011), Gertler and Karadi (2012), among others. The equilibrium in these works is derived by log-linearizing around the steady state and assuming the constraint is always binding. Instead, I use a global method to solve the model, and emphasize that accounting for occasionally binding constraint is very important for quantifying asset pricing implications of the model. My work contributes to the literature by arguing that quantitative analysis on macroeconomic effects and policy evaluations of financial frictions should take into account the importance of occasionally binding constraint on asset price dynamics, which lie in the center of the propagation mechanism of financial frictions.

The remainder of the paper is organized as follows: I present the model setup and define the competitive equilibrium in Section 1.2. In Section 1.3, I outline model solution, computation and discuss some analytical results in asset pricing. Section 1.4 presents benchmark model's performance in various aspects. Section 1.5 provides some additional asset pricing implications, and Section 1.6 concludes and lays down several extensions on my research agenda. Model derivations, data sources and computation details are provided in the Appendix A.

1.2   The Model Setup

I embed a financial intermediary sector with a leverage constraint à la Gertler and Kiyotaki (2010) into a standard Lucas (1978) endowment economy .

There are three sectors in the economy, namely, households, financial intermediaries (banks), and non-financial firms. I assume households cannot invest directly in the risky asset market by holding the equity of non-financial firms. There is a limited market participation, also see Mankiw and Zeldes (1991), Basak and Cuoco (1998), or Vissing-Jorgensen (2002). Instead, households can only save through a risk-free deposit account with banks. Each household owns a unit mass of banks coming in overlapping generations. Banks borrow short-term debt from households[3], and invest in the equity of the firms. In addition to assisting in channeling funds from households to non-financial firms, banks engage in maturity transformation. They hold long term assets and fund these assets with short term liabilities (beyond their own equity capital). In addition, the banking sector in this model is meant to capture the entire banking sector, including commercial banks, investment banks as well as hedge funds.

Time is discrete and infinite, $t = 0, 1, 2, \cdots$. The non-financial firms in this economy are modeled as in a Lucas (1978) tree economy which pays aggregate output every period. The aggregate output is denoted by $Y_0, Y_1, Y_2, \cdots$. The log growth rate of the output process is given by

$$\log\left(\frac{Y_{t+1}}{Y_t}\right) = \mu_y + \sigma\varepsilon_{y,t+1},$$

---

[3] To motivate a limited enforcement argument later, it is best to think of banks only obtaining deposits from households who do not own them.

in which $\varepsilon_{y,t+1}$ is an i.i.d. random variable with mean zero and unit variance, modeled as a finite-state Markov chain. The parameter σ captures the aggregate consumption volatility.

I use Q_t to denote the price of the Lucas tree at period t, and thus the total return on the Lucas tree, $R_{y,t+1}$, is defined as

$$R_{y,t+1} = \frac{Q_{t+1} + Y_{t+1}}{Q_t}.$$

1.2.1   Households

There is a unit mass of identical households who makes intertemporal consumption and saving decisions. I collapse all households into a single representative household. He is infinitely lived and maximizes the objective function,

$$\max_{\{C_t, B_t\}_{t=0}^{\infty}} E_0 \left[\sum_{t=0}^{\infty} \beta^t u(C_t) \right], \tag{1.1}$$

where C_t is the period t consumption. I consider a constant relative risk aversion (CRRA, hereafter) instantaneous utility function with risk aversion parameter γ, $u(C_t) = \frac{1}{1-\gamma} C_t^{1-\gamma}$. In the subsequent quantitative analysis (Section 1.4), I use more general recursive preferences (Kreps and Porteus, 1978; Epstein and Zin, 1989), which disentangle the risk aversion with IES. This is quantitatively important for asset pricing implications as discussed in Section 1.4.1 and Section 1.4.5. More details about recursive preferences are provided in Appendix A.1.

The household can only save through a risk-free deposit account with banks. Let $\{\pi_t\}_{t=0}^{\infty}$ denote the stream of (stochastic) income that the household receives, and

$R_{f,t}$ denote the one-period risk-free interest rate for a loan (made by the household to the banks) that pays off on date $t+1$. A set of budget constraints (1.2) is described as the following:

$$
\begin{aligned}
C_0 + B_0 &= \pi_0, \qquad\qquad\qquad\qquad\qquad\qquad\qquad (1.2) \\
C_t + B_t &= B_{t-1} R_{f,t-1} + \pi_t, \ t \geq 1.
\end{aligned}
$$

In the above formulation, the household receives a stream of income, $\{\pi_t\}_{t=0}^{\infty}$ and makes his consumption and saving decisions. C_t is the period t consumption choice, and B_t is the amount he deposits in the one-period risk-free bond, which pays a gross interest rate $R_{f,t}$ in the next period. I will show later on, π_t is the amount of wealth transferred from the banking sector to the household at period t. That is, his ownership of the banks pays off over time as an income stream $\{\pi_t\}_{t=0}^{\infty}$. Technically, the $\{\pi_t\}_{t=0}^{\infty}$ sequence is constructed so that it can be easily verified that $C_t = Y_t$ satisfies the budget constraint.

1.2.2 Financial Intermediaries

The banks come in overlapping generations. Denote n_{t+j}^t to be the total amount of net worth held by all generation t banks at period $t+j$, and s_{t+j}^t, the total number of shares in the Lucas tree held by all generation t banks at period $t+j$. I use Λ_t to denote the Arrow-Debreu price of one unit of consumption good at period t denominated in terms of time 0 consumption goods. Under this notation, the price of a unit of consumption good at period $t+j$ denominated in terms of period t consumption good is $\frac{\Lambda_{t+j}}{\Lambda_t}$. Given the price system $\{\Lambda_t\}_{t=0}^{\infty}$, a generation t bank

maximizes the present value of its future cash flow by choose:

$$\max_{\{s^t_{t+j}, n^t_{t+j+1}\}^\infty_{j=0}} E_t \left[\sum_{j=1}^{\infty} \frac{\Lambda_{t+j}}{\Lambda_t} (1-\lambda)^{j-1} \lambda n^t_{t+j} \right]. \tag{1.3}$$

In each period, a fraction λ of the bank is forced to liquidate, in which case, their net worth is paid off as dividend. The remaining fraction $(1-\lambda)$ will survive to the next period. The liquidation fraction/probability is i.i.d. across banks and time. As a result, the total fraction of a generation t survived until period $t+j$ is $(1-\lambda)^{j-1}$, and a fraction λ is paid out as dividend. Note that the bank and household share the same stochastic discount factor, Gertler and Kiyotaki (2010) provide an "insurance story" to justify this.

Equation (1.4) is the initial condition of banks' net worth. The initial generation starts with initial net worth N_0. After that, in each period, the household uses a fraction δ of the Lucas tree to set up new banks, as assumed in Gertler and Kiyotaki (2010). Therefore, $\delta[Q_t + Y_t]$ is the initial net worth of the generation t bank at period t.

$$n^t_t = \delta[Q_t + Y_t] \ if \ t \geq 1, \ n^0_0 = N_0. \tag{1.4}$$

Equation (1.5) is the law of motion of net worth. At period $t+j$, the bank started with net worth n^t_{t+j} and chooses hold s^t_{t+j} shares of the stock. Each share pays $Q_{t+j+1} + Y_{t+j+1}$ in the next period, which is the first term on the right hand side of (1.5). However, the bank has to borrow $s^t_{t+j}Q_{t+j} - n^t_{t+j}$ from the household in order to finance the purchase of the stock. The second term on the right hand side

13

of (1.5) is the amount of loan repayment the bank has to deliver to the household in period $t + j + 1$.

$$n_{t+j+1}^t = s_{t+j}^t \left[Q_{t+j+1} + Y_{t+j+1} \right] - \left[s_{t+j}^t Q_{t+j} - n_{t+j}^t \right] R_{f,t+j}, \quad for\ all\ \ j \geq 0. \quad (1.5)$$

Equation (1.6) is the participation constraint, motivated by a limited enforcement argument in Gertler and Kiyotaki (2010). At period $t + k$ in the future, the banker has an opportunity to divert a θ fraction of bank assets at its market price and default on its debt. And the depositors can only recover $(1-\theta)$ fraction of bank asset, due to limited enforcement. Because the depositors recognize the bank's incentive to divert funds, they will restrict the amount they lend. In this way a participation constraint arises: we need to make sure that the value of the bank must exceed the banker's outside option in all future periods. Note that there are infinitely many participation constraints from period t on into the future.

$$E_{t+k} \left[\sum_{j=1}^{\infty} \Lambda_{t+k+j} \left(1 - \lambda \right)^{j-1} \lambda n_{t+j}^t \right] \geq \theta s_{t+k}^t Q_{t+k}, \quad for\ all\ \ k \geq 0. \quad (1.6)$$

1.2.3 Competitive Equilibrium

A competitive equilibrium is a collection of prices, $\{Q_t, R_{f,t}, \Lambda_t, \pi_t\}_{t=0}^{\infty}$, and quantities $\left\{ \left\{ s_{t+j}^t, n_{t+j}^t \right\}_{j=0}^{\infty}, N_t \right\}_{t=0}^{\infty}$ that satisfy 1) household utility maximization; 2) banks of each generation maximize profit; 3) market clearing conditions; 4) a set of consistency conditions.

The market clearing conditions include:

$$C_t = Y_t, \quad (1.7)$$

$$\sum_{j=0}^{t} s_t^{t-j} = 1, \tag{1.8}$$

$$\sum_{j=0}^{t} n_t^{t-j} = N_t, \tag{1.9}$$

$$N_t = (1 - \lambda)\left[Q_t + Y_t - (Q_{t-1} - N_{t-1}) R_{f,t-1}\right] + \delta\left[Q_t + Y_t\right], \tag{1.10}$$

$$Q_t = N_t + B_t, \tag{1.11}$$

$$\pi_0 = Q_0 + Y_0 - N_0, \tag{1.12}$$

$$\pi_t = \lambda\left[Q_t + Y_t - (Q_{t-1} - N_{t-1}) R_{f,t-1}\right] - \delta\left[Q_t + Y_t\right], \text{ for } t \geq 1. \tag{1.13}$$

Here (1.8) says all shares owned by existing generations of banks must sum up to 1, and (1.9) says the total net worth of banks of all generations must sum up to N_t. Equation (1.10) is the accounting identity. In equation (1.10), I use N_t to denote the total net worth of the banking sector (the total amount of wealth held by all the banks). The first part of the right-hand side of equation (1.10) is the total amount of net worth of all banks at date t that comes from existing banks (banks of generation $t-1$ and older): At period t-1 all existing banks together own one share of the Lucas tree, which pays off $Q_t + Y_t$. They have net worth N_{t-1}, and borrowed $Q_{t-1} - N_{t-1}$ to buy the tree. Consequently, $(Q_{t-1} - N_{t-1}) R_{f,t-1}$ is the amount of interest they have to return to the household. The second part of the right-hand side of equation (1.10), $\delta[Q_t + Y_t]$, is the amount of net worth that is newly injected into the banking sector at period t. Recall that each period the household use δ fraction of the Lucas tree to set up a new generation of banks.

The market clearing condition also include (1.12) and (1.13). Equation (1.12) implies the household and the bankers together own the Lucas tree. In particular,

15

the household owns part of the Lucas tree directly, through π_0, and owns part of the Lucas tree indirectly, through the banks, which is N_0. Equation (1.13) has the following interpretation: in period t, a λ fraction of all existing banks are forced to liquidate, and their net worth flows into the household. At the same time, the household also used δ fraction of the value of the Lucas tree to set up new banks. This completes the discussion of the market clearing conditions. Of course, given the budget constraint and market clearing conditions, one of (in each period) is redundant according to Walras' law.

I also need certain consistency condition:

$$\Lambda_t = \frac{\beta^t u'(C_t)}{u'(C_0)},$$

which captures the "insurance story" that Gertler and Kiyotaki (2010) tells.

1.3   Model Solution

In this section, I outline the main steps in deriving the solution, highlighting the economic mechanism linking intermediary equity capital and the asset prices. Detailed derivations are provided in the Appendix A.2.

1.3.1   State Variable and its Dynamics

In this economy, financial intermediary equity capital is an important state variable that affects asset prices. As I comment below, in the discrete time setup, it turns out to be more convenient to use normalized debt, instead of net worth, as the state variable. Both debt and net worth measure the capitalization of financial

16

intermediary sector, and therefore, I may use them interchangeably in explaining the model intuitions.

I define normalized debt level as,

$$\mathbf{b}_t = \frac{B_{t-1} R_{f,t-1}}{Y_t},$$

as the state variable of the economy, where

$$B_{t-1} = Q_{t-1} - N_{t-1},$$

is the total amount of debt that the banks borrow from the household sector in period $t-1$. Because this is a growth economy, I normalize quantities and prices by total output, and denote:

$$q(\mathbf{b}_t) = \frac{Q(\mathbf{b}_t)}{Y_t}; \ \frac{Y_{t+1}}{Y_t} = g_{t+1}; \ \hat{n}_t = \frac{N_t}{Y_t}, \tag{1.14}$$

The law of motion for the state variable is therefore,

$$\mathbf{b}_{t+1} = \frac{R_{f,t}}{g_{t+1}} \left\{ (1-\lambda)\, \mathbf{b}_t + (\lambda - \delta)\, q(\mathbf{b}_t) - (1 - \lambda + \delta) \right\}. \tag{1.15}$$

One advantage of using $\mathbf{b}_t$ as the state variable is that: given today's $\mathbf{b}_t$ and an initial guess of the price functional $q(\cdot)$, the law of motion (1.15) determines $\mathbf{b}_{t+1}$ in close form. This property facilitates an iterative procedure to compute the equilibrium, as discussed in Section 1.3.4. However, if I use normalized net worth $\hat{n}_t$ as the state variable, a choice in Maggiori (2012) and He and Krishnamurthy (2012), I find that the law of motion of normalized net worth $\hat{n}_t$ in this discrete time context is not in closed form.

17

I can now express the current period net worth, $\widehat{n}_t$, as a function of the state variable $\mathbf{b}_t$:

$$
\begin{aligned}
\widehat{n}_t &= q\left(\mathbf{b}_t\right) - \frac{\mathbf{b}_{t+1} g_{t+1}}{R_{f,t}}, \\
&= (1 - \lambda + \delta) q\left(\mathbf{b}_t\right) - \left[(1 - \lambda) \mathbf{b}_t - (1 - \lambda + \delta)\right].
\end{aligned}
\tag{1.16}
$$

1.3.2 Recursive Formulation of Bank's Problem

I first set up some notations. I use M_{t+1} to denote the one-period stochastic discount factor implied by household problem, as standard in the asset pricing literature. That is,

$$
M_{t+1} = \frac{\Lambda_{t+1}}{\Lambda_t} = \beta \frac{u'\left(C_{t+1}\right)}{u'\left(C_t\right)}.
$$

Euler equation implies:

$$
E\left(M_{t+1}\right) R_{f,t} = 1.
$$

Both M_{t+1} and $R_{f,t}$ do not depend on $\mathbf{b}_t$, rather, they are determined by aggregate consumption growth process in the equilibrium. Note that with i.i.d. consumption growth, $R_{f,t}$ is a constant, which I denoted as R_f.

The bank's optimization problem has a recursive representation:

$$
V\left(\mathbf{b}_t, n_t\right) = \max_{\{s_t, n_{t+1}\}} E_t\left[M_{t+1}\left\{\lambda n_{t+1} + (1 - \lambda) V\left(\mathbf{b}_{t+1}, n_{t+1}\right)\right\}\right]
\tag{1.17}
$$

$$
subject\ to: n_{t+1} = s_t\left[Q\left(\mathbf{b}_{t+1}\right) + Y_{t+1}\right] - \left[s_t Q\left(\mathbf{b}_t\right) - n_t\right] R_f,
\tag{1.18}
$$

$$
E_t\left[M_{t+1}\left\{\lambda n_{t+1} + (1 - \lambda) V\left(\mathbf{b}_{t+1}, n_{t+1}\right)\right\}\right] \geq \theta s_t Q\left(\mathbf{b}_t\right).
\tag{1.19}
$$

Given initial wealth n_t and the current state $\mathbf{b}_t$, the bank chooses control variables $\left(s_t, n_{t+1}\right)$, subject to the constraints. The constraint (1.18) essentially determines

n_{t+1} given the choice s_t and the realization of the random variables exogenous to the maximization problem. I do not substitute out n_{t+1} just to save notation. Since n_{t+1} depends on s_t, constraint (1.19) restricts the choice of s_t.

I conjecture that $V(\mathbf{b}_{t+1}, n_{t+1})$ is of the form[4]:

$$V(\mathbf{b}_{t+1}, n_{t+1}) = \mu(\mathbf{b}_{t+1}) n_{t+1}, \tag{1.20}$$

in which $\mu(\mathbf{b}_{t+1})$ is the shadow price of net worth at time $t+1$. In this case, the maximization problem can be written as:

$$V(\mathbf{b}_t, n_t) = \max_{\{s_t, n_{t+1}\}} E_t\left[M_{t+1}\left\{\lambda + (1-\lambda)\mu(\mathbf{b}_{t+1})\right\} n_{t+1}\right] \tag{1.21}$$

$$subject\ to: n_{t+1} = s_t\left[Q(\mathbf{b}_{t+1}) + Y_{t+1}\right] - \left[s_t Q(\mathbf{b}_t) - n_t\right] R_f,$$

$$E_t\left[M_{t+1}\left\{\lambda + (1-\lambda)\mu(\mathbf{b}_{t+1})\right\} n_{t+1}\right] \geq \theta s_t Q(\mathbf{b}_t).$$

Given $\{\mu(\mathbf{b}_{t+1}), Q(\mathbf{b}_{t+1})\}$, I define

$$v(\mathbf{b}_t) = \lambda + (1-\lambda) E_t\left[M_{t+1}\mu(\mathbf{b}_{t+1})\right] R_f, \tag{1.22}$$

in which $v(\mathbf{b}_t)$ is the shadow price of net worth at date t if the participation constraint is not binding for any bank. Also, I define

$$P(\mathbf{b}_t) = \frac{E_t\left[M_{t+1}\left\{\lambda + (1-\lambda)\mu(\mathbf{b}_{t+1})\right\}(Q(\mathbf{b}_{t+1}) + Y_{t+1})\right]}{v(\mathbf{b}_t)}. \tag{1.23}$$

in which $P(\mathbf{b}_t)$ is the equilibrium price of the Lucas tree in the case where the participation constraint does not bind for any bank. Note that $v(\mathbf{b}_t)$ and $P(\mathbf{b}_t)$ is completely determined once the functional form of $\{\mu(\mathbf{b}_{t+1}), Q(\mathbf{b}_{t+1})\}$ is known.

---

[4] Note this is not saying that the equilibrium solution is nonlinear. It says, given equilibrium prices, the bank's value function is linear. The equilibrium prices are highly nonlinear, and are determined by some nonlinear method.

As shown in the Appendix A.2, I can summarize the equilibrium conditions with a compact notation.

$$Q\left(\mathbf{b}_t\right) = \frac{v\left(\mathbf{b}_t\right) P_t\left(\mathbf{b}_t\right) + v\left(\mathbf{b}_t\right) N_t\left(\mathbf{b}_t\right) \wedge \theta P_t\left(\mathbf{b}_t\right)}{v\left(\mathbf{b}_t\right) + \theta}. \tag{1.24}$$

Also,

$$\mu\left(\mathbf{b}_t\right) = v\left(\mathbf{b}_t\right) \vee \frac{\theta Q\left(\mathbf{b}_t\right)}{N_t}. \tag{1.25}$$

in which $P\left(\mathbf{b}_t\right)$ and $v\left(\mathbf{b}_t\right)$ are given by (1.22) and (1.23). Here I used the short-hand notation $x \wedge y \equiv \min\left\{x, y\right\}$ and $x \vee y = \max\left\{x, y\right\}$. Obviously, $Q\left(\mathbf{b}_t\right) \leq P\left(\mathbf{b}_t\right)$ and $\mu\left(\mathbf{b}_t\right) \geq \nu\left(\mathbf{b}_t\right)$, and strict inequality holds if and only if the participation constraint is binding.

1.3.3 Parameter Requirement

Parameter Assumption: I focus on the parameter that the lowest possible realization of consumption growth, g_L, is bounded by:

$$(1 - \lambda) R_f < g_L < \frac{(1 - \lambda) R_f}{(1 - \lambda + \delta)}.$$

The first part of the inequality implies that the minimum consumption growth rate of the economy cannot be too low. The intuition is that if the shocks are too low, a long enough sequence of bad shocks will send the total debt level in the banking sector to infinity, which cannot be consistent with any equilibrium. This observation has important consequences. For example, it implies that it would be inappropriate to consider a discrete time model with normal shocks, because

20

the shocks are unbounded. The log-linearization method ignores this equilibrium restriction.[5]

In this economy, a λ fraction of net worth exits the banking sector and a δ fraction of the market value of the Lucas tree is injected back into the banking sector in each period. If λ is small enough, or δ is large enough, the bank will eventually get out of the constraint. The second part of the inequality makes sure that we focus on the interesting case that λ is large enough and δ is small enough, so that the economy will not grow out of the constraint with probability one.

The theoretical results on the parameter assumptions are provided in Ai, Bansal and Li (2012).

1.3.4 Computation

The literature[6] usually uses a local approximation method to solve the model with Gertler and Kiyotaki (2010) type of participation constraint, imposing the assumption that the constraint is always binding around the steady state. One exception is Maggiori (2012), which features a analytical global solution up to a system of ordinary differential equations (ODEs) in a continuous time setting with log utility. In this paper, I use a global method to solve the model with recursive preferences in a discrete time context, allowing for occasionally binding constraint. In Section 1.4.1, I use quantitative experiments to show that the global method allowing for occasionally binding constraint is critical to quantify the asset pricing implications in such a

---

[5] This is not an issue in continuous time given Maggiori(2012)'s experiment, as in continuous time, as time interval shrinks, so does the size of the shocks.

[6] An incomplete list includes Gertler and Kiyotaki (2010), Gertler and Karadi (2011), Gertler, Kiyotaki and Queralto (2011), Gertler and Karadi (2012), among others.

model.

There are several reasons which make the model computation special. First, this model features an incomplete market, and thus the competitive equilibrium defined in Section 1.2.3 does not correspond to a social planner's solution. Instead, we need to solve the competitive equilibrium directly. Second, because of the occasionally binding constraint (1.19), standard local approximation methods, for instance, perturbation method, cannot be used, unless we impose the assumption that the constraint always binds around the steady state. As such, I use a recursive method, the theoretical underpinnings of which are developed in Ai, Bansal and Li (2012), to construct the global solution. Third, because of the nonlinearity of the model and my focus on nonlinearity-sensitivity of asset prices with the state variable, I solve the model on a large number of grid points to ensure accuracy.

To summarize the intuition of an iterative procedure to solve the model, the following system (1.26), (1.27), (1.28), and (1.29) defines a mapping $\{\mu(\mathbf{b}'), q(\mathbf{b}')\} \Longrightarrow \{\mu(\mathbf{b}), q(\mathbf{b})\}$, in which I use the convention that "'" denotes next period quantities. This system is normalized version of the system (1.22)-(1.25).

$$v(\mathbf{b}) = \lambda + (1 - \lambda) E[M'\mu(\mathbf{b}')] R_f, \tag{1.26}$$

$$p(\mathbf{b}) = \frac{E[M'\{\lambda + (1-\lambda)\mu(\mathbf{b}')\}\{q(\mathbf{b}') + 1\}g']}{v(\mathbf{b})}, \tag{1.27}$$

$$q(\mathbf{b}) = \frac{\nu(\mathbf{b})p(\mathbf{b}) + v(\mathbf{b})n(\mathbf{b}) \wedge \theta p(\mathbf{b})}{v(\mathbf{b}) + \theta}, \tag{1.28}$$

$$\mu(\mathbf{b}) = v(\mathbf{b}) \vee \frac{\theta q(\mathbf{b})}{\hat{n}}. \tag{1.29}$$

22

in which $\hat{n}$ is determined by equation (1.16), and the law of motion of the state variable, b, is given by equation (1.14).

If we find pricing functions $\{\mu(\mathbf{b}), q(\mathbf{b})\}$ that satisfy the above "functional equations", and under the equilibrium pricing functions, $\widehat{n}_t$ stays strictly positive for all t starting from any initial condition[7], then we can use these pricing functions to construct equilibrium. The basic intuition for an iterative procedure is the following: Given $\mathbf{b}$, I conjecture the pricing functions $\{\mu(\mathbf{b}), q(\mathbf{b})\}$, and solve $\mathbf{b}'$ in close form from the law of motion (1.15) and hence $\{\mu(\mathbf{b}'), q(\mathbf{b}')\}$. I then use the equilibrium conditions summarized in the above system (1.26), (1.27), (1.28), and (1.29) to solve a new set of market clearing prices. Start with the new prices, and do iterations. The equilibrium prices are the fixed points suggested by this iteration procedure. Some additional details of the computation procedure are provided in the Appendix A.4.

1.3.5 Asset Pricing

In this section, I discuss the asset pricing implications of the model from the equilibrium conditions. To save notations, I do not explicitly express asset prices as functions of the state variable $\mathbf{b}$ when there is not confusion, instead I summarize this dependence in the time subscript "t", as standard in the literature.

The Microfoundation of a Leverage Constraint

At the equilibrium, by the property of value function (1.20), the participation constraint can be expressed as

$$\mu_t n_t \geq \theta Q_t s_t.$$

---

[7] As shown in Lemma 4 of Appendix refA3:sec8.1, only strictly positive $\widehat{n}_t$ can be supported by the equilibrium.

Therefore, the participation constraint provides a microfoundation for a leverage constraint:

$$\frac{Q_t s_t}{n_t} \leq \frac{\mu_t}{\theta},$$

in which a bank's leverage ratio is defined as its total assets over net worth. First, the bank's maximum leverage ratio, $\frac{\mu_t}{\theta}$, does not depend on bank-specific factors. This nice property allows me to sum across individual banks to obtain the relation for the demand for total bank assets as a function of total net worth,

$$Q_t \leq \frac{\mu_t}{\theta} N_t.$$

Note that the demand for total bank assets is equal to Q_t, because all the firm equity is concentrated in the banking sector, and the total number of shares of equity is normalized to 1.

Second, the maximum leverage ratio depends on the aggregate state variable $\mathbf{b}_t$, and is countercyclical, as the shadow price of net worth μ_t is high in bad times when net worth is scarce. This model feature is consistent with the empirical evidence on the leverage ratio of the aggregate intermediary sector, as shown in Fig. 1.2.

Expecting that a bank will be able to abscond with stocks purchased with loans from household, household will require a collateral posted against the loans. Therefore, the participation constraint can be also rewritten/reinterpreted and aggregated as a collateral constraint, as follows:

$$B_t \leq \left(\frac{\mu_{\mathbf{t}}}{\theta} - 1\right) N_t. \tag{1.30}$$

On the left hand side of (1.30), the aggregate loans from household sector, B_t, is equal to $Q_t - N_t$, as one of the market clearing conditions (1.11). The right hand of

(1.30) is equal to aggregate net worth of the banking sector with a multiplier. It can be considered as the collateral required by the household to post against the loans.

Setup of the Asset Market

I posit a retail interbank market where the banks can trade Arrow-Debreu securities (in zero net supply) that pay one unit of net worth given a certain state among themselves. Suppose that the banks have a better enforcement/monitoring technology than households, therefore, the Arrow-Debreu securities are traded frictionless, i.e. no banks can default on them. Due to zero net supply, the market clearing condition pins down the Arrow-Debreu prices. In this sense, the stochastic discount factor suggested by the banks' portfolio choice problem (defined in equation (1.32)) can price all the assets traded frictionlessly among banks, with their payoffs being replicated by the Arrow-Debreu securities. Two classes of such assets of my interest are discussed in order.

First, risky assets. I distinguish between the unobservable return on a claim to aggregation output (consumption), $R_{y,t+1}$, and the observable return on the market portfolio, $R_{m,t+1}$; the latter is the return on the aggregate dividend claim. As in Campbell and Cochrane (1999) and Bansal and Yaron (2004), I model aggregate consumption and aggregate dividend as two separate processes. In particular, the log growth rate of aggregate dividend is specified as:

$$\log\left(\frac{D_{t+1}}{D_t}\right) = \mu_d + \varphi\sigma\varepsilon_{y,t+1} + \varphi_d\sigma\varepsilon_{d,t+1}.$$

in which $\varepsilon_{y,t+1}$ is the consumption shock specified as an i.i.d. random variable with finite state Markov chain as before, and $\varepsilon_{d,t+1}$ is standard Normally distributed, and

captures the dividend growth shock that is uncorrelated with consumption growth shock. Two additional parameters $\varphi > 1$ and $\varphi_d > 1$ allow me the calibrate the overall volatility of dividends (which is larger than that of consumption in the data) and its correlation with consumption. I use $Q_{d,t}$ to denote the price of the dividend claim, and the market return is thus defined as,

$$R_{m,t+1} = \frac{Q_{d,t+1} + D_{t+1}}{Q_{d,t}}.$$

Second, the interbank loans that lend one unit of net worth today and return (pay back) $R_{f,t}^L$ units in the next period, which $R_{f,t}^L$ denotes the gross interest rate.

Asset Pricing

In this section, I discuss the equilibrium conditions that determines the returns of three kinds of assets, namely, interest rates on household and interbank loans, and the returns for risky assets.

The interest rate on household loans is determined by the Euler equation of household problem, unaffected by frictions and has the standard interpretation of the optimal trade-off between consumption and savings.

Lemma 1. *The interest rate for the loans from the household sector, $R_{f,t}$, must satisfy*

$$E\left[M_{t+1}\right] R_{f,t} = 1.$$

Under the asset market structure in the interbank market discussed in last section, although the banks re constrained in obtaining household deposits, they are unconstrained in choosing risky assets and interbank loans. The stochastic discount

factor suggested by the bank's portfolio choice problem price the risky assets and the interbank loans.

Lemma 2. *The returns, R_{t+1}, for any assets that financial intermediary can trade frictionlessly among themselves (i.e. "frictionless" means that bank cannot default on them), including $R_{m,t+1}, R_{y,t+1}$ and $R_{f,t}^L$, must satisfy*

$$E\left[M_{t+1}\left\{\lambda + (1-\lambda)\mu_{t+1}\right\}R_{t+1}\right] = \Omega_t, \qquad (1.31)$$

in which

$$\Omega_t = \nu_t \frac{P_t}{Q_t},$$

$$= \nu_t + \theta\left(1 - \frac{\nu_t}{\mu_t}\right).$$

I use $\widetilde{M}_{t+1}$ to denote the "augmented stochastic discount factor" implied by bank's optimization problem,

$$\widetilde{M}_{t+1} = M_{t+1}\frac{\lambda + (1-\lambda)\mu_{t+1}}{\Omega_t}, \qquad (1.32)$$

which can price all the assets traded frictionlessly among banks. Beside M_{t+1}, the intertemporal marginal rate of substitution of consumption, $\widetilde{M}_{t+1}$ also depends on an additional component, Φ_{t+1}, which I define as:

$$\Phi_{t+1} = \frac{\lambda + (1-\lambda)\mu_{t+1}}{\Omega_t}. \qquad (1.33)$$

The term, $\lambda + (1-\lambda)\mu_{t+1}$, is a measure of shadow price of net worth at the next period, which is a weighted average of marginal value of net worth given the bank

27

is forced to liquidate or not. Based on the equation (1.31), Ω_t can be interpreted as the (risk adjusted) present value (in term of consumption good) of investing one unit of net worth for one period, which is a measure of the marginal value of net worth at current period. Thus, we can think of the second component, Φ_{t+1}, as the shadow price appreciation from period t to $t+1$. And the augmented stochastic discount factor has the interpretation of the intertemporal marginal rate of substitution with respect to additional unit of net worth. $\widetilde{M}_{t+1}$ depends not only on household consumption, but also on intermediary equity capital. The banker dislikes assets with low return when aggregate consumption is low, and when his financial intermediary has low net worth/high debt.

Up to a log-normal approximation[8], I use m, π and r denote the logarithm terms, and derive a two-factor model for risk premium for all assets:

$$E_t\left(r_{t+1} - r_{f,t}^L\right) + \frac{1}{2}var_t\left(r_{t+1}\right) = -cov_t\left(m_{t+1}, r_{t+1}\right) - cov_t\left(\phi_{t+1}, r_{t+1}\right). \qquad (1.34)$$

One the right hand side of equation (1.34), the first term, $-cov_t\left(m_{t+1}, r_{t+1}\right)$, is standard as in the economy without frictions. The second term, $-cov_t\left(\phi_{t+1}, r_{t+1}\right)$, is responsible for asset pricing impacts for the additional channel of a leverage constraint. As shown in Section 1.4.3, the non-linear sensitivity of the marginal value of net worth, μ_{t+1}, with respect to a fundamental shock, translates into countercyclical exposure of ϕ_{t+1} to the shock, and therefore, generates countercyclical market price of risk.

---

[8] The log-normality assumption may not be a good approximation here, as the model endogenously generates negative skewness and excess kurtosis to asset prices. This assumption facilitates to obtain a two-factor asset pricing equation for expressional purpose. The model computation and qualitative results in the paper do not rely on this assumption.

Note that two interest rates are priced by different stochastic discount factors, therefore, there is an interest rate spread, as stated in the following lemma:

Lemma 3. *The interest rate spread, defined as the difference between interest rate on interbank loans, $R^L_{f,t}$, and interest rate on household, $R_{f,t}$, is equal to zero when participation constraint is not binding, but becomes strictly positive when the constraint binds.*

First, we have $R_{f,t} = R^L_{f,t}$ whenever the constraint is not binding, because in this case, the leverage constraint is slack and both loans act as a perfect substitute. Second, we have $R_{f,t} \leq R^L_{f,t}$ when the intermediary sector is constrained. From the demand perspective, interbank borrowing is very attractive. It allows banks to invest in the stock without affecting their debt capacity with the household. As a result, all banks want to borrow from each other on the interbank market. Market clearing requires interest rate to go up to clear the market. I will provide more intuitions on the interest rate spread in Section 1.4.3 through quantitative results.

1.4   Quantitative Results

In this section, I calibrate the model at an annual frequency and evaluate its ability to replicate key moments of both cash flow dynamics and asset returns. I focus on a long sample of U.S. annual data $(1930 - 2011)$, including pre-war data, whenever the data is available. I begin with evaluating the model performance with CRRA utility, and compare the simulation accuracy between the global method and a third order local approximation method. Then, I focus on the benchmark model with recursive preferences, based on calibrated parameters reported in Table 1.1, and

extensively discuss its quantitative asset pricing implications. Appendix A.3 provides more details on the data sources.

1.4.1   Quantitative Evaluation the Solution Method

I begin with the model with CRRA utility and compare the performance of the global method used in this paper with a third order local approximation method. I argue that using a global method which allows for occasionally binding constraint is critical to quantify the asset pricing implications of financial frictions.

First, I focus on CRRA utility case at different levels of risk aversion, namely, $\gamma = 1$ (log utility), $\gamma = 2$ and $\gamma = 5$, which are commonly used in the macroeconomics literature. For each calibration experiment, I keep all the other parameters the same as in the benchmark calibration, summarized in Table 1.1, and I compare the same model with the global method and a third order local approximation method implemented by the Dynare++ package. For each experiment, the moments from different solution methods are listed in two adjacent columns. The results are reported in Table 1.2.

I make the following observations. First, even with CRRA utility at low levels of risk aversion, for instance, $\gamma = 1$ (log utility), or $\gamma = 2$, the probability of constrained region is still low, around $20 - 30\%$. When risk aversion increases, the probability of constrained region rapidly decreases. Second, it is surprising but interesting to see that the model's implied equity premium decreases with risk aversion, and this pattern behaves in the opposite direction as compared with the standard Lucas economy without frictions. In CRRA utility case, the IES, as the reciprocal of the risk aversion, decreases with risk aversion, and leads the average leverage ratio

to decrease dramatically, and in turn makes the volatility of shadow price of net worth to decrease rapidly. Since the dampening effect from the volatility of shadow price of net worth dominates the marginal rate of substitution of consumption, the first component in $\widetilde{M}_{t+1}$ as defined in equation (1.32), the augmented stochastic discount factor becomes less volatile and equity premium decreases. This experiment conveys the message that with CRRA utility, the financial frictions are not likely to have large asset pricing implications, because there is a strong trade-off between the contributions of two components in the augmented stochastic discount factor to the market price of risk. In Section 1.4.5, I will come back to this point and argue that when we incorporate recursive preferences with an IES larger than 1, the dampening effect discussed here is much weaker, and financial frictions generate significant impacts on asset prices.

It is also noteworthy that as the probability of constrained region decreases with risk aversion, the model's simulated moments suggested by the local approximation method have larger discrepancies with those of the global method. To further illustrate this point, in Table 1.3, I fix the risk aversion at $\gamma = 2$, and compare the model results for different bank asset divertible fractions $\theta = 0.2, 0.4$, and 0.8. As above, for each experiment, I keep all the other parameters the same as in the benchmark calibration, summarized in Table 1.1. Since the parameter θ directly affects the incentive for banks to divert by increasing its outside option value, the probability of constrained region is monotonically increasing with θ. As suggested by the global solution, in the high θ case ($\theta = 0.8$), the constraint is almost always binding, while in the low θ case ($\theta = 0.2$), the prob(binding) is as low as 0.03. Clearly, in the high θ case, the third order local approximation solution performs very well, and reports

very close moments to the global solution. However, in the low θ case in which the constraint rarely binds, the local approximation solution which imposes the assumption that the constraint always binds around steady state, greatly exaggerate the asset price volatilities, and therefore overstate the equity premium. In particular, in the low theta case ($\theta = 0.2$), the volatilities of price-dividend ratio and interbank interest rate are overestimated by more than twice and 10 times, respectively. And the equity premium is overestimated by more than 5 times.

I use the Den Haan and Marcet simulation accuracy test (1994) to compare the computation accuracy of the two solution methods. The basic idea is to construct the test statistic to measure the distance of simulated Euler equation error from zero. Under null hypothesis of exact numerical solution, the test statistic follows a χ^2 distribution. Additional details on constructing the test statistic are provided in Appendix A.5. Fig. 1.3 and Fig. 1.4 report the results for high θ case. In particular, they plot the empirical cumulative distribution of test statistic (based on 500 simulations of 1000 annual observations) versus its true χ^2 distribution under the null hypothesis for the global method and the local approximation method respectively. Both figures show that the empirical cumulative distributions are close to the true distribution under the null hypothesis. This implies that a third order local approximation method works well when $prob(binding)$ is high. Fig. 1.5 and Fig. 1.6 compare the results for low θ case. Fig. 1.5 shows that the global method still works well, however, the local approximation method fails in the sense that the empirical cumulative distribution of simulation accuracy test statistic is far from its true distribution under the null hypothesis.

In sum, in order to quantify the asset pricing implications of financial intermedi-

ary, we need to go to recursive preferences that which allow for a separation between the IES and risk aversion, and consequently permit both parameters to be simultaneously larger than 1, and use a global solution method which accounts for occasionally binding constraint.

1.4.2   Parameter Values

In this section, I discuss the parameter values in the benchmark calibration, which are summarized in Table 1.1.

Following Bansal and Yaron (2004), I set the relative rate of risk aversion, γ, to be 10, and the elasticity of intertemporal substitution, ψ, to be 1.5. I set the discount factor, β, to be 0.994 to match the level of risk-free interest rate for the household loans in the data.

In the log output growth process, the parameters μ_y and σ are calibrated to match the mean and volatility of the consumption growth in the data. Similarly, μ_d matches the average log dividend growth rate. Two additional parameters in the log dividend growth process, $\varphi > 1$ and $\varphi_d > 1$ allow me to match the overall volatility of dividends (which is larger than that of consumption in the data) and its correlation with consumption. The parameter φ, captures the loading of the log dividend growth process on the consumption growth shock. As in Abel (1999), φ can be interpreted as the leverage ratio on consumption growth.

There are three parameters for the financial sector: the annual liquidation/exit probability of banks, λ ; the transfer parameter for new banks, δ, and the fraction of bank asset divertible, θ . I set $\lambda = 0.12$, implying that banks survive for 8.33 years on average, similar to the number used in Gertler and Kiyotaki (2010).

There are no direct empirical counterparts in the data to pin down the rest two parameters, δ and θ. I choose these two parameters indirectly to match the following two targets: an average leverage ratio of 4 for economy-wide financial intermediary sector, and a standard deviation of interest rate spread of 0.55% per annum, consistent with that of TED spread.

Several considerations are noteworthy. First, as in Gerlter and Kiyotaki (2010), the model treats the entire intermediary sector as a group of identical institutions. Note that in the model the capital structure of the intermediary plays a central role in asset prices determination. It is important to match the leverage ratio because it affects how consumption shocks get magnified and the probability of being in the constrained versus unconstrained region. I follow the composition of the financial intermediary sector defined in Adrian, Moench and Shin (2011) to compute the leverage ratio of the aggregate financial intermediary from the Flow of Funds Table [9]. The average leverage ratio over the sample period $1945 - 2011$ is 3.67. I calibrate the parameters so that the model produces an average leverage ratio of 4.

Second, I calibrate the parameters based on the second moment, instead of the mean, of the TED spread. In the model, when the constraint is not binding, the interest rate spread is equal to zero, however, in the data, the TED spread is largely positive but smooth when the financial intermediary is well capitalized. Therefore, I match the volatility of TED spread. In the model, the volatility of interest rate spread is closely related to the probability of being in the constrained region. This moment provides a strong discipline on how much chance that the constraint is

[9] Publicly available at the online data library of Federal Reserve Board, http://www.federalreserve.gov/releases/z1/.

binding.

1.4.3 Basic Properties of the Model's Solution

In this section, I show the basic properties of the model's solution. In particular, I present the equilibrium prices, conditional volatilities of the market return and stochastic discount factor, and the equilibrium market return and risk-free interest rates, as functions of the state variable in this economy, i.e. the normalized debt level, $\mathbf{b}$.

Equilibrium Prices

Fig. 1.7 shows the equilibrium price-dividend ratio and marginal value of net worth as functions of normalized debt, $\mathbf{b}$, of the banking sector.

I make the following observations: First, I assume the realized consumption growth is bounded and satisfies the parameter restrictions as discussed in Section 1.3.3. This is important, otherwise, the equilibrium may not exist as shown in Ai, Bansal and Li (2012). In other words, if we assume that shocks are conditionally (log) Normal as in typical RBC models, there will be no equilibrium although the log-linearization method in Gertler and Kiyotaki (2010) still produces a solution. As a result of that assumption, the equilibrium level of debt will always be bounded between $\mathbf{b}_{MIN}$ and $\mathbf{b}_{MAX}$.

Second, the top panel shows that the equilibrium price-dividend ratio is monotonically decreasing in $\mathbf{b}$. As a comparison, the price-dividend ratio is a constant in the Lucas economy without frictions. The intermediary normalized debt level strongly affects asset prices through an adverse dynamic feedback: a negative fundamental

35

shock causes the losses of net worth and the accumulation of more debt, lowers the borrowing capacity of the intermediary today and into the future, and thus lowers the investment in risky asset market and depresses the stock prices, which further lowers the net worth. Importantly, note that the price-dividend ratio is low even when the constraint is not binding. The possibility of a binding constraint in the future lower the bank's capacity to invest in the stock today, and consequently lowers the market price of the stock. This implies that the amplification effect on risk premium is in action even in the unconstrained region, although the magnitude is smaller than in the constrained region.

With similar intuitions, the bottom panel shows that the marginal value of net worth is monotonically increasing in **b**. Note that in the standard Lucas economy it is a constant, and equal to 1.

Furthermore, the dashed line in bold in Fig. 1.7 depicts the the equilibrium prices in the constrained region. In the region where the constraint is binding, the price-dividend ratio decreases sharply and the marginal value of net worth increases sharply. This implies that the effects of intermediary debt on asset prices are non-linear and are especially large in bad times when the intermediary debt is high. That is, when the intermediary sector is extremely financially constrained, a negative fundamental shock is amplified to have large effects.

Conditional Volatility of Returns

Fig. 1.9 presents the conditional standard deviation of the market return (in log units) as a function of normalized debt level **b**. As the banking sector becomes more financially constrained, the conditional volatility of market return increases. Due to

the nonlinear sensitivity of price-dividend ratio with respect to the intermediary debt level as shown in the top panel of Fig. 1.7, the conditional volatility of market return increases more sharply when the banking sector is more levered. The increasing conditional volatility with the adversity of the state implies that the exposure of market return on the consumption shock (i.e. return beta) is increasing in bad times, which is one of the important channels to generate higher equity premium in bad times. As a comparison, the conditional variance of the return is constant in the Lucas economy without frictions since the price-dividend ratio is a constant.

The model endogenously produces several effects that have been emphasized in the empirical literature. First, the conditional variance in stock returns is persistent. The state variable, $\mathbf{b}$, is persistent, and it translates into a persistent conditional variance of stock returns. Second, the model endogenously generates a "leverage effect", that is, a consumption shock, as a negative innovation to market return, is a positive innovation to return volatility. Third, the conditional volatility of stock returns is countercyclical, and is higher when the intermediary net worth is low.

Conditional Volatility of Stochastic Discount Factor

Fig. 1.8 presents the conditional standard deviation of stochastic discount factor (in log units) as a function of normalized debt level $\mathbf{b}$. The conditional volatility of the stochastic discount factor determines the maximal Sharpe ratio. As the banking sector becomes more financially constrained, the conditional variance of stochastic discount factor increases. As discussed in Section 1.3.5, the stochastic discount factor depends not only on the aggregate consumption, but also on the shadow price of net worth. The second component increases more sharply when the leverage of the

37

intermediary sector is high as shown in the bottom panel of Fig. 1.7 and translates into higher volatility of the stochastic discount factor. The increasing conditional volatility of the stochastic discount factor with the adversity of the state implies that the market price of consumption shock is increasing in bad times. This is an important channel for generating countercyclical equity premium. As a comparison, the conditional volatility of the stochastic discount factor is constant in the Lucas economy without frictions since the shadow price of net worth is a constant at 1, and the consumption growth is homoscedastic.

Equity Premium

Fig. 1.10 presents the expected market return on levered dividend claim and two risk-free interest rates, i.e. the interbank interest rate, and the interest rate on household loans, as functions of the normalized debt level.

I define the equity premium as the spread between expected market return and interbank interest rate, $E_t\left(r_{m,t+1} - r_{f,t}^L\right)$, as it is determined by the covariance of the augmented stochastic discount factor $\widetilde{m}_{t+1}$ and the market return $r_{m,t+1}$. I make the following two observations. First, the equity premium increases with intermediary sector's normalized debt level, **b**. Second, the behavior of increases in the equity premium is asymmetric, namely, it increases much faster in the constrained region than in the unconstrained region. Both observations are explained by the fact that the equilibrium asset prices are more sensitive to the fundamental shocks when the intermediary net worth is low. As the financial intermediary sector becomes more financially constrained, both the exposure of market return to consumption shock (i.e. return beta) and the market price of the shock increase, and thus contribute to a

38

higher equity premium. And the equity premium increases faster when intermediary is extremely under-capitalized.

Interest Rate Spread

Fig. 1.10 shows two interest rates as functions of the normalized debt level, **b**. The interest rate on household loans, $r_{f,t}$, is a constant, and does not depend on the state variable **b**, as stated in Lemma 1. The interest rate on interbank loans $r_{f,t}^L$ is identical to $r_{f,t}$ when the constraint does not bind. However, when the constraint binds, the interest rate spread, denoted as $\left(r_{f,t}^L - r_{f,t}\right)$, becomes strictly positive, and increases with the state variable **b**. This pattern is consistent with the empirical evidence that in bad times when the banking sector is under-capitalized, the TED spread spikes.

In order to understand the response of interbank interest rate $r_{f,t}^L$, it is important to focus on the conditional mean of stochastic discount factor (in log units), i.e. $logE_t\left[\exp\left(m_{t+1} + \phi_{t+1}\right)\right]$, which is equal to $-r_{f,t}^L$ (up to a log-normal approximation).

$$\log E_t\left[\exp\left(m_{t+1} + \phi_{t+1}\right)\right] \tag{1.35}$$

$$= \quad E_t\left(m_{t+1}\right) + \frac{1}{2}var_t\left(m_{t+1}\right)$$

$$+ E_t\left(\phi_{t+1}\right) + \frac{1}{2}var_t\left(\phi_{t+1}\right) + cov_t\left(m_{t+1}, \phi_{t+1}\right).$$

In the i.i.d. consumption growth case, the first term $E_t\left(m_{t+1}\right) + \frac{1}{2}var_t\left(m_{t+1}\right)$ is constant. Fig. 1.11 plots a decomposition of the rest two terms in the conditional mean of stochastic discount factor, i.e. $E_t\left(\phi_{t+1}\right)$ and $\frac{1}{2}var_t\left(\phi_{t+1}\right) + cov_t\left(m_{t+1}, \phi_{t+1}\right)$. Clearly, there are two forces determining the response of the interbank interest rate.

First, the top panel shows that $E_t(\phi_{t+1})$ is decreasing in **b**. In the bad state with a negative shock which leads to a higher debt level, the net worth becomes more valuable today than the next period. Thus, the banks are very reluctant to lend net worth to others, instead they have strong incentive to borrow net worth and invest. Due to zero net supply, the market clearing condition drives up the inter-bank interest rate. Second, the bottom panel shows the second moment component $\frac{1}{2}var_t(\phi_{t+1}) + cov_t(m_{t+1}, \phi_{t+1})$ increases in **b**. The precautionary savings effect decreases the interbank interest rate. As shown by the magnitude of two panels, the first effect dominates the precautionary savings effect, and thus overall the interest rate on interbank loans increases in response to a negative fundamental shock, when the constraint is binding.

1.4.4 The Performance of Benchmark Model

I repeatedly simulate 1000 artificial samples from the model, each with 81 annual observations. For each data moment, I report the median value, $2.5, 5, 95$, and 97.5 percentiles, as well as the population value from a very long simulation (a long simulation of 10000 annual observations). The results are summarized in Table 1.4.

Designed by the calibration procedure, the model matches the aggregate consumption and dividend dynamics very well. It is noteworthy that by choosing two parameters, φ and φ_d, i.e. the loadings of aggregate dividend growth on consumption growth shock and its own shock, the model roughly matches the correlation between consumption and dividend growth, and the overall volatility of dividend process.

I use two asset pricing moments, namely, the leverage ratio and the volatility of interest rate spread to calibrate the model. Not surprisingly, the model matches

these two moments very well.

The model also performs very well in matching other asset pricing moments which are not targeted in the calibration. First, the model internally generates a persistent fluctuations of price-dividend ratio with first autocorrelation of 65%, even though the driving consumption growth process is i.i.d. Note that in the Lucas economy without frictions, price-dividend ratio is a constant. In this economy, intermediary's debt level is a state variable that affects asset prices, and thus price-dividend ratio inherits its positive serial correlation.

Second, the model produces a high equity premium (in log units) of 4.1%, a significant share (78%) of the equity premium observed in the data, and a stock market volatility of 16.5%, only slightly lower than a volatility of 19.8% in the data.

However, we also notice that there are some discrepancies between the model implied moments with the data. The model implied average interest rate spread is 0.15%, lower than 0.64% in the data. As I argued in Section 1.4.2, the model predicts zero interest rate spread when constraint does not bind, however, in the data, the TED spread is largely always positive even when the banking sector is well-capitalized. What's more, we only have TED spread for a short sample (1986 − 2011), therefore, the average spread may be driven high due to the inclusion of the recent financial crisis period when the TED spread was enormously high. Another discrepancy is that the model understates the volatility of the log price-dividend ratio. In the model, the standard deviation of the log price-dividend ratio is 0.12, as compared with 0.45 in the annual data. Historical stock prices display low-frequency variation relative to cash flow, which is not captured in the model. The historical standard deviation of log price-dividend ratio is this high in part because stock prices

41

were persistently high at the end of the sample period. In Bansal and Yaron (2004), the sample period ends at 1998, they obtain a lower standard deviation of 0.29 in the data, but still somewhat higher than in the model here.

Overall speaking, Table 1.4 suggests that the model performs relatively well to match both cash flow dynamics and asset pricing moments for U.S. data, given the driving force is an i.i.d. process. I could introduce a predictable component in expected consumption and dividend growth to further improve the persistence and standard volatility of price-dividend ratio.

1.4.5 Comparative Statics

To shed more light on the economic mechanisms in the model, Table 1.6 conducts five comparative statics by varying the key parameters: (1) risk aversion decreased from 10 in the benchmark calibration to 5; (2) the volatility of consumption growth, changed from 2.20% to 1.56% per annum to match the post-World War II aggregate consumption data; (3) IES ψ changed from 1.5 to 0.5; (4) the fraction of banks forced to liquidate in each period, λ, from 0.12 to 0.16; (5) the fraction of bank asset divertible, θ, from 0.4 to 0.6. In each experiment, except for the parameter being perturbed, all the other parameters are kept the same as in the benchmark calibration. In Table 1.6, all moments are reported from a very long simulation of data from the model at the annual frequency. The first column corresponding to the benchmark calibration as reported in Table 1.6.

Different Risk Aversion γ and Consumption Volatility σ

The first two variations consider changes in the risk aversion γ and the consumption volatility σ, the moments of which are reported in the second and third column, respectively.

Relative to the benchmark calibration of $\gamma = 10$, setting $\gamma = 5$ decreases the equity premium and the probability of entering the constrained region. When lower risk aversion, the intermediary sector is less conservative, and is willing to take a more risky portfolio, i.e. it has a higher average leverage ratio. Hence, the same consumption volatility is translated into a greater volatility of net worth, and the economy is more likely to hit a binding constraint state. This is a risk-taking effect. There is also a general equilibrium effect reinforcing the risk-taking effect. Due to a lower risk aversion, the market price of risk falls and causes the intermediary to be compensated less per unit of risk, and therefore, the intermediary sector on average retains less earnings, and has a lower average net worth level, which in turns leads the constraint to bind more often. Furthermore, we also observe that lower risk aversion increases the interest rate spread due to the dampening of the precautionary saving effect.

When lowering the consumption volatility to post-World War II level, it is intuitive to observe that the equity premium, the volatility of net worth and stock return all decreases. However, a surprising result in the case is the probability of constrained region increases with a lower level of fundamental shock. The reason is that the increasing price-consumption ratio (i.e. increasing the right hand side of the constraint) makes the intermediary has more incentive to divert bank assets.

Different IES ψ

Relative to an $IES = 1.5$, setting $IES = 0.5$ has important asset pricing implications. The moments of this experiment are presented in the fourth column of Table 1.6. First, lower IES leads to a higher risk-free interest rate for the loans from the household sector. Second, an IES smaller than 1 implies a much lower average leverage ratio, meaning that the banking sector holds a less risky portfolio. As expected, the same fundamental volatility is translated to a smaller volatility of net worth and a lower equity premium. Surprisingly, the probability of constrained region is increasing. This is because the general equilibrium effect dominates the risk-taking effect: the price of risk falls and the banking sector is compensated less per unit of risk, and hence it has a lower average net worth level, which in turn leads to the constraint to bind more frequently.

Different Liquidation/Exit Probability λ

In the fifth column of Table 1.6, I increase λ, the fraction of banks forced to liquidate each period, from 0.12 to 0.16. This implies that the average survival duration decreases from 8.33 years to 6.25 years. As we can see from the Table 1.6, since every period there is a larger fraction of net worth paid back to the household sector, the banking sector tends to be more financially constrained, and have a higher average leverage ratio. Following the same "risk taking" story as stated above, higher risky position is translated into a higher volatility of the net worth and a higher equity premium. Higher volatility of the net worth leads the economy to enter the constrained region more often. This effect is also reinforced by the lower average net worth of the banking sector.

It is noteworthy that this experiment also reflects a amplification and persistence trade-off. With a higher λ, that is, a larger fraction of aggregation net worth paid back to the household sector each period, the equilibrium premium increases, however, the price-dividend ratio is less persistence, translated by a less persistent net worth process. This case is expected to feature a less return predictability.

Different Bank Asset Divertible Fraction θ

In the experiment shown in the last column, I increase the parameter θ, which dictates the fraction of bank asset divertible, from 0.4 to 0.6. This mainly affects the average leverage ratio of the banking sector. As the banking sector can divert a larger fraction of bank assets, the leverage constraint allows a much lower average leverage ratio. As a result, the volatilities of net worth and of shadow price of net worth decrease, which leads to a decrease in equity premium and stock market volatility. Despite of lower average leverage, the probability of constrained region is still larger than in the benchmark case. This is because the right hand side threshold of the constraint increases, which makes it to bind more frequently.

Conditional Moments

Table 1.5 shows the model implied moments conditional on the leverage constraint being binding or not. Each panel of the table corresponds to a comparative statics experiment discussed above. As shown in the table, for all cases, the leverage ratio, Sharpe ratio and interest rate spread conditional on the constraint being binding is higher than those moments in the unconstrained region.

1.5 Additional Asset Pricing Implications

1.5.1 Variance Decomposition of Price-Dividend Ratio

In this section, I replicate the variance decomposition of price-dividend ratio as in Cochrane (1992) and Campbell and Cochrane (1999). Table 1.7 presents the estimation results. Consistent with previous research, the estimates in the data find that more than 100 percent of the price-dividend ratio variance is attributed to expected return variation. A high price-dividend ratio signals a decline in subsequent real dividends, so it must signal a large decline in expected returns. The model is consistent with this feature in the data. Almost all (over 90%) the variation in price-dividend ratio is due to changing expected returns. This evidence repeats the intuition discussed above: the expected dividend growth is our model is constant over time, however, a negative fundamental shock, which causes the loss of net worth (or the accumulation of net debt), provides an endogenous channel of a discount rate shock, that greatly and persistently lowers the expected return.

An interesting point of comparison for my result is to the habit model in Cochrane and Campbell (1999). In that model, they modify the utility function of a representative investor to exhibit time-varying risk aversion, and therefore a negative fundamental shock is a discount rate shock by construction. Differently, I work on CRRA utility and recursive preferences as a more general utility function to disentangle risk aversion with IES, but generate an endogenous channel of time-varying equity premium as a function of the frictions in the economy.

46

1.5.2 Return Predictability

In this section, I provide the valuation on model's ability to endogenously generate return predictability. The left panel of Table 1.8 reports the results on predictability of mutli-period excess returns by the log price-dividend ratio. Consistent with evidence in earlier papers, In the data, the R^2 rises with maturity, from 4% at one year horizon to about 31% at the five year horizon. The model-implied predictability of equity return is somewhat lower. The slope coefficients in the multi-horizon return projections implied by the model are of the right sign and magnitude compared to those in the data.

The right panel of Table 1.8 shows evidence on predictability of multi-period excess returns by the log leverage ratio of the aggregate financial intermediary sector. In the data, the R^2 rises with maturity, from 9% at one year horizon to about 28% at the five year horizon. The model-implied predictability of equity return is comparable to those in the data, and the slope coefficients in the multi-horizon return projections implied by the model are of the right sign as those in the data. In sum, the empirical evidence presented in this section shows that the leverage constraint channel endogenously generates significant variation in equity premium.

1.5.3 Correlation Structure of Leverage Ratio

The economic mechanism in the model has strong implications for the correlation of leverage ratio with various asset market moments. In Table 1.9, I reports the correlations of leverage growth with price-dividend ratio, excess stock return, stock market integrated volatility and financial asset growth of the intermediary sector.

In the literature, there are some discussions about the cyclicality of leverage ra-

47

tio. In particular, Adrian and Shin (2010) documents that the leverage ratio of security broker-dealers is highly procyclical, by showing that leverage ratio of this particular type of financial intermediary, constructed from Flows and Fund Table in U.S., is positively correlated with its asset growth. He, Khang and Krishnamurthy (2010) shows that there is large heterogeneity among different types of financial intermediary. In particular, they document that in the period of 2007q1 to 2009q1, the broker-dealers shed assets, consistent with Adrian and Shin (2010)$'s$ evidence, however, the commercial banking sector increased asset holdings over this period significantly, and therefore, increased its leverage ratio. In this paper, the model intermediary sector is meant to capture the entire financial intermediary sector. Thus, I follow the definition in Adrian Moench and Shin (2011) to construct the leverage ratio of aggregate intermediary sector, with a coverage consistent with the model. The details about data construction are shown in the Appendix A.3. I find, in the data, the leverage growth of aggregate intermediary sector is negatively correlated with its asset growth, which suggests the leverage ratio is countercyclical. This is consistent with the model.

The data also suggests that in bad times when leverage ratio increases, stock price is low, the stock return decreases in the contemporaneous period, and stock market volatility increases. The benchmark model fits these correlation patterns in the data well.

1.5.4 Correlation Structure of Interest Rate Spread

As a distinct prediction, the model draws strong implications for the correlation of interest rate spread between interbank and household loans with price-dividend ratio,

price-earnings ratio and the stock market volatility. Fig. 1.1 shows the periods of significant widening of TED spread coincide with those of dramatic increases in stock market volatility, and large decreases in price-dividend and price-earning ratios. In Table 1.10, I confirm these correlations. As I discussed in Section 1.4.3, the model is consistent with the correlation patterns in the data well. In the model, the interest rate spread, as a measure of the tightness of the credit constraint, spikes when the intermediary sector are extremely financial constrained. The banks are constrained, and do not have liquidity to lend out to others, thus, the market clearing drives up the interest rate. On the other hand, low intermediary net worth depresses the stock market, and increases the stock market volatility, as we discussed above. These model predictions explain the empirical evidence very well.

1.5.5 Backward Looking Regression

In this section, I follow Bansal, Kiku and Yaron (2012) to evaluate the model by examining the link between price-dividend ratio and consumption growth. I replicate their empirical procedure and run the following regression:

$$p_{t+1} - d_{t+1} = \alpha_0 + \sum_{j=1}^{L} \alpha_j \Delta c_{t+1-j} + u_{t+1}.$$

In the actual data and in the simulated data, I regress the log of price dividend ratio on L lags ($L = 1, 2, ..., 5$) of consumption growth. In the data, at all lag-lengths, predictability of the price dividend ratio by lagged consumption growth is close to zero. However, in the model, price dividend ratio predictability by lagged consumption has an R^2 of 42%, see Fig. . This is not surprising as prices in this model are driven primarily by the net worth, and hence, by movements in the lagged

49

consumption, and a reduction in growth rates causes the loss in net worth, and thus increases the equity premium, and provide an endogenous positive discount shock, leading to a fall in current price-dividend ratio. This feature of the model is similar to the habit model in Campbell and Cochrane (1999). Both models are backward looking, in the sense that backward consumption plays an important role in determining current prices. The empirical evidence presented in this section proposes a challenge for asset pricing models with financial intermediary.

1.6 Conclusion

In this study, I show financial frictions are important for understanding a wide variety of dynamic asset pricing phenomena. I build a financial intermediary sector with a leverage constraint à la Gertler and Kiyotaki (2010) into a standard endowment economy with recursive preferences and an independently and identically distributed consumption growth process. Quantitatively, the model generates a high and countercyclical equity premium, a low and smooth risk-free interest rate and a procyclical and persistent variation of price-dividend ratio. As a distinct prediction from the model, when the intermediary sector is financially constrained, the interest rate spread between interbank and household loans spikes, stock market valuation ratio falls and the market volatility rises dramatically. This pattern is consistent with the empirical evidence that high TED spread coincides with low stock price and high stock market volatility, which I document in the paper.

I use a recursive method to construct the global solution, and argue that accounting for occasionally binding constraint is important for quantifying the asset pricing implications through a careful quantitative evaluation. A local approximation

method assuming the constraint always binds around steady state tend to greatly exaggerate the asset price volatilities and equity premium.

Several extensions are on my research agenda. First, I can introduce a predictable component of expected growth into the consumption dynamics, i.e. long-run risk (Bansal and Yaron, 2004). In this context, the long-run risk is both a growth rate shock and a discount rate shock, because it causes the loss of intermediary capital and therefore endogenously affects the expected return. Second, it is interesting to study the asset pricing implications of financial intermediary in a production economy, in which consumption and investment decisions are endogenous. In this framework, financial frictions affect not only asset prices, but also real activities, and the leverage constraint is potentially an endogenous channel to generate long-run risks and rare disasters in consumption growth, and thus provides some interesting insights in the context of a production based asset pricing model.

Table 1.1: Parameter Values in the Benchmark Calibration at the Annual Frequency

Notation	Parameters	Value
Recursive preferences		
β	Time discount factor	0.994
γ	Relative risk aversion	10
ψ	The elasticity of intertemporal substitution	1.5
Financial Sector		
λ	Liquidation/exit probability of banks	0.12
θ	Fraction of bank assets divertible	0.40
δ	Transfer to entering banks	0.02
Consumption and Dividend Dynamics		
μ_c	Unconditional mean of consumption growth	0.18
σ	Unconditional volatility of consumption growth	0.022
μ_d	Unconditional mean of dividend growth	0.105
φ	Dividend growth's loading on consumption growth shock	3
φ_d	Dividend growth's loading on dividend growth shock	4

This table reports the parameter values used for benchmark calibration at the annual frequency.

Table 1.2: CRRA Utility: Different RRA γ

	Data	$\gamma = 1$		$\gamma = 2$		$\gamma = 5$	
		Global	Local	Global	Local	Global	Local
Avg.Leverage	3.67	4.38	2.72	3.76	3.72	2.51	2.60
$E[log(\hat{n})]$	-	2.49	2.70	2.02	2.02	1.61	1.58
$E(r_m - r_f^L)$	4.58	1.75	1.92	1.09	1.40	0.99	1.57
$E(r_f^L - r_f)$	0.64	0.44	-0.05	0.56	0.37	0.12	-0.48
$\sigma[log(\hat{n})]$	-	0.30	0.44	0.23	0.26	0.114	0.16
$\sigma(p - d)]$	0.45	0.07	0.07	0.05	0.06	0.02	0.05
$\sigma(r_m)$	19.79	17.34	17.55	16.15	16.82	14.06	15.24
$\sigma(r_f^L)$	0.55	0.98	2.51	1.09	1.86	0.49	1.49
$prob(binding)$		0.28		0.35		0.11	

This table presents selected moments implied by the model with CRRA utility at different risk aversion parameters. Other parameters are kept the same as in the benchmark calibration in Table 1.1. All the moments reported are computed from a very long sample of simulated data. In columns "global", the moments are based on the global solution. In columns "Local", the moments are based on a third order local approximation method implemented using dynare++ package. Means and volatilities of returns and growth rates are expressed in percentage terms.

Table 1.3: CRRA Utility: Different Bank Assets Divertible Fraction θ

	Data	Model					
$\gamma = 2$		$\theta = 0.2$		$\theta = 0.4$		$\theta = 0.8$	
		Global	Local	Global	Local	Global	Local
$E[log(\hat{n})]$	-	1.96	3.16	2.02	2.02	2.08	2.08
$E(r_m - r_f^L)$	4.58	0.96	5.54	1.09	1.40	0.44	0.35
$E(r_f^L - r_f)$	0.64	0.05	-3.85	0.56	0.37	2.30	2.32
$\sigma[log(\hat{n})]$	-	0.31	1.39	0.23	0.26	0.11	0.11
$\sigma(p - d)$	0.45	0.06	0.12	0.05	0.06	0.03	0.03
$\sigma(r_m)$	19.79	16.21	23.65	16.15	16.82	14.67	14.55
$\sigma(r_f^L)$	0.55	0.36	5.77	1.09	1.86	0.71	0.71
$prob(binding)$		0.03		0.35		1.00	

This table presents selected moments implied by the model with CRRA utility of risk aversion parameter of 2, at different fractions of bank assets divertible, θ. Other parameters are kept the same as in the benchmark calibration in Table 1.1. All the moments reported are computed from a very long sample of simulated data. In columns "global", the moments are based on the global solution. In columns "Local", the moments are based on a third order local approximation method implemented using dynare++ package. Means and volatilities of returns and growth rates are expressed in percentage terms.

Table 1.4: Dynamics of Growth Rates and Prices Based on Benchmark Calibration

	Data	Benchmark Model					
	Estimate	Median	2.50%	5%	95%	97.50%	Pop
$E(\Delta c)$	1.83	1.78	1.30	1.39	2.18	2.30	1.80
$\sigma(\Delta c)$	2.19	2.18	1.92	1.96	2.41	2.44	2.22
$E(\Delta d)$	1.08	1.09	-1.22	-0.80	3.02	3.39	1.27
$\sigma(\Delta d)$	10.98	10.90	9.27	9.69	12.35	12.54	10.93
$corr(\Delta c, \Delta d)$	0.56	0.60	0.45	0.47	0.71	0.72	0.59
$avg.leverage$	3.67	4.00	3.61	3.66	4.51	4.69	4.02
$\sigma(leverage)$	1.65	0.93	0.50	0.54	2.32	2.78	1.22
$E(r_m - r_f)$	5.22	4.04	1.25	1.60	6.62	7.06	4.07
$E(r_m - r_f^L)$	4.58	3.86	1.01	1.31	6.57	6.96	3.90
$\sigma(r_m)$	19.79	16.54	14.00	14.42	18.75	19.41	16.69
$E(p - d)$	3.38	3.12	2.88	2.92	3.32	3.36	3.12
$\sigma(p - d)$	0.45	0.12	0.08	0.08	0.15	0.16	0.12
$AC1(p - d)$	0.86	0.62	0.43	0.47	0.75	0.77	0.65
$E(r_f^L - r_f)$	0.64	0.15	0.03	0.03	0.40	0.45	0.17
$\sigma(r_f^L)$	0.55	0.52	0.13	0.18	0.94	1.09	0.58

This table presents descriptive statistics for aggregate consumption growth, dividends, prices, the interest rate spread (i.e. the spread between interest rates for interbank and household loans). The data are real, sampled at an annual frequency and cover the sample period from 1930 to 2011, whenever the data are available. The sample period for leverage ratio is from 1945 to 2011. The sample period for interbank interest rate is from 1986 to 2011. The "Model" panel presents the corresponding moments implied by the model. The first five columns in the right panel represent percentiles of finite sample Monte-Carlo distributions. Population values (Pop) are computed from a very long sample of simulated data. Means and volatilities of returns and growth rates are expressed in percentage terms.

Table 1.5: Model Implied Conditional Moments

	Unconstrained	Constrained
Panel A: Benchmark		
Probability	0.86	0.14
Leverage Ratio	3.69	6.02
Sharpe Ratio	0.21	0.46
Interest Rate Spread	0.00	1.21
Panel B: $\gamma = 5$		
Probability	0.77	0.23
Leverage Ratio	3.78	6.04
Sharpe Ratio	0.12	0.35
Interest Rate Spread	0.00	1.47
Panel C: $\sigma = 0.0156$		
Probability	0.58	0.42
Leverage Ratio	3.76	4.86
Sharpe Ratio	0.16	0.32
Interest Rate Spread	0.00	1.25
Panel D: $\lambda = 0.16$		
Probability	0.72	0.28
Leverage Ratio	4.22	6.34
Sharpe Ratio	0.23	0.45
Interest Rate Spread	0.00	1.55
Panel E: $\theta = 0.6$		
Probability	0.33	0.67
Leverage Ratio	2.92	3.65
Sharpe Ratio	0.19	0.31
Interest Rate Spread	0.00	1.34

This table presents selected moments implied by the model conditional on being in the unconstrained versus constrained regions. Each panel corresponds to a comparative statics experiment in Table 1.6. All the moments reported are computed from a very long sample of simulated data. Means and volatilities of returns and growth rates are expressed in percentage terms.

Table 1.6: Comparative Statics

	Benchmark	$\gamma = 5$	$\sigma = 0.0156$	IES $= 0.5$	$\lambda = 0.16$	$\theta = 0.6$
avg.leverage	4.02	4.30	4.23	3.49	4.81	3.41
$E[log(\hat{n})]$	2.86	2.82	2.88	2.15	2.38	2.69
$\sigma[log(\hat{n})]$	0.25	0.28	0.19	0.20	0.26	0.19
$AC1(p-d)$	0.65	0.65	0.63	0.70	0.55	0.67
$\sigma(\mu)$	0.47	0.53	0.29	0.23	0.56	0.33
$\sigma(\phi)$	0.14	0.16	0.11	0.10	0.17	0.10
$E(r_m - r_f)$	4.07	2.96	2.64	3.33	4.83	4.19
$E(r_m - r_f^L)$	3.90	2.62	2.11	3.03	4.39	3.29
$\sigma(r_m)$	16.69	16.96	11.56	15.67	16.47	15.78
$E(r_f^L - r_f)$	0.17	0.34	0.53	0.30	0.43	0.90
$\sigma(r_f^L)$	0.58	0.87	0.93	0.78	0.98	1.01
$prob(binding)$	0.14	0.23	0.42	0.23	0.28	0.67
$amp.eff.$	2.68	3.61	2.89	2.09	3.03	2.26

This table presents selected moments implied by the model for comparative statics experiments. The first column reports the moments based on benchmark calibration. Each of the rest 5 columns report the moments by changing one parameter, while keeping all the other parameters the same as in the benchmark calibration. All the moments reported are computed from a very long sample of simulated data. Means and volatilities of returns and growth rates are expressed in percentage terms. $\sigma(\mu)$ denotes the volatility of shadow value of net worth. $\sigma(\phi)$ denotes the volatility of $log(\Phi)$, defined in equation (1.33).

Table 1.7: Variance Decomposition of Price-Dividend Ratio

Source	Data	S.E.	Model
Dividends	-6%	(31%)	2%
Returns	108%	(42%)	90.45%

This table reports the percentage of $var(p - d)$ accounted for by returns and dividend growth rates:

$$100 \sum_{j=1}^{15} \Omega^j \frac{cov_t (p_t - d_t, x_{t+j})}{var_t (p_t - d_t)}$$

$x = -r$ and Δd, respectively, and $\Omega = \frac{1}{1+E(r)}$. The "model" column is based on a very long simulation of annual observations from the model with benchmark calibration.

Table 1.8: Return Predictability

Predictor	p-d			log leverage		
	Data	(S.E.)	Model	Data	(S.E.)	Model
$B(1)$	-0.09	(0.07)	-0.27	0.09	(0.05)	0.05
$B(3)$	-0.27	(0.16)	-0.43	0.22	(0.09)	0.09
$B(5)$	-0.43	(0.21)	-0.66	0.28	(0.11)	0.10
$R^2(1)$	0.04	(0.04)	0.05	0.02	(0.02)	0.06
$R^2(3)$	0.19	(0.13)	0.09	0.09	(0.04)	0.10
$R^2(5)$	0.31	(0.15)	0.15	0.11	(0.05)	0.16

This table provides evidence on predictability of future excess return by log price-dividend ratio, and log leverage ratio of the aggregate intermediary sector. The entries correspond to regressing

$$r_{t+1}^e + r_{t+2}^e + ... + r_{t+j}^e = \alpha(j) + B(j)x_t + v_{t+j}$$

where r_{t+1}^e is the excess return, j denotes the forecast horizon in years. x_t denotes log price-dividend ratio for the left panel, and denotes log leverage ratio for the right panel. The entries for the model are based on 1000 simulations each with 81 annual observations. Standard errors are Newey-West corrected using 10 lags.

Table 1.9: Correlations of Aggregate Leverage Ratio and Asset Prices

	Data	S.E.	Model
$corr(\Delta lev, p-d)$	-0.71	0.20	-0.44
$corr(\Delta lev, r_m - r_f^L)$	-0.75	0.19	-0.93
$corr(\Delta lev, IV)$	0.38	0.16	-
$corr(\Delta lev, asset-growth)$	-0.60	0.16	-

This table shows the correlations between log leverage growth of aggregate intermediary sector with asset market moments, including price-dividend ratio, stock excess return, stock market integrated volatility and financial asset growth in the aggregate intermediary sector. The data are sampled at the annual frequency, ranging from 1945 to 2011. Data constructions are described in the Appendix A.3. The numbers reported in "S.E." column are based on GMM Newey-West standard errors. The corresponding model implied correlations are reported whenever applicable, based on a very long sample of simulated data.

Table 1.10: Correlations of Interest Rate Spread and Asset Prices

	Data	S.E.	Model
$corr(r_f^L - r_f, \Delta lev)$	0.11	0.06	0.46
$corr(r_f^L - r_f, p-d)$	-0.42	0.20	-0.77
$corr(r_f^L - r_f, IV)$	0.32	0.15	0.40

This table shows the correlations between TED spread with asset market moments, including log leverage growth of the intermediary sector, log price-dividend ratio, log price-earnings ratio and stock market integrated volatility. The data are sampled at the annual frequency, ranging from 1986 to 2011. Data constructions are described in the Appendix A.3. The numbers reported in "S.E." column are based on GMM Newey-West standard errors. The corresponding model implied correlations are reported whenever applicable, based on a very long sample of simulated data.
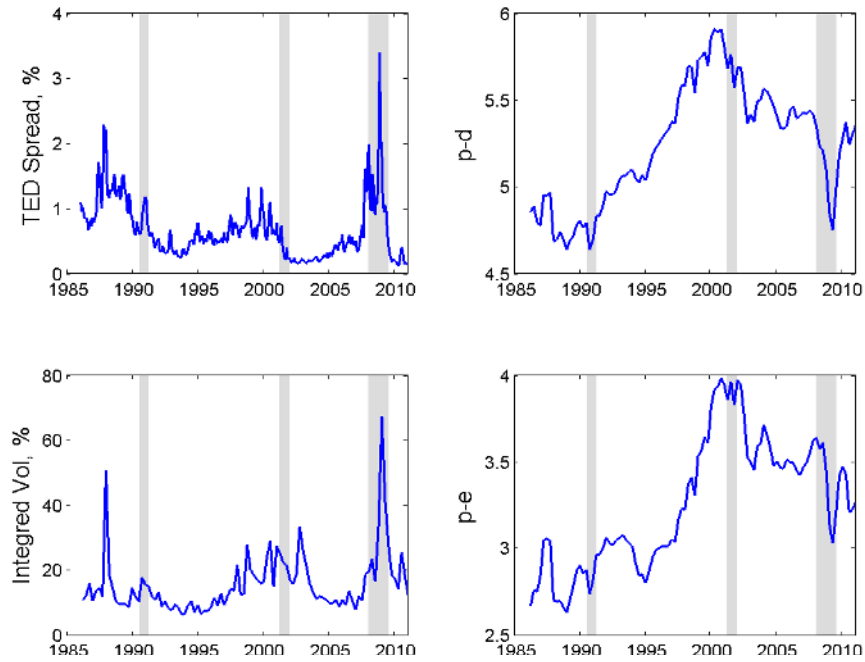
FIGURE 1.1: TED Spread, p-d ratio, p-e ratio and Integrated Volatility

This figure plots TED spread, log p-d ratio, log p-e ratio and integrated volatility over the sample period 1986 to 2011. TED spread and integrated volatility are in annualized percentage. Shaded areas refer to NBER dated recessions. Data constructions are described in Appendix A.3.

61
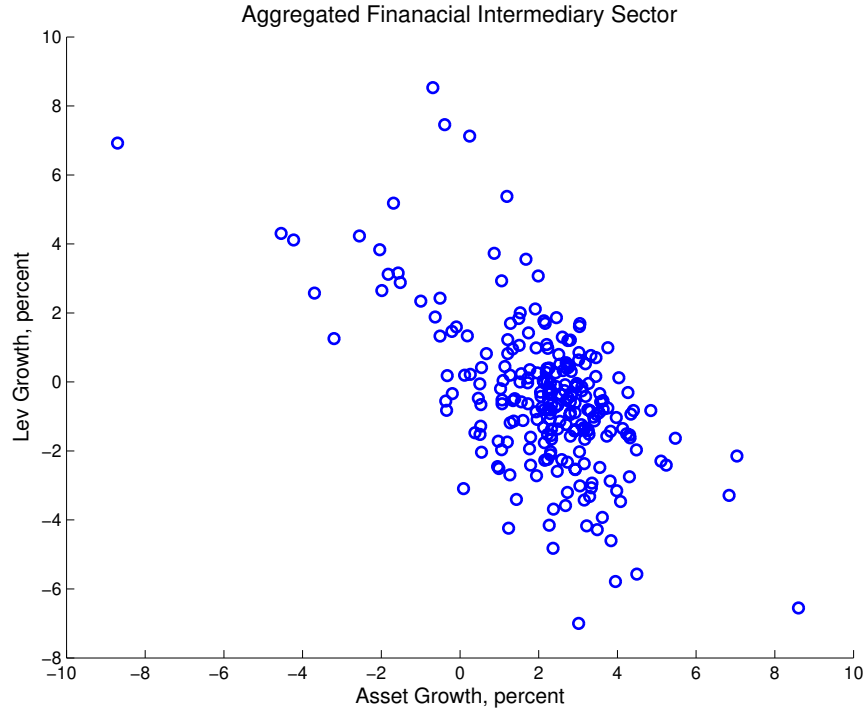
Figure 1.2: Leverage Ratio of Aggregate Financial Intermediary Sector

This figure shows scatter plots of the growth rate of financial assets (horizontal axis) versus the growth rate of leverage ratio (vertical axis) of the aggregate financial intermediary sector. The sample is at quarterly frequency, ranging from 1952q2 to 2011q4. Both axes are measured in percentage. The constructions of the total financial assets and leverage ratio of the aggregate financial intermediary sector are described in Appendix A.3.
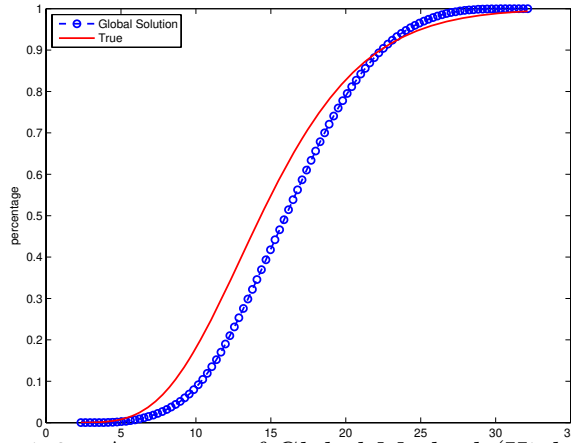
FIGURE 1.3:   Accuracy of Global Method (High θ Case)

This figure shows the cumulative distribution function of the simulation accuracy test statistics suggested by Den Haan and Marcet (1994) and the corresponding χ^2 distribution under the null hypothesis. The realizations of the test statistics are based on 500 simulation paths, each with 1000 annual observations. The simulations are based on the global solution with high θ case ($\theta = 0.8$).
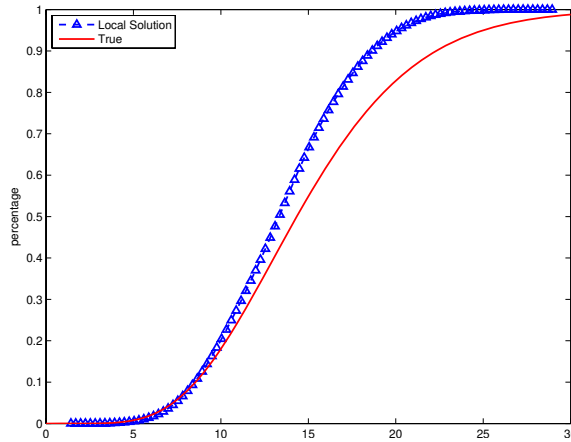
FIGURE 1.4:   Accuracy of Local Approximation Method (High θ Case)

This figure shows the cumulative distribution function of the simulation accuracy test statistics suggested by Den Haan and Marcet (1994) and the corresponding χ^2 distribution under the null hypothesis. The realizations of the test statistics are based on 500 simulation paths, each with 1000 annual observations. The simulations are based on the third order local approximation solution with high θ case ($\theta = 0.8$).

63
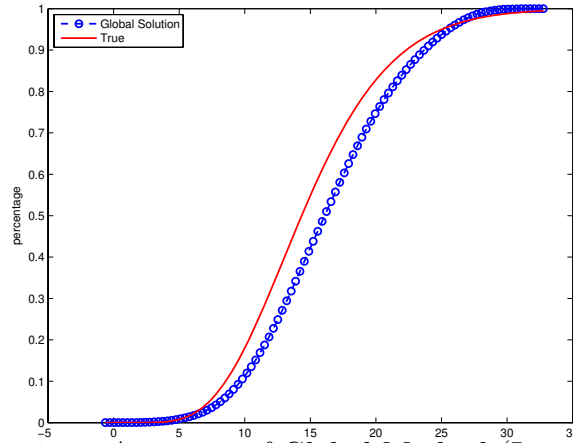
FIGURE 1.5: Accuracy of Global Method (Low θ Case)

This figure shows the cumulative distribution function of the simulation accuracy test statistics suggested by Den Haan and Marcet (1994) and the corresponding χ^2 distribution under the null hypothesis. The realizations of the test statistics are based on 500 simulation paths, each with 1000 annual observations. The simulations are based on the global solution with low θ case ($\theta = 0.2$).
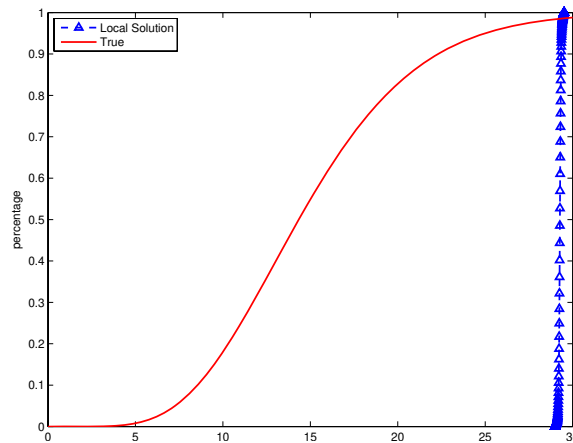
FIGURE 1.6: Accuracy of Local Approximation Method (Low θ Case)

This figure shows the cumulative distribution function of the simulation accuracy test statistic suggested by Den Haan and Marcet (1994) and the corresponding χ^2 distribution under the null hypothesis. The realizations of the test statistics are based on 500 simulation paths, each with 1000 annual observations. The simulations are based on the third order local approximation solution with low θ case ($\theta = 0.2$).

64
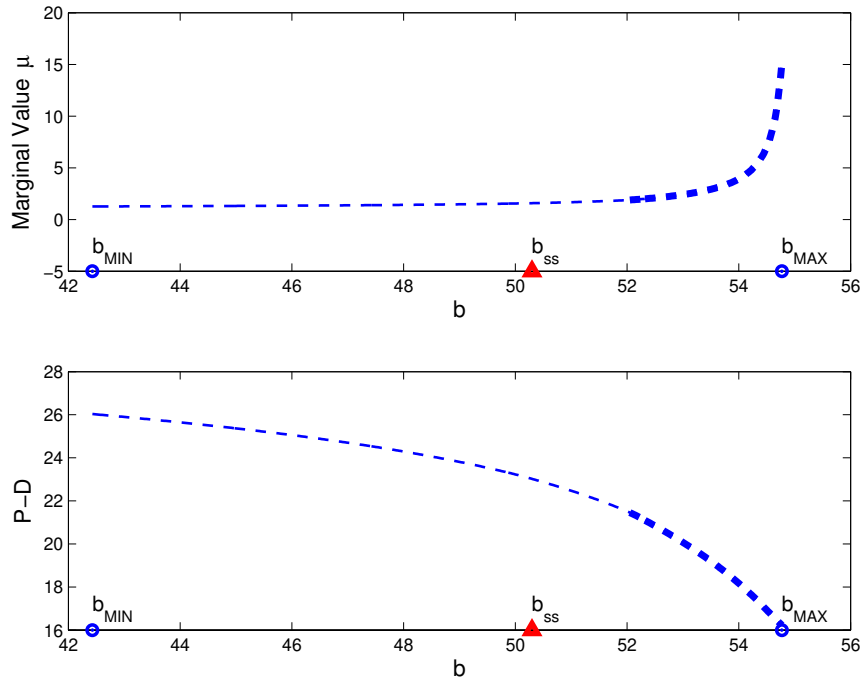
FIGURE 1.7: Equilibrium Prices as Functions of Normalized Debt Level, **b**

This figure shows the p-d ratio on aggregate dividend claim and the shadow price of net worth as functions of the state variable b. b_{MIN} and b_{MAX} denote the boundaries of the equilibrium debt level. b_{ss} denotes the average debt level in this economy, suggested by a long simulation from the model. The part of curves highlighted in bold denotes the region at which the constraint is binding. The parameters are based on the benchmark calibration summarized in Table 1.1.

65
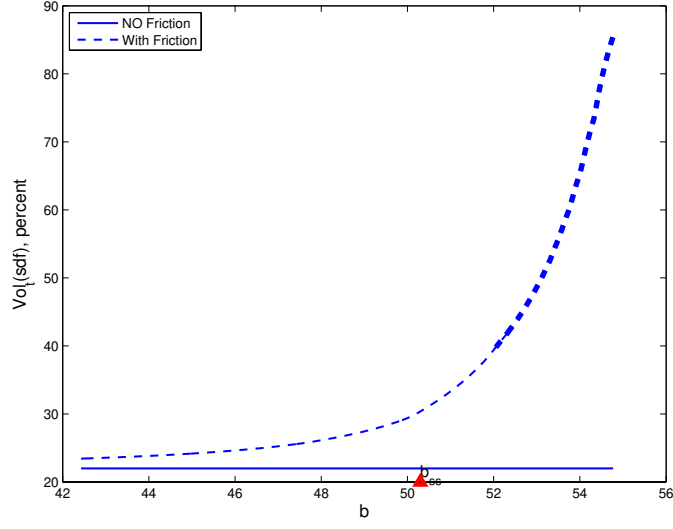
FIGURE 1.8: Conditional Volatility of Log SDF as a Function of Normalized Debt Level, **b**

This figure shows conditional volatilities of stochastic discount factor with and without frictions as functions of the state variable b. b_{ss} denotes the average debt level in this economy, suggested by a long simulation from the model. The part of curves highlighted in bold denotes the region at which the constraint is binding. The vertical axis is measured in annualized percentage. The parameters are based on the benchmark calibration summarized in Table 1.1.
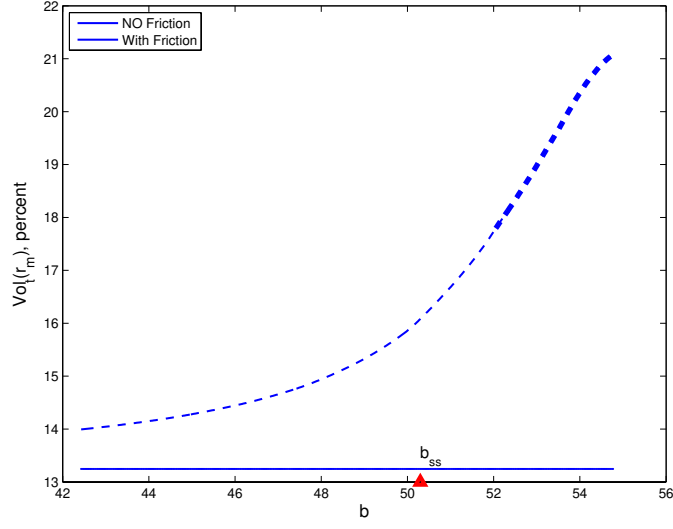
FIGURE 1.9: Conditional Volatility of Return as a Function of Normalized Debt Level, **b**

This figure shows conditional volatilities of market return with and without frictions as functions of the state variable b. b_{ss} denotes the average debt level in this economy, suggested by a long simulation from the model. The part of curves highlighted in bold denote the region at which the constraint is binding. The vertical axis is measured in annualized percentage. The parameters are based on the benchmark calibration summarized in Table 1.1.
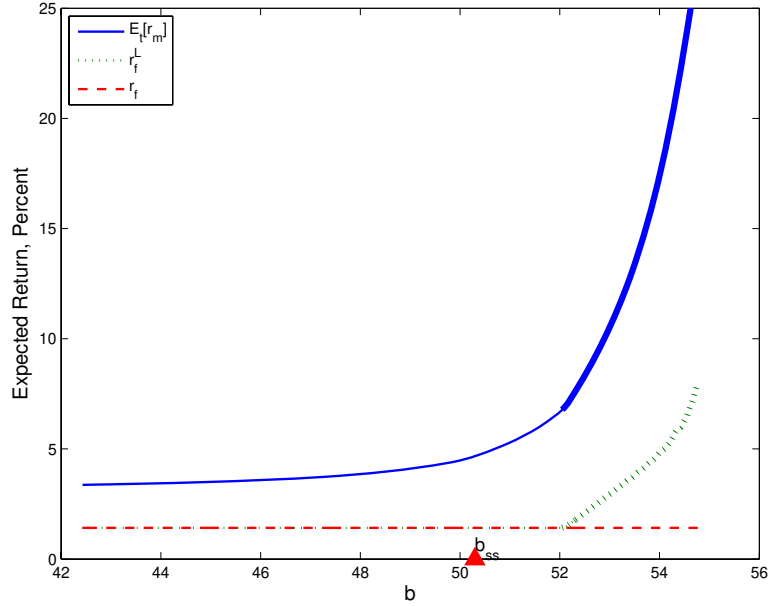
FIGURE 1.10: Expected Returns as Functions of Normalized Debt Level, **b**

This figure shows the expected market return, interbank interest rate, and the interest rate on household loans as functions of the state variable b. b_{ss} denotes the average debt level in this economy, suggested by a long simulation from the model. The part of curves highlighted in bold denotes the region at which the constraint is binding. The vertical axis is measured in annualized percentage. The parameters are based on the benchmark calibration summarized in Table 1.1.
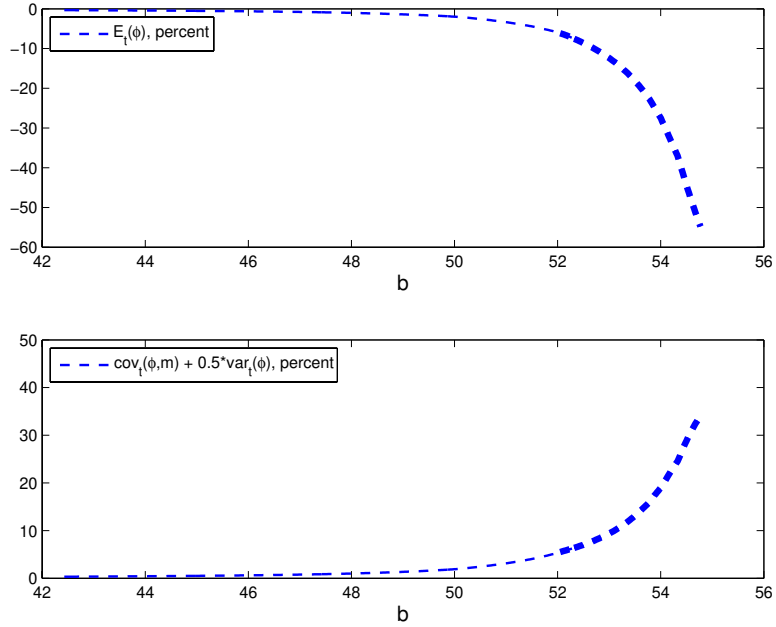
FIGURE 1.11: Decomposition of Stochastic Discount Factor

This figure shows the $E_t(\phi_{t+1})$ and $\frac{1}{2}var_t(\phi_{t+1}) + cov_t(m_{t+1}, \phi_{t+1})$, two components in the decomposition of the conditional mean of augmented stochastic discount factor, as shown in equation (1.35). b_{ss} denotes the average debt level in this economy, suggested by a long simulation from the model. The part of curves highlighted in bold denotes the region at which the constraint is binding. The vertical axis is measured in annualized percentage. The parameters are based on the benchmark calibration summarized in Table 1.1.
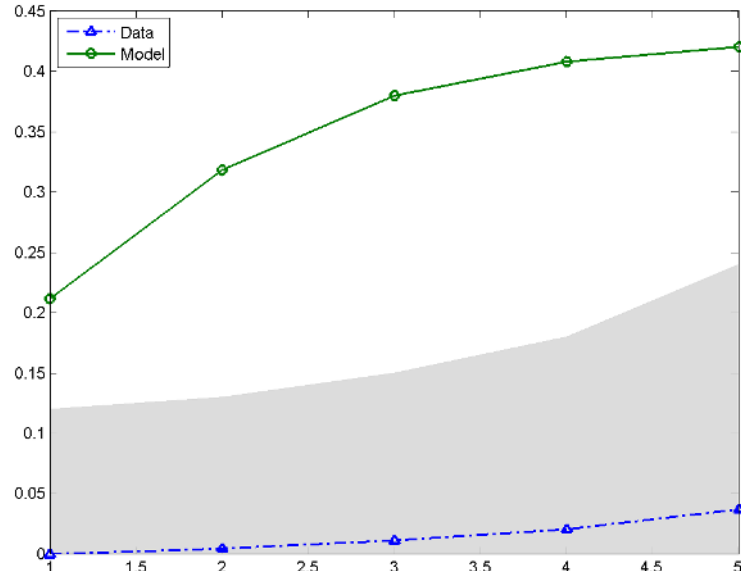
FIGURE 1.12: Price-dividend Ratio and Backward Consumption Growth

This figure plots the R^2 for regressing future log price-dividend ratio onto distributed lags of consumption growth:

$$p_{t+1} - d_{t+1} = \alpha_0 + \sum_{j=1}^{L} \alpha_j \Delta c_{t+1-j} + u_{t+1}$$

where L, the number of lags, is depicted on the horizontal-axis. The shaded area in the figure corresponds to the 95% confidence band in which data-based standard errors are constructed using a block-bootstrap. The data employed in the estimation are real, compounded continuously, sampled on an annual frequency and cover the period from 1930 to 2011. The "model" panel presents the predictability evidence implied by the model, based on a very long path of simulated data and the benchmark calibration.

2

Toward a Quantitative General Equilibrium Asset Pricing Model with Intangible Capital

2.1   Introduction

Historically, stocks with high book-to-market ratios, that is, value stocks, earn a higher average return than those with low book-to-market ratios, that is, growth stocks (Fama and French 1992, 1995). The difference in log units is approximately 4.3% per year and is known as the value premium. The market-to-book ratio of a firm is often viewed as a measure of the intensity of future growth options relative to assets currently in place. Interpreted this way, the empirical evidence on value premium suggests that the average spread between the return on physical assets in place and growth options is comparable to the aggregate stock market equity premium.

In this article we propose a quantitative general equilibrium model in which

71

growth options form intangible capital. When calibrated to standard statistics of the dynamics of macroeconomic quantities, our model is able to reproduce key features of asset returns data, including the difference in the average return on installed physical capital and future growth opportunities. Our model generates a high equity premium (5.66% per year for the market return, in log units) with a risk aversion of ten and a low and smooth risk-free interest rate. Our results are comparable to those obtained by the standard real business cycle (RBC) models in terms of the second moments of aggregate consumption, investment, and hours worked. Furthermore, the expected annual log return on growth options is 4.08% lower than that on installed physical capital, a significant share of the observed value premium in the data.

We follow Ai (2009) and model growth options as intangible capital in an otherwise standard neoclassical production economy. In contrast to assets in place, growth options do not produce consumption goods, and hence their payoff is not directly linked to aggregate productivity shocks. Rather, they represent an investment opportunity that allows their owner to build new production units using physical investment goods. Higher aggregate investment enables a greater fraction of growth options to be implemented and yield a higher payoff. Thus, in our model, the returns of growth options and physical capital depend on different risk factors and hence feature different risk premiums in equilibrium.

We make two major modifications to the Ai (2009) model. First, we adopt recursive preferences and an aggregate productivity process with long-run risk as in Croce (2008). This allows us to generate a highly volatile pricing kernel. More importantly, we show that in our model physical capital endogenously has a much higher exposure to long-run risk than intangible capital. Our production-based model thus rational-

72

izes the empirical findings on the cross-section of equity returns in Bansal, Dittmar, and Lundblad (2005), Hansen, Heaton, and Li (2008), and Kiku (2006).

Second, focusing on U.S. microeconomic data we document that the productivity of new vintages of capital is less sensitive to aggregate productivity shocks than that of older vintages. Based on this novel empirical finding, our model features heterogeneous productivity of vintage capital, with young vintages having lower exposure to aggregate shocks, as in the data. As a result, in our economy the response of physical investment with respect to unexpected fluctuations in aggregate productivity (short-run shocks) is positive, as in standard RBC models, but it is negative with respect to news about future productivity shocks (long-run shocks). These findings provide a crucial explanation of the high equity premium, large spread between the return on growth options and assets in place, and significant investment volatility observed in the data.

In our setup, the elasticity of substitution between tangible investment and intangible capital is high, implying that the adjustment of tangible capital is not costly. Consequently, investment responds strongly to contemporaneous productivity shocks, as it does in standard RBC models. The response of investment to long-run shocks, however, is sluggish for two reasons. First, news shocks predict future productivity growth but do not affect current output. Because of consumption-smoothing motives, the agent tends to avoid dramatic changes in investment, as they cause fluctuations in consumption in the opposite direction. Second, because new investments are less exposed to aggregate shocks because of their young age, their productivity is affected by news shocks only with a delay. The agent, therefore, finds it optimal to postpone the adjustment of investment with respect to such shocks. In equilibrium,

73

after a long-run productivity shock, the price of physical capital responds immediately and sharply, whereas physical investment and the return on growth options do not. This feature of our model is novel and allows us to reproduce both the equity premium and the value premium observed in the data, while maintaining the appealing features of the traditional RBC models on the quantity side.

Our analysis contributes to several strands of literature. We follow Hansen, Heaton, and Li (2005) and Li (2009) and interpret the spread in the return on book-to-market-sorted portfolios as evidence for the difference in the risk premiums of tangible and intangible capital. Hansen, Heaton, and Li (2005) believe that this observation "has potentially important ramifications for how to build explicit economic models to use in constructing measures of the intangible capital stock." The purpose of our article is to develop such a model and provide a unified framework to both measure and price intangible assets.

Our article is related to the literature on real options and the cross-section of equity returns (see, e.g., Berk, Green, and Naik 1999; Gomes, Kogan, and Zhang 2003; Carlson, Fisher, and Giammarino 2004; Cooper 2006) and the literature on adjustment costs and value premium (Zhang 2005; Gala 2005). However, our study differs from the above literature along several dimensions. First, in our economy, growth options are less risky than assets in place, whereas in previous real options–based models the opposite is true. The real options–based literature, by and large, explains the observed value premium by postulating that value firms are option intensive, while growth firms are assets in place intensive. Empirical evidence, however, suggests that growth firms are option intensive. Typically, growth firms have higher R&D investment (Li and Liu 2010) and a higher capital-expenditure-to-sales

74

ratio (Da, Guo, and Jagannathan 2012), two commonly used empirical proxies for firms' growth opportunities. Growth firms also feature longer cash-flow duration than value firms (see, e.g., Dechow, Sloan, and Soliman 2004, Da 2006, and Santos and Veronesi 2010), consistent with the interpretation that their assets consist mainly of options rather than installed physical capital. More recently, Kogan and Papanikolaou (2009, 2010) provide direct empirical evidence for the lower average return of growth options relative to assets in place. Our framework is consistent with the above empirical findings, because in our economy assets in place have both higher returns and shorter duration than growth options.

Second, we work in general equilibrium and study the quantitative implications of our model for asset prices as well as the joint dynamics of consumption, investment, and hours worked. Many of the above-mentioned articles, however, present partial equilibrium models. Although Gomes et al. (2003) and Gala (2005) adopt a general equilibrium approach, they do not focus on standard RBC moments. In contrast, we use the empirical evidence on the quantity side of the economy to discipline our model of production technology and, therefore, its asset pricing implications. Our unified neoclassical framework combines the success of the RBC models on the quantity side with the success of long-run risk–based models on the cross-section of equity returns obtained in endowment economies.

Third, our model assumes a long-run component in productivity and endogenously generates a long-run component in consumption growth. We show that value stocks are more exposed to long-run shocks than are growth assets. This feature of our model is consistent with the empirical evidence presented in Bansal, Dittmar, and Lundblad (2005), Hansen, Heaton, and Li (2008), and Kiku (2006).

Similarly to our approach, Ai and Kiku (2012) also explore conditions under which growth options are less risky than assets in place because of lower exposure to long-run risk. Their analysis differs from ours, however, in that in their model the creation of intangible assets is exogenous, and they do not confront the model with empirical evidence on macroeconomic quantities, such as investment or hours worked.

Our article builds on the literature on asset pricing in production economies, which was recently surveyed by Kogan and Papanikolaou (2011). Our work differs from previous articles in two significant respects. First, our model addresses the equity premium puzzle, as does the rest of the literature, but more importantly we also study the spread between the returns on tangible and intangible capital. Second, this literature typically relies on capital adjustment costs or other frictions in investments to generate variations in the price of physical capital. However, strong adjustment costs, although necessary to generate a sizeable equity premium, are often associated with either a counterfactually low volatility of investment or a counterfactually high volatility of the risk-free interest rate. Our model simultaneously produces a low volatility of the risk-free interest rate, a significant volatility of stock market returns, and a high volatility of investment, as in the data.

In a recent study, Borovička et al. (2011) develop methods to analyze the sensitivity of quantities and asset prices with respect to macroeconomic shocks in dynamic stochastic general equilibrium models. Borovička and Hansen (2011) focus on the discrete-time case and examine the shock-exposure and shock-price elasticities of tangible and intangible capital generated by our model. They reinterpret the difference in the productivity of old and young capital vintages as an investment-specific

76

shock, in the spirit of Papanikolaou (2011). In contrast to Papanikolaou, we abstract away from independent productivity shocks in the investment sector and provide a microfoundation for the adjustments in the relative price of capital vintages. We show that these adjustments arise endogenously when capital vintages have different exposure to conventional total factor productivity shocks, consistent with our empirical findings.

Finally, our article also relates to the literature that emphasizes the importance of intangible capital in understanding macroeconomic quantity dynamics and asset prices. Hall (2001) infers the quantity of intangible capital in the U.S. economy from a capital adjustment cost model. McGrattan and Prescott (2010a, 2010b) emphasize the importance of intangible capital in understanding economic fluctuations. Jovanovic (2008) models intangible capital as investment options and investigates its implications on aggregate Tobin's Q. Gourio and Rudanko (2010) focus on the relationship between customer capital, investment, and aggregate Tobin's Q. Lin (2010) studies intangible capital and stock returns in a partial equilibrium model with capital adjustment cost. Eisfeldt and Papanikolaou (2012) analyze organization capital and the cross-section of expected returns. Although providing insights on intangible capital, these articles do not study the difference in the expected return of value and growth stocks.[1]

The remainder of the article is organized as follows. We present the model and some analytical results in Sections 2.2 and 2.3. In Section 2.4, we provide empirical

---

[1] We do not intend to claim that all forms of intangible capital are less risky than physical capital. In fact, several of the above-mentioned papers suggest that certain forms of intangible capital may be riskier than physical capital. Same as Hansen et al. (2006), we believe that the historical difference in the average return of value and growth stocks calls for a theory in which intangible capital can be less risky than physical capital.

evidence on the lower risk exposure of new investments relative to physical capital of older vintages. We discuss the quantitative implications of our benchmark model in Section 2.5 and consider relevant extensions in Section 2.6. Section 2.7 concludes. Proofs of the theorems and the robustness analysis of the empirical results can be found in the Appendix B.

2.2   Model Setup

2.2.1   Preferences

Time is discrete and infinite, $t = 1, 2, 3, \cdots$. The representative agent has Kreps and Porteus (1978) preferences, as in Epstein and Zin (1989):

$$V_t = \left\{ (1 - \beta) \, u \, (C_t, N_t)^{1 - \frac{1}{\psi}} + \beta \left(E_t \left[V_{t+1}^{1-\gamma} \right] \right)^{\frac{1 - 1/\psi}{1 - \gamma}} \right\}^{\frac{1}{1 - 1/\psi}},$$

where C_t and N_t denote, respectively, the total consumption and total hours worked at time t. For simplicity, we assume an inelastic labor supply and set $u \, (C_t, N_t) = C_t$. We relax this assumption in Appendix B.3

2.2.2   Production technology

Production units.   Consumption goods are produced by production units of overlapping generations. Production units created at time τ are called generation-τ production units and begin operation at time $\tau + 1$. Each generation-τ production unit uses labor, n_t^τ, as the only input of production and pays a competitive real wage w_t. For $t \geq \tau + 1$, let A_t^τ denote the time t labor productivity level common to all the production units belonging to generation τ. The output of a generation-τ production

unit at time t, y_t^τ, is given by

$$y_t^\tau = (A_t^\tau n_t^\tau)^{1-\alpha}, \quad \forall t \geq \tau + 1.$$

At the equilibrium, the cash flow of a generation-τ production unit at time t is given by

$$\pi_t^\tau = \max_n \left\{ (A_t^\tau n)^{1-\alpha} - w_t n \right\}.$$

In our setup, labor productivity, A_t^τ, is generation-specific and captures the heterogenous exposure of production units of different vintages to aggregate productivity shocks. The productivity processes are specified as follows. First, we assume that the log growth rate of the productivity process for the initial generation of production units, Δa_{t+1}, is given by

$$\log \frac{A_{t+1}}{A_t} \equiv \Delta a_{t+1} = \mu + x_t + \sigma_a \varepsilon_{a,t+1}, \qquad (2.1)$$

$$x_{t+1} = \rho x_t + \sigma_x \varepsilon_{x,t+1},$$

$$\begin{bmatrix} \varepsilon_{a,t+1} \\ \varepsilon_{x,t+1} \end{bmatrix} \sim i.i.d.N \left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \right), \quad t = 0, 1, 2, \cdots.$$

This specification follows Croce (2008) and captures long-run productivity risks.

Second, we impose that the growth rate of the productivity of production units of age $j = 0, 1, ..., t-1$ is given by

$$\frac{A_{t+1}^{t-j}}{A_t^{t-j}} = e^{\mu + \phi_j (\Delta a_{t+1} - \mu)}. \qquad (2.2)$$

Under the above specification, production units of all generations have the same unconditional expected growth rate. We also set $A_t^t = A_t$ to ensure that new pro-

duction units are on average as productive as older ones.[2] Heterogeneity hence is driven solely by differences in exposure to aggregate productivity risk, ϕ_j. Our empirical investigation in Section 2.4 suggests that ϕ_j is increasing in j, that is, older production units are more exposed to aggregate productivity shocks than younger ones.[3] To capture this empirical fact, we adopt a parsimonious specification of the ϕ_j function as follows:

$$\phi_j = \begin{cases} 0 & j = 0 \\ 1 & j = 1, 2, \cdots \end{cases}.$$

That is, new production units are not exposed to aggregate productivity shocks in the initial period of their life, and afterwards their exposure to aggregate productivity shocks is identical to that of all other existing generations.

We discuss the empirical evidence on heterogeneous exposure in Section 2.4, and we consider more general specifications of the ϕ_j function in Section 2.6. Providing a microeconomic foundation for this feature of the model is beyond the scope of this study. However, we note that both our empirical evidence and the specification of ϕ_j are consistent with the learning model of Pastor and Veronesi (2009). In their economy, young firms are subject to substantial idiosyncratic risks but have very little exposure to aggregate shocks. The reason is that young firms are embedded with new technologies, which are highly uncertain. It is not optimal to operate these new technologies on a large scale until the uncertainty is reduced with learning. As

---

[2] Generation-t production units are not active until period $t + 1$; therefore, the level of A_t^t does not affect the total production of the economy in period t.

[3] In the data, the productivity process of young firms has a higher idiosyncratic volatility than that of older firms. To capture this fact, generation-specific shocks should be included in Equation (2.2). After solving the model with these additional shocks, however, we find only negligible differences in our results. We therefore choose not to include this additional source of shocks for parsimony.

a result, shocks to young firms have little impact on aggregate quantities. Over time, as young firms age, their productivity becomes more correlated with aggregate output because their technologies are adopted on a larger scale.

In our economy, it is convenient to measure production units of all generations in terms of their generation-0 equivalents. As we show in Appendix B.1, our specification of the productivity process implies that the output and cash flow of a generation-t production unit are ϖ_{t+1} times greater than those of a generation-0, where

$$\varpi_{t+1} = \left(\frac{A_{t+1}^t}{A_{t+1}}\right)^{\frac{1-\alpha}{\alpha}} = e^{-\frac{1-\alpha}{\alpha}(x_t + \sigma_a \varepsilon_{a,t+1})} \quad \forall t. \tag{2.3}$$

We use $p_{K,t+1}$ to denote the cum-dividend value of a generation-0 production unit at time $t+1$. Because the cash flow of generation-t production units is ϖ_{t+1} times that of a generation-0 production unit, the value of a new production unit created at time t measured in time-t consumption numeraire is $E_t[\Lambda_{t,t+1}\varpi_{t+1}p_{K,t+1}]$, where $\Lambda_{t,t+1}$ denotes the stochastic discount factor. We also assume that a production unit dies with probability δ_K at the end of each period, and death shocks are *i.i.d.* across production units and over time.

Blueprints. The only way to construct a new production unit in this economy is to implement a blueprint. Implementing a blueprint at time t costs $\frac{1}{\theta_t}$ units of physical investment goods. We call θ the quality of a blueprint, because blueprints with high θ are more efficient in constructing production units. We allow θ_t to differ across blueprints and evolve stochastically over time to capture idiosyncratic shocks to the profitability of blueprints. At the beginning of each period t, first the value of θ_t is

revealed, and then the owner of the blueprint makes the decision of whether or not to implement it. A blueprint can only be implemented once, and implementation is irreversible. If not implemented immediately, a blueprint dies with probability δ_S at the end of the period, and death shocks are *i.i.d.* across blueprints and over time.

In our setup, at any time t, the owner of a blueprint faces an optimal stopping problem. She can choose to build a production unit immediately at cost $\frac{1}{\theta_t}$. Alternatively, she may delay the implementation decision into the future. If we denote the value of a blueprint with quality θ_t at time t as $p_{S,t}(\theta_t)$, then the following recursive relation holds:

$$p_{S,t}(\theta_t) = \max\left\{ E_t\left[\Lambda_{t,t+1}\varpi_{t+1}p_{K,t+1}\right] - \frac{1}{\theta_t}, \quad (1-\delta_S)E_t\left[\Lambda_{t,t+1}p_{S,t+1}(\theta_{t+1})\right] \right\}. \quad (2.4)$$

The first term in the brackets is the payoff of immediate option exercise: Implementing a blueprint with quality θ_t at time t costs $\frac{1}{\theta_t}$ of the amount of general output and creates a generation-t production unit whose value is $E_t\left[\Lambda_{t,t+1}\varpi_{t+1}p_{K,t+1}\right]$. The second term is the payoff associated with delaying option exercise: With probability $1-\delta_S$ the blueprint survives to the next period and obtains another draw of θ_{t+1}.

In our economy, the supply of blueprints is endogenous. At time t, a total measure J_t of new blueprints can be produced by investing J_t units of output. Blueprints created at time t can be used to build production units starting from period $t+1$.

Interpretation. Production units are the building blocks of assets in place. Their creation requires physical output and their value is reflected in the accounting books. They produce final goods directly and generate payoffs immediately.

Blueprints are growth options. They capture key features of innovations and

new investment opportunities. They are subject to substantial idiosyncratic risk (θ) and are implemented only if their quality becomes high enough. Blueprints do not produce any consumption goods immediately; they only start to do so after being implemented. They are intangible in the sense that they are claims to future output and lack physical embodiment. According to U.S. accounting rules, the cost of developing new blueprints, such as innovations and new investment opportunities, is typically expensed rather than capitalized. For this reason, we think of J_t as an intangible investment. In the rest of the article, we use the terms blueprints and growth options, and the terms production units and assets in place, interchangeably.

Both production units and blueprints constitute a form of capital because they can be stored and thus allow investors to trade off current-period consumption against future consumption. Specifically, production units are tangible capital, and blueprints are intangible capital. We are interested in understanding how the different roles played by tangibles and intangibles in aggregate production determine their expected returns.

In our setup, stocks feature high book-to-market ratios (value stocks) if they consist mainly of claims to tangible capital. Conversely, low book-to-market ratio stocks (growth stocks) are intangible capital intensive. At the equilibrium, the value premium reflects the difference in the expected returns on tangible and intangible capital.

Our notions of value and growth are also consistent with the empirical evidence on the negative relation between cash-flow duration and book-to-market value (see, e.g., Dechow, Sloan, and Soliman 2004 and Da 2006). Our value stocks feature short cash-flow duration because they are mainly claims to assets in place that pay off

immediately. Growth stocks, in our model, are long-duration assets, because they load heavily on growth options, which generate cash flows only in the distant future after they are implemented and become production units.

2.2.3 Aggregation

Tangible capital. We use M_t to denote the total measure of production units created at time t and use K_t to denote the productivity-adjusted total measure of production units expressed in generation-0 equivalents. The advantage of using K_t as a state variable is that the aggregate production function is of the Cobb-Douglas form despite the heterogeneity across vintages. If we let Y_t denote aggregate output, then the following holds:

$$Y_t = \sum_{\tau=0}^{t-1} (1 - \delta_K)^{t-\tau-1} M_\tau y_t^\tau = K_t^\alpha (A_t N_t)^{1-\alpha}, \tag{2.5}$$

where A_t is the labor productivity of generation-0 production units. In Appendix B.1, we show that the law of motion of the productivity-adjusted measure of tangible capital, K_t, takes the following simple form:

$$K_1 = M_0, \quad K_{t+1} = (1 - \delta_K) K_t + \varpi_{t+1} M_t, \quad t = 1, 2, \cdots.$$

Our specification of productivity has two advantages. First, it provides a parsimonious way to incorporate the empirical fact that new investments are less exposed to aggregate productivity shocks than capital of older vintages. Second, it maintains tractability at the aggregate level.

84

Intangible capital. To avoid having to keep track of the distribution of θ_t as an infinite dimensional state variable, we assume that θ_t is *i.i.d.* among blueprints and over time. For simplicity, we also assume that the distribution of θ_t has a continuous density, denoted as f. As shown in Ai (2009), in this case the mass of newly created production units, M_t, depends only on the total measure of all available blueprints at time t, denoted as S_t, and the total amount of tangible investment goods, I_t, through the following relation:

$$M_t = G(I_t, S_t) = \max_{\theta_t^*} \left\{ S_t \times \int_{\theta_t^*}^\infty f(\theta)\, d\theta \right\}, \text{ subject to} \qquad (2.6)$$

$$S_t \times \int_{\theta_t^*}^\infty \frac{1}{\theta} f(\theta)\, d\theta \le I_t,$$

where the function G is defined as the value function of the optimization problem (2.6).

Intuitively, optimal option exercise follows a simple cutoff rule: Blueprints are implemented in period t if and only if their quality exceeds θ_t^*. Ai (2009) provides a formal proof of this claim and shows that

$$\theta_t^* = G_I(I_t, S_t). \qquad (2.7)$$

In each period, the agent chooses tangible investment, I_t, and exercises top-quality options until the exhaustion of all physical investment goods. Therefore, given the resource constraint in Equations (2.6) and (2.7), both M_t and θ_t^* are fully determined by I_t and S_t.

Note that one blueprint transforms into exactly one production unit after implementation. Therefore, $G(I_t, S_t)$ is the total measure of both the newly created

production units and the blueprints implemented. Taking into account the amount of new blueprints created, J_t, the dynamics of the intangible stock, S_t, is

$$S_{t+1} = [S_t - G(I_t, S_t)](1 - \delta_S) + J_t. \tag{2.8}$$

Using Equation (2.6), the law of motion of K_t can be written as

$$K_{t+1} = (1 - \delta_K) K_t + \varpi_{t+1} G(I_t, S_t). \tag{2.9}$$

Finally, we assume that general output can be transformed frictionlessly into consumption, C_t, tangible investment, I_t, and intangible investment goods, J_t, so that the implied aggregate resource constraint is given by

$$C_t + I_t + J_t \le K_t^\alpha (A_t N_t)^{1-\alpha}. \tag{2.10}$$

2.2.4 Relation to the literature

Our model of growth options follows the general equilibrium setup in Ai (2009) and differs from existing studies in several respects. First, unexercised growth options can be stored and potentially implemented in the future. The storability of unexercised growth options makes them a type of capital distinct from physical assets in place. In contrast, Berk, Green, and Naik (1999) and Gomes, Kogan, and Zhang (2003) assume that options disappear if not immediately exercised.

Second, the creation of new growth options in our model is endogenously determined by the optimal choice of the agent. This allows not only the price but also the quantity of intangible capital to adjust to productivity shocks in general equilibrium. The endogenous quantity channel increases the representative agent's ability to smooth consumption and allows options to be *less* risky than assets in place. In

contrast, partial equilibrium–based real-option models (e.g., Berk, Green, and Naik 1999; Gomes, Kogan, and Zhang 2003; and Carlson, Fisher, and Giammarino 2004) typically assume exogenous arrival of growth options and abstract from the quantity adjustment channel. As a result, options are *more* risky than assets in place in these models.

Third, our intangible capital is the stock of growth options and does not immediately produce output, as does tangible capital. This feature links the cross-sectional differences in both stock returns and their cash-flow duration to production technology. The macroeconomic literature that focuses on the quantity dynamics of intangible capital, in contrast, typically assumes that both intangible and tangible capital affect output directly. For example, the aggregate production function in McGrattan and Prescott (2010a, 2010b) and Corrado, Hulten, and Sichel (2006) are of the form $Y_t = F(A_t, K_t, S_t, N_t)$, where K_t and S_t denote tangible and intangible capital, respectively. This specification implies that the payments to tangible and intangible capital have similar duration and are both perfectly conditionally correlated with aggregate productivity shocks, thus allowing little room for differences in expected returns.

Finally, the incorporation of intangible capital presents additional challenges to general equilibrium asset pricing models with production. Because of the well-known difficulty in generating a high equity premium in production economies, one might be tempted to assume that intangible capital is much riskier than physical capital and propose this as a resolution of the equity premium puzzle. However, as argued by Hansen, Heaton, and Li (2005), the empirical evidence on the value premium suggests the exact opposite. In the United States, the portfolios of firms with low

87

book-to-market ratios pay substantially lower returns than those of firms with high book-to-market ratios. This suggests that intangible capital earns a much lower risk premium than tangible capital, making it even harder to account for the overall market equity premium. We turn now to the solution of the model and discuss a mechanism that simultaneously generates a high equity premium and a high value premium.

2.3 Model Solution

2.3.1 The social planner's problem

We consider a competitive equilibrium with complete markets in which claims to production units and blueprints are traded. The equilibrium allocation and prices can be constructed from the solution to the social planner's problem that maximizes the representative agent's utility:

$$
V\left(K_{t}, S_{t}, x_{t}, A_{t}\right)=\max_{C_{t}, I_{t}, J_{t} \geq 0}\left\{(1-\beta) C_{t}^{1-\frac{1}{\psi}}+\beta\left(E\left[V\left(K_{t+1}, S_{t+1}, x_{t+1}, A_{t+1}\right)^{1-\gamma} \mid x_{t}, A_{t}\right]\right)^{\frac{1-1/\psi}{1-\gamma}}\right\}^{\frac{1}{1-1/\psi}},
$$

subject to the evolution of productivity (Equations (2.1) and (2.2)), the resource constraint (Equation (2.10)), and the laws of motion of S_t and K_t (Equations (2.8) and (2.9)). We refer the reader to Ai (2009) for a formal proof of the equivalence between the competitive equilibrium allocation and Pareto optimality.

Despite the heterogeneity in productivity of production units and quality of blueprints, our formulation of the social planner's problem does not use cross-sectional distributions. Our model hence maintains the tractability of standard RBC models—relevant to study macroeconomic quantity dynamics—and simultaneously allows us to study both option-exercise and the cross-section of physical and intangible capital

88

returns in general equilibrium.

2.3.2   *Asset prices*

Given equilibrium allocations, the stochastic discount factor of the economy, $\Lambda_{t,t+1}$, can be represented by the ratio of marginal utilities at time t and $t + 1$:

$$\Lambda_{t,t+1} = \beta \left(\frac{C_{t+1}}{C_t}\right)^{-\frac{1}{\psi}} \left[\frac{V_{t+1}}{\left(E_t\left[V_{t+1}^{1-\gamma}\right]\right)^{\frac{1}{1-\gamma}}}\right]^{\frac{1}{\psi}-\gamma}. \tag{2.11}$$

Let $p_{K,t}$ and $q_{K,t}$ denote the time-t cum- and ex-dividend price of a generation-0 production unit, respectively. Let $p_{S,t}(\theta)$ denote the value of a blueprint with quality θ at time t before the option exercise decision is made. Because shocks to θ are *i.i.d.* over time, the time-t price of an ex ante identical blueprint before the revelation of θ_t, denoted $p_{S,t}$, is

$$p_{S,t} = \int_0^\infty p_{S,t}(\theta) f(\theta) d\theta,$$

and can be interpreted as the per-unit value of the perfectly diversified aggregate stock of blueprints. We also use $q_{S,t}$ to denote the price of a newly created blueprint at time t.

We use the first-order and envelope conditions of the social planner's problem to characterize the price of growth options and assets in place, as stated in the following proposition.

Proposition 1. (Equilibrium Conditions) Assets in place are priced as follows:

$$\begin{aligned}
p_{K,t} &= \alpha K_t^{\alpha-1}(A_t N_t)^{1-\alpha} + (1-\delta_K) q_{K,t}, & (2.12) \\
q_{K,t} &= E_t\left[\Lambda_{t,t+1} p_{K,t+1}\right].
\end{aligned}$$

A blueprint with quality θ is implemented at time t if and only if $\theta \geq \theta_t^*$, where θ_t^* satisfies

$$E_t \left[\Lambda_{t,t+1} \varpi_{t+1} p_{K,t+1} \right] - \frac{1}{\theta^*(t)} = (1 - \delta_S) E_t \left[\Lambda_{t,t+1} p_{S,t+1} \right]. \tag{2.13}$$

The price of growth options is determined as follows:

$$p_{S,t} = \frac{G_S(I_{t+1}, S_{t+1})}{G_I(I_{t+1}, S_{t+1})} + (1 - \delta_S) q_{S,t} \tag{2.14}$$

$$q_{S,t} = E_t \left[\Lambda_{t,t+1} p_{S,t+1}(\theta_{t+1}) \right] = 1.$$

Proof. See Ai (2009). $\square$

Together, the two equations in (2.12) constitute a recursive relation that can be used to solve for $p_{K,t}$ given equilibrium quantities. The interpretation is that the value of a unit of tangible capital is equal to the present value of its marginal product.

Because a blueprint is implemented at time t if and only if its quality exceeds the threshold level, θ_t^*, Equation (2.13) implies that the owner of a marginal blueprint with quality θ_t^* must be indifferent between immediate option exercise and delaying implementation into the future.

Equation (2.14) provides a decomposition of option value into in-the-money and out-of-the-money payoff components. The value of an unexercised option is $(1 - \delta_S) q_{S,t}$ after accounting for the death shock. The term $\frac{G_S(I_t, S_t)}{G_I(I_t, S_t)}$ can be interpreted as the expected payoff of an in-the-money option and is an increasing function of $\frac{I}{S}$ by the homogeneity and the concavity of G. Intuitively, a rise in $\frac{I}{S}$ increases the probability of growth options to be exercised and therefore their payoff rises as well. From

90

the social planner's perspective, $\frac{G_S(I,S)}{G_I(I,S)}$ can be interpreted as the marginal product of intangible capital: $G_S(I,S)$ is the number of new production units that can be produced by an additional growth option, and $G_I(I,S)^{-1} = \frac{1}{\theta^*}$ is the price of a marginal production unit measured in current-period consumption goods. The value of an unexercised growth option, q_S, is always one because one unit of general output can always be transformed into one unit of new blueprints at time t.

Finally, note that Equations (2.7)–(2.14) completely characterize both aggregate quantities and prices in the economy. Given aggregate quantities, Equation (2.7) can be used to solve for the optimal option-exercise threshold for blueprints.

The returns of tangible and intangible capital can be therefore written as

$$r_{K,t+1} = \frac{p_{K,t+1}}{q_{K,t}} = \frac{\alpha K_{t+1}^{\alpha-1}(A_{t+1}N_{t+1})^{1-\alpha} + (1-\delta_K)q_{K,t+1}}{q_{K,t}}, \qquad (2.15)$$

and

$$r_{S,t+1} = \frac{p_{S,t+1}}{q_{S,t}} = \frac{G_S(I_{t+1}, S_{t+1})}{G_I(I_{t+1}, S_{t+1})} + (1-\delta_S), \qquad (2.16)$$

respectively. Equations (2.15) and (2.16) are the key to understanding the expected returns on tangible and intangible capital. Equation (2.15) implies, as is common in standard RBC models, that the return on assets in place is monotonic in aggregate productivity shocks. In contrast, the return on growth options does not depend directly on productivity shocks, and in fact it is a function only of the I_{t+1}/S_{t+1} ratio. The return on intangibles is high in states in which the demand for options is large, that is, when I is large relative to the total supply of growth options, S. Our choice to model intangibles as growth options thus allows the return on

physical and intangible capital to depend on different risk factors and, consequently, to command different risk premiums in equilibrium. In Section 2.5, we show that physical investment I is not responsive to long-run productivity shocks. As a result, the return on intangible capital has little exposure to long-run risk, whereas physical capital is highly risky.

2.4 Firms' Exposure to Aggregate Risks

In this section we provide empirical evidence supporting the claim that new production units are less sensitive to aggregate productivity shocks than are older vintages of physical capital. A production unit in our model should be interpreted as any investment project generating cash flows. Because it is difficult to identify both productivity and age of individual projects within firms, we adopt an indirect approach and work with firm-level data. Specifically, for each firm in our data set we estimate the time series of its productivity growth rate and compute two alternative measures of the age of its assets in place. We find that the correlation between firm-level and aggregate productivity growth is statistically smaller for firms with younger vintages of physical capital.

2.4.1 Data and firm-level productivity estimation

Data description. We consider publicly traded companies on U.S. stock exchanges listed in both the Compustat and the Center for Research in Security Prices (CRSP) databases for the period 1950–2008. In what follows, we report Compustat items in parentheses and define industry at the level of two-digit SIC codes. The output, or value added, of firm i in industry j at time t, $y_{i,j,t}$, is calculated as sales (sales) minus

the cost of goods sold (`cogs`) and is deflated by the aggregate gross domestic product (GDP) deflator from the U.S. National Income and Product Accounts (NIPA). We measure the capital stock of the firm, $k_{i,j,t}$, as the total book value of assets (`at`) minus current assets (`act`). This allows us to exclude cash and other liquid assets that may not be appropriate components of physical capital. We use the number of employees in a firm (`emp`) to proxy for its labor input, $n_{i,j,t}$, because data for total hours worked are not available.

We construct two measures of the age of assets in place of firm i at time t. Our first measure is simply the age of firms, calculated using founding years from Ritter and Loughran (2004) and Jovanovic and Rousseau (2001). This procedure enables us to form a large data set with 8,084 different firms and 83,089 observations.

Our second measure is capital age, $KAge_{i,t}$, which we compute as follows:

$$KAge_{i,t} = \frac{\sum_{l=1}^{T}(1-\delta_i)^l \cdot I_{i,t-l} \cdot l}{\sum_{l=1}^{T}(1-\delta_i)^l \cdot I_{i,t-l}}, \tag{2.17}$$

where $I_{i,t}$ measures capital expenditures (`capx`), and δ_i is the firm-specific depreciation rate (depreciation expenses (`xdp`) divided by book value of property, plant, and equipment (`ppent`)) averaged over time. When data on depreciation expenses are not available, we measure depreciation by Compustat depreciation (`dp`) minus amortization of intangibles (`am`). According to the above definition, the capital age of a firm is the weighted average age of its capital vintage if we set $T = \infty$. Empirically, we can only choose a finite T and face the following trade-off: A large T provides a better approximation of the age of capital vintage, but it considerably reduces the number of observations in our data set.

In Table 2.1, we sort all observations in our panel into four firm-age quantiles and present summary statistics. For each quantile, we report median firm age (Column 2) and median capital age calculated using $T = 5$, $T = 8$, and $T = 15$ (Column 3-5, respectively). All measures of capital age are increasing in firm age, indicating that they are consistent with each other. Across all our measures, the average age in the first quartile is statistically different from that in the fourth quartile. Therefore, our age quartiles capture a significant amount of the dispersion in both capital and firm age.

Table 2.1 explicitly shows the trade-off related to the choice of T. If we use the average annual depreciation rate from Compustat of 15%, setting $T = 15$ implies that we account for roughly 92% of the firms' total capital stock. This choice of T provides a fairly good approximation of the true capital vintage of the firms, but it only allows us to compute capital age for 36% of the $8,084$ firms for which firm age is available. On the other hand, setting $T = 5$ permits us to retain all our firms, but this captures only 62% of firms' most recent capital stock. To keep our discussion focused, we present our empirical evidence using firm age as the main proxy for the age of firms' production units. In Appendix B.2, we show that our empirical results are robust to different measures of capital age.

Estimation of firm-level productivity. We assume that the production function at the firm level is Cobb-Douglas and allow the parameters of the production function to be industry-specific:

$$y_{i,j,t} = A_{i,j,t} k_{i,j,t}^{\alpha_{1,j}} n_{i,j,t}^{\alpha_{2,j}}, \tag{2.18}$$

where $A_{i,j,t}$ is the firm-specific productivity level at time t. This is consistent with our original specification because the observed physical capital stock, $k_{i,j,t}$, corresponds to the mass of production units owned by the firm.

We estimate the industry-specific capital share, $\alpha_{1,j}$, and labor share, $\alpha_{2,j}$, using the dynamic error component model adopted in Blundell and Bond (2000) to correct for endogeneity. Details are provided in Appendix B.2. Given the industry-level estimates for $\widehat{\alpha}_{1,j}$ and $\widehat{\alpha}_{2,j}$, the estimated log productivity of firm i is computed as follows:

$$\ln \widehat{A}_{i,j,t} = \ln y_{i,j,t} - \widehat{\alpha}_{1,j} \cdot \ln k_{i,j,t} - \widehat{\alpha}_{2,j} \cdot \ln n_{i,j,t}.$$

We allow for $\alpha_{1,j} + \alpha_{2,j} \neq 1$, but our results hold also when we impose constant returns to scale in the estimation, that is, $\alpha_{1,j} + \alpha_{2,j} = 1$.

We use the multifactor productivity index for the private nonfarm business sector from the Bureau of Labor Statistics (BLS) as the measure of aggregate productivity.

2.4.2 Empirical results

Here, we present our estimates on the link between firm exposure to aggregate productivity and firm age. We provide additional robustness analyses of our results in Appendix B.2. We consider the following baseline regression:

$$\Delta \ln A_{i,j,t} = \xi_{0i} + \xi_1 \Delta \ln \overline{A}_t + \xi_2 AGE_{i,j,t} + \xi_3 AGE_{i,j,t} \cdot \Delta \ln \overline{A}_t + \xi_4 B/M_{i,j,t} + \varepsilon_{i,j,t}, \quad (2.19)$$

where ξ_{0i} is a firm-specific fixed effect, $\Delta \ln \overline{A}_t$ is the growth rate of aggregate productivity as measured by the BLS, and $B/M_{i,j,t}$ measures firm book-to-market ratio. We introduce the book-to-market ratio to control for the difference in the composition of tangible and intangible assets across firms. The key parameter of interest here

is the coefficient ξ_3, which captures the age effect on firm sensitivity to aggregate productivity growth. If the average age of investment projects is increasing in firm age, then under the null of our model ξ_3 is positive.

We find strong empirical evidence in favor of our specification of firm productivity (Table 2.2). In our baseline estimation (regression (1)), the estimated coefficient ξ_3 is both positive and statistically significant. Furthermore, we obtain very similar point estimates in regressions (2) and (3), where we correct for possible sample selection bias induced by firm exits.

If exits caused by exposure to negative aggregate productivity shocks are correlated with firm age, they could induce an upward bias in our estimate of ξ_3 in regression (1). Consider a hypothetical scenario in which young firms are more exposed to negative aggregate productivity shocks than are older firms. In such a case, the estimate of ξ_3 obtained from regression (1) would be biased upwards, because young firms would be more likely to exit our database in years with large negative aggregate productivity shocks.

In regression (2) we correct for sample-selection bias by adopting the Heckman (1979) two-stage sample-selection estimator. In regression (3), we instead estimate Equation (2.19), excluding all the observations from years with negative aggregate productivity shocks. The details of these robustness analyses can be found in Appendix B.2, where we also adopt an additional estimation procedure for the coefficients of the production function. Across all these specifications, our estimates of ξ_3 are very robust: They are consistently positive, statistically significant, and comparable in magnitude.

Note also that the estimate of ξ_4 is consistently negative across all specifications,

implying that the productivity growth rate of growth firms is always higher than that of value firms. This is consistent with the view that growth firms have a longer cash-flow duration than value firms, a fact that our model replicates and that we address in Section 2.5.

Our specification of firms' productivity processes is not only qualitatively consistent with the pattern in the data, but also quantitatively plausible. In fact, our calibration matches well the magnitude of firms' transition from low to high exposure to aggregate productivity shocks. We denote by ϕ_Y (ϕ_O) the regression coefficient of the productivity growth of the young (old) capital vintages on aggregate productivity growth rates. In our model, $\phi_Y = 0$ and $\phi_O = 1.12$. To see why $\phi_O = 1.12$, note that the aggregate productivity growth rate is a weighted average of that of the new capital vintage, $A_{t+1}^t / A_t^t = e^\mu$, and the common growth rate of all older vintages, A_{t+1}/A_t:

$$\Delta \ln \bar{A}_{t+1} = (1 - \lambda_t) \Delta \ln A_{t+1} + \lambda_t \mu, \quad 1 > \lambda_t > 0.$$

The regression coefficient of $\Delta \ln A_{t+1}$ on aggregate productivity growth $\Delta \ln \overline{A}_{t+1}$ is therefore $\frac{1}{1-\overline{\lambda}}$. Assuming an annual death rate of 11% in the spirit of our calibration, $\overline{\lambda} = 11\%$ in steady state, and $\frac{1}{1-\overline{\lambda}} = 1.12$.

In the data, we estimate ϕ_Y and ϕ_O using the following regressions:

$$\Delta \ln A_{i,j,t} = \begin{cases} \xi_{0i} + \phi_Y \Delta \ln \overline{A}_t + \xi_{1i} B/M_{i,j,t} + \widetilde{\varepsilon}_{i,j,t} & i \in \text{Young} \\ \xi_{0i} + \phi_O \Delta \ln \overline{A}_t + \xi_{1i} B/M_{i,j,t} + \widetilde{\varepsilon}_{i,j,t} & \text{otherwise.} \end{cases} \quad (2.20)$$

In each period a firm is classified as Young if it belongs to the set of the 25% youngest firms in our sample. We report our estimation results in Table 2.3. Overall, our estimate of ϕ_Y is not statistically different from zero, and that of ϕ_O is positive and

significant. The difference in productivity exposure, $OMY = \phi_O - \phi_Y$, is positive and statistically significant, and the point estimate is close to its model counterpart, 1.12.

Our choice of the ϕ_j process is likely to understate the duration of the transition from low to high exposure. Our model assumes that production units have full exposure to aggregate productivity shocks after one period, whereas the median capital age of the young firms for $T = 15$ is 4.47 years, which suggests that the transition from low exposure to high exposure takes an average of 3.47 years. In Section 2.6, we extend our model to allow for more general specifications of the ϕ_j process and show that longer transitions further enhance the equity and value premiums generated by our model. Our current specification for the ϕ_j process reflects a conservative calibration.

2.5 Quantitative Implications of the Model

In this section, we calibrate our model at an annual frequency and evaluate its ability to replicate key moments of both macroeconomic quantities and asset returns. We focus on a long sample of U.S. annual data, including pre-World War II data. All macroeconomic variables are real and per capita. Consumption and physical investment data are from the Bureau of Economic Analysis (BEA), whereas intangible investment (J_t) is measured as in Corrado, Hulten, and Sichel (2006) by aggregating expenses in brand equity, firm-specific resources, R&D, and computerized information. As in the U.S. NIPA, we treat intangible investment as an expense and define measured output, $Y_{M,t}$, as $C_t + I_t$. Annual data on asset returns are from the Fama-French data set. We use the Fama-French *HML* factor as a measure of the spread

between tangible and intangible capital. Appendix B.2 provides more details on our data sources.

2.5.1 Parameter values

Our model has three major components: heterogeneous productivity of vintage capital, long-run productivity risk, and intangible capital. To determine the importance of each component, we compare four different calibrations. The benchmark model comprises all three components and is our preferred calibration. Model 1 lacks heterogeneous productivity of vintage capital (we set $\phi_0 = 1$) but retains the other features of the benchmark model, namely, long-run productivity risk and intangible capital. In model 2, we further exclude fluctuations in long-run productivity growth (by setting $\sigma_x = 0$). Finally, we consider the case without intangible capital in model 3. Essentially, model 3 is the neoclassical growth model with recursive preferences and $i.i.d.$ productivity growth rates. The details of the four models are summarized in Table 2.4.

The parameters of the models can be divided into three groups. The first group includes risk aversion, γ; intertemporal elasticity of substitution, ψ; capital share, α; depreciation rates, δ_K and δ_S; average growth rate of the economy, μ; and the first-order autocorrelation of the predictable component in productivity growth, ρ. These parameters are identical across all four calibrations. We choose the parameters for risk aversion, $\gamma = 10$, and intertemporal elasticity of substitution, $\psi = 2$, in line with the long-run risk literature. We set the capital share $\alpha = 0.3$ and the annual depreciation rate of physical capital $\delta_K = 11\%$, consistent with the RBC literature (Kydland and Prescott 1982). We choose the same rate of depreciation for intangible

99

capital, $\delta_S = 11\%$. The measured depreciation of intangible capital in our model also includes implemented blueprints, $G(I_t, S_t)$, and ranges from 40% to 60% per year across the calibrations. Although the empirical evidence on depreciation of intangibles is sparse (Hand and Lev 2003), these numbers are consistent with the empirical estimate in Corrado, Hulten, and Sichel (2006). Our sensitivity analysis suggests that δ_S only modestly affects our asset pricing results. We calibrate $\mu = 2\%$ per year, consistent with the average annual real growth rate of the U.S. economy. We set $\rho = 0.93$, which is the point estimate obtained in Croce (2008).

The second group of parameters includes the discount factor, β; the standard deviation of the persistent component of productivity growth, σ_x; and the short-run shock volatility, σ_a. In all calibrations, we set the discount factor β to match the level of the risk-free interest rate in the data if possible. An exception is model 3, which lacks sufficient parameters to match both the level of the risk-free rate and the consumption–tangible investment ratio. We therefore choose β in model 3 to match the consumption–tangible investment ratio but not the level of the risk-free rate. We set σ_a and σ_x in both the benchmark model and model 1 to approximately match the standard deviation and the first-order autocorrelation of the annual growth rate of measured output. In both models 2 and 3, we impose $\sigma_x = 0$ and set σ_a to match the standard deviation of the annual growth rate of measured output.

The third group of parameters describes the functional form of the aggregator $G(I_t, S_t)$ function. As shown in Ai (2009), for any smooth G function that is concave and homogeneous of degree 1, there is a unique density function $f(\theta)$ such that G is the aggregator of the option-exercise problem described in Equations (2.4) and (2.6). We focus our attention on density functions that generate the following CES

100

aggregator:

$$G(I, S) = \left(\nu I^{1 - \frac{1}{\eta}} + (1 - \nu) S^{1 - \frac{1}{\eta}} \right)^{\frac{1}{1 - 1/\eta}}. \tag{2.21}$$

We choose the two parameters ν and η to approximately match the steady-state consumption–tangible investment ratio and also the consumption–intangible investment ratio across all models, insofar as possible.[4] In Appendix B.3, we derive the associated density function, f, and the implied cross-sectional distribution of the book-to-market ratio of newly implemented blueprints. We show that our aggregator G, although calibrated to match aggregate moments, conforms well with the microeconomic evidence on the distribution of the book-to-market ratios of new IPO firms in the United States.

The calibrated parameter values are summarized in Table 2.5, and the steady-state moments used to calibrate the parameters are reported in Table 2.6. We solve the model using a second-order local approximation computed using the `dynare++` package. Our results are consistent with those of Borovička and Hansen (2011) , who adopt alternative numerical procedures to analyze shock-price and shock-exposure elasticities generated by our model. We also solve our models numerically using a finite element–based global approximation method to check the accuracy of the local approximation method. Overall, the two numerical solutions produce very similar results.

---

[4] Model 3 does not have intangible capital, so $E[I/J]$ is not defined. In model 2, the parameter η has only minor effects on the stochastic steady state; therefore, it is not possible to match both $E[C/I]$ and $E[I/J]$ simultaneously. In model 2, we follow the RBC literature and set ν to match the consumption–physical investment ratio observed in the data.

2.5.2 Quantity dynamics

In this section, we show that all four models produce largely similar macroeconomic quantity dynamics and that our benchmark model improves slightly upon the RBC model (model 3) along several dimensions. In this sense, our model inherits the success of the RBC models on the quantity side of the economy.

The quantity dynamics produced by our calibrations are shown in the top panel of Table 2.7. All four calibrations produce a small volatility of consumption growth and a high volatility of tangible investment growth, consistent with the data. Recall that model 3 is essentially the standard RBC model with recursive preferences. We know from Tallarini (2000) that the risk aversion parameter of the recursive preference has little effect on the quantity dynamics. Therefore, on the quantity side, the model behaves just like the standard RBC model with CRRA preferences, where $\gamma = \frac{1}{\psi} = 0.5$. The second moments generated by model 3 are consistent with those in Kydland and Prescott (1982). In particular, this model produces a small standard deviation of consumption (2.14% per year) and a standard deviation of investment about six times larger (15.33% per year).

Comparing models 2 and 3, we see that the addition of intangible capital to the standard RBC model reduces the volatility of physical investment growth. This is because the concavity of the aggregator G implies decreasing marginal production of physical investment and affects the volatility of physical investment similarly to adjustment cost functions in neoclassical models. Therefore, to generate a high volatility of tangible investment, therefore, the curvature of $G(I, S)$ needs to be low, or, equivalently, the elasticity of substitution between I and S, η, needs to

be sufficiently high. All of our calibrated models with intangible capital have this feature. Adding long-run shocks and heterogenous productivity of capital vintages increases the volatility of investment. In model 2, investment growth volatility is almost 11%, and in the benchmark model it reaches a level of 14.18%, consistent with the data.

The persistence of the growth rates of macroeconomic quantities produced by our model is similar to that in the data. In models 2 and 3, both output and consumption are autocorrelated, even if productivity growth is not. This result is generated by the persistent fluctuations of our endogenous state variables, K and S (as in Kaltenbrunner and Lochstoer 2010). The persistence generated in these two models, however, is smaller than that in the data. The addition of long-run productivity risk increases the autocorrelation of consumption and output growth rates (Croce 2008). Because both the benchmark model and model 1 feature long-run productivity shocks, they produce a higher autocorrelation in output growth (Table 2.6) and consumption growth (Table 2.7) than models 2 and 3. The introduction of long-run productivity shocks, therefore, brings our model closer to the data.

The correlation between consumption and physical investment growth rates in our benchmark calibration is consistent with its empirical pattern, that is, it is moderate at an annual frequency and high over the long horizon. This is an improvement with respect to standard RBC models, which are notorious for producing large correlations of consumption and investment growth even over short horizons. Standard RBC models have only one source of shocks, the short-run productivity shocks. Because both consumption and investment comove with this shock, the correlation of their growth rates is quite high. In contrast, our model also features news about

103

future productivity shocks that have no effect on current total output. Because of the resource constraint, consumption and total investment must move in opposite directions in response to these shocks, reducing their unconditional correlation.

In Section 2.6 and Appendix B.3, we show that extensions of our model are capable of matching a broader set of moments, including intangible investment volatility and the dynamics of hours worked.

2.5.3 Asset price dynamics

In this section we examine the asset pricing implications of our model. Although the quantity dynamics of the benchmark model inherits the basic features of the standard RBC model, thanks to the lagged risk exposure of new vintage capital, asset returns in our model respond to long-run risks similarly to endowment-based long-run risk models, for example, that of Bansal and Yaron (2004). More importantly, our model is able to produce a large spread between the expected return on tangible and intangible capital.

Campbell (2000) summarizes the challenge to general equilibrium asset pricing models as three puzzles: the equity premium puzzle (Mehra and Prescott 1985), the stock market volatility puzzle (Campbell 1999), and the risk-free rate puzzle (Weil 1989). These puzzles are even more difficult to solve in production economies, as models must (1) not only generate a pricing kernel that is sufficiently volatile but also endogenously produce a high risk exposure of the stock market returns, and (2) be consistent with the empirical evidence from the quantity side of the economy. The literature has relied primarily on adjustment cost or other forms of rigidity in investment to generate the variation in the price of physical capital. In the next

104

sections we show that it is difficult to reconcile the high risk exposure of the market returns and the high volatility of tangible investment when relying on rigidity in investment as the only means by which to generate variations in the price of physical capital. In our benchmark model, however, thanks to heterogeneous productivity of capital of different vintages, we can simultaneously produce volatile stock market returns and aggregate physical investment. The adoption of recursive preferences with high intertemporal elasticity of substitution also allows us to solve the risk-free rate puzzle.

The empirical evidence on the value premium imposes a strong discipline on general equilibrium asset pricing models with intangible capital. Stocks with high book-to-market ratios earn higher returns than stocks with low book-to-market ratios, and the difference between market value and book value can be attributed to the value of intangible capital owned by the firm. This evidence suggests that intangible capital earns a lower average return than physical capital. Qualitatively, the benchmark model and models 1 and 2 are consistent with intangible capital being less risky than physical capital (Ai 2009). Quantitatively, however, only the benchmark model is capable of producing a significant value premium. The interaction between lagged risk exposure of new vintage capital and long-run productivity risk is the main driver of this result.

In the following subsections, we first discuss the common features of all four calibrations and then examine the models' implications for the returns on physical capital r_K. Finally, we study the models' implications for the value premium. The asset pricing implications of all four calibrations are summarized in Table 2.7.

105

Common features. All calibrations, except model 1, are able to generate a low and relatively smooth, risk-free interest rate.[5] The volatilities of the risk-free interest rates are low because we adopt an intertemporal elasticity of substitution greater than one: Because agents are very willing to substitute consumption across time, fluctuations in the expected consumption growth rate produce only small variations in the equilibrium interest rate.

All four models produce a fairly high volatility of the stochastic discount factor. Because the representative agent is endowed with recursive preferences, fluctuations in expected consumption growth (long-run risk, in the language of Bansal and Yaron 2004) strongly affect marginal utility. Models 2 and 3 feature predictability in consumption growth because of the endogenous fluctuations in K and S. The introduction of long-run productivity shocks in both the benchmark model and model 1 almost doubles the volatility of the stochastic discount factor.

Investment dynamics and physical capital returns. As shown in Kaltenbrunner and Lochstoer (2010) and Croce (2008), an important challenge for the long-run risk–based asset pricing model with production is to account for the high volatility of investment and stock returns simultaneously. Although recursive preferences generate a high volatility of the stochastic discount factor, the return to physical capital is typically very smooth, unless one is willing to assume a large adjustment cost. High levels of adjustment cost, however, are typically associated with counterfactually low levels of volatility in investment growth.

[5] In standard RBC models, there is always a tension between simultaneously producing a high consumption–physical investment ratio and a low level of the risk-free rate through the subjective discount factor β. This explains why in model 3 we are not able to match the level of the risk-free rate, because we set β to reproduce the consumption-investment ratio observed in the data.

This tension is present in models 1, 2, and 3, but it is resolved in our benchmark model, where the annual volatility of the unlevered returns on physical capital is 2.00% and investment is as volatile as in an RBC model. To explain our results, we plot in Fig. 2.1 and Fig. 2.2 the impulse response functions of quantities and prices, respectively, generated by both short-run and long-run shocks in the benchmark model and model 1.

The left panels of Fig. 2.1 and Fig. 2.2 show that the introduction of heterogeneous productivity of vintage capitals does not significantly alter the model's response to short-run shocks. This result has two important implications. First, because the quantity dynamics in the benchmark model are mostly driven by short-run shocks, they inherit the success of standard RBC models with *i.i.d.* productivity growth (model 3). Second, the risk premiums associated with short-run shocks are small in both models. Therefore, to understand the success of our benchmark model in accounting for both equity and value premiums, we must focus on the interaction between long-run shocks and the heterogeneous productivity of vintage capitals.

As shown in the right-hand panels of Fig. 2.1 and Fig. 2.2, the impulse responses to long-run shocks are significantly different across model 1 and the benchmark model. With a one-standard-deviation change in the long-run productivity shock, the return on physical capital, r_K, in the benchmark model increases by about 1.5%, whereas the change in r_K in model 1 is barely visible. This implies that the exposure to the long-run productivity risk of physical capital is very small in model 1, whereas in the benchmark model it is larger by several orders of magnitude.

To explain the different behavior of r_K across the benchmark model and model 1, we focus our attention on the ex-dividend price of physical capital, $q_{K,t}$ (see Fig. 2.2,

107

fourth panel, right column). Iterating Equation (2.12) forward, we can express $q_{K,t}$ as the present value of the infinite sum of all future payoffs:

$$q_{K,t} = \sum_{j=1}^{\infty} (1 - \delta_K)^j E_t \left[\Lambda_{t,t+j} \alpha K_{t+j}^{\alpha-1} \left(A_{t+j} \right)^{1-\alpha} \right], \qquad (2.22)$$

Equation (2.22) implies that the price of physical capital, $q_{K,t}$, is the present value of the marginal product of physical capital in all future periods. This equation holds in model 1 as well. A positive innovation in the long-run productivity component x_t has two effects on the future marginal product of physical capital. The first is a direct effect: Keeping everything else constant, an increase in x_t raises the marginal product of physical capital by increasing all future A_{t+j} for $j = 1, 2, \cdots$. The second effect comes from the general equilibrium. An increase in the marginal productivity of capital also triggers more investment, which augments K_{t+j} in all future periods. Because of the decreasing returns to scale ($\alpha < 1$), an increase in K_{t+j} mitigates the direct effect.

In model 1, the elasticity of substitution between physical investment and intangible capital, η, is set to 1.4. This implies that the supply of physical investment is quite elastic. Consequently, the return on physical capital responds very little to long-run shocks. To see this point more clearly, note that without overlapping generations of vintage capital, we have $\varpi_t = 1 \quad \forall t$, and Equation (2.13) can be written as

$$q_{K,t} - (1 - \delta_S) = \frac{1}{G_I (I_t, S_t)} = \frac{1}{\nu} \left(\frac{I_t}{G (I_t, S_t)} \right)^{\frac{1}{\eta}}. \qquad (2.23)$$

By Equation (2.23), as η increases, I_t becomes more sensitive to changes in $q_{K,t}$. Equation (2.22) implies that if investment adjusts elastically to productivity shocks,

108

then the effect of the long-run productivity shock on $q_{K,t}$ is small, due to decreasing return to scale of physical capital. This intuition is confirmed by our impulse response functions. Innovations in the long-run productivity component are accompanied by a nearly permanent increase in the I/S ratio (Fig. 2.1, third panel, right column, solid line). As a result, the changes in q_K after a long-run productivity shock are almost negligible (Fig. 2.2, fourth panel, right column). To summarize, in model 1 the return on physical capital responds little to long-run productivity shocks because the direct effect on the price of physical capital is mostly offset by movements in investment (the general equilibrium effect). As with standard adjustment cost models, it is difficult to simultaneously produce a high volatility of both investment growth and returns on physical capital in model 1.

In the benchmark model, however, after a long-run productivity shock, investment rises, but after a substantial delay, whereas the return on physical capital increases immediately and sharply. The I/S ratio initially drops and then starts to rise, always staying below the level obtained in model 1 (Fig. 2.1, fourth panel, right column). The last panel in the right column of Figure 1 plots the impulse response of physical capital stock normalized by productivity ($k_t = K_t/A_t$) after a long-run shock. Because of the lagged response of investment, the level of physical capital in the benchmark model stays nearly permanently behind that obtained in model 1. Because the marginal product of capital, $\alpha k_t^{-(1-\alpha)}$, is a decreasing function of normalized capital stock, in the benchmark model the marginal product of physical capital remains almost permanently above that observed in model 1, producing a strong increase in $q_{K,t}$.

In this case, the direct and general equilibrium effects of long-run productivity

109

shocks affect $q_{K,t}$ in the same way, thereby reinforcing each other. The marginal product of capital increases both because a positive shock in x_t increases A_{t+j} in all future periods and because the sluggish response of investment to long-run shocks results in a nearly permanent reduction of physical capital stock relative to that in model 1.

To understand the lagged response of investment to long-run news in the benchmark model, note that a long-run shock increases the productivity of all existing vintages of capital almost permanently but affects the productivity of the new production units only after a delay. This generates an incentive to postpone the exercise of new growth options. As a result, a long-run productivity shock immediately produces a strong income effect (the agent anticipates a persistent increase in the productivity of all existing vintages of capital and prefers to consume more) without generating a significant substitution effect (the return on new physical investment is unaffected by long-run productivity shocks for an extended period of time). At time 1, when a positive long-run shock materializes, the net effect is an immediate increase in consumption and a decrease in investment, exactly the opposite of what happens in model 1, in which the substitution effect dominates the income effect and investment increases. This feature of the model is consistent with recent empirical findings of Barsky and Sims (2011) and Kurmann and Otrok (2010).

In the benchmark model, positive long-run shocks, although small, have quite significant and prolonged negative effects on physical investment. This sluggish response of investment is generated by the persistence of the long-run shocks: After positive long-run news, the relative productivity of new investment remains behind that of existing vintages for an extended period of time, thereby discouraging a fast

and full recovery of investment.

Value premium. We report the value premium in the data and the model in the last row of Table 2.7. In the data, *HML* is calculated as the average return of the *HML* factor as constructed by Fama and French (1995). The model counterpart of *HML* is calculated as the difference in the leveraged return on tangible and intangible capital.[6]

To understand the difference in the expected returns of tangible and intangible capital, we can use the functional form of $G(I, S)$ in Equation (2.21) and write the returns of intangible capital in Equation (2.16) as

$$r_{S,t+1} = \frac{1 - \nu}{\nu} \left(\frac{I_{t+1}}{S_{t+1}} \right)^{\frac{1}{\eta}} + (1 - \delta_S). \tag{2.24}$$

As explained in Section 2.3.2, the term $\frac{1-\nu}{\nu} \left(\frac{I_{t+1}}{S_{t+1}} \right)^{\frac{1}{\eta}}$ can be interpreted as the expected payoff of in-the-money options in period $t + 1$. Because S_{t+1} is determined in period t, innovations in the return on intangible capital respond positively to innovations in I_{t+1}. The intuition for this result is that an increase in the $\frac{I}{S}$ ratio lowers the option-exercise threshold $\theta_t^* = G_I(I_t, S_t)$ and raises the probability of option-exercise, thereby enhancing the payoff of growth options. As shown in Fig. 2.1 and 2.2, in our benchmark model, $\frac{I}{S}$ responds negatively to long-run productivity shocks. Therefore, our model is able to account for the empirical fact that growth stocks

---

[6] In our analysis, we abstract away from both financial and operative leverage. Garca-Feijo and Jorgensen's (2010) estimates suggest a degree of total leverage of four; we set leverage to three to be conservative.

are less exposed to long-run economic risks, as documented in Bansal, Dittmar, and Lundblad (2005), Hansen, Heaton, and Li (2008), and Kiku (2006).

To understand the lower exposure of growth options with respect to long-run productivity shocks compared to assets in place, note that the payoff of growth options can be replicated by long positions in assets in place and short positions in the cost of the strike asset. Exercising growth options at time t costs $\frac{1}{\theta}$ unit of investment goods and produces a generation-t production unit, which is equivalent to ϖ_{t+1} generation-0 production units. Because installed physical capital in this economy is measured in terms of generation-0 production unit equivalents, the expected cost of creating an additional unit of K_t on date t is $E_t\left[\varpi_{t+1}^{-1}\frac{1}{\theta}\right] = \frac{1}{\theta}e^{\frac{1-\alpha}{\alpha}\left(x_t+\frac{1}{2}\sigma_a^2\right)}$, $0 < \alpha < 1$. Growth options have low exposure to long-run risk, x_t, because the cost of exercising them, $E_t\left[\varpi_{t+1}^{-1}\frac{1}{\theta}\right]$, covaries positively with x_t and acts as a hedge. Good news for the productivity of existing productions units is bad news for unimplemented blueprints because it is more expensive to create new production units as productive as those of old vintages. As a result, both physical investment and option returns respond negatively to long-run productivity shocks.

The implications of our model for the value premium are summarized in the bottom panel of Table 2.7. We make the following observations. First, all models with intangible capital yield a higher return for physical capital than for intangible capital. Second, despite the introduction of long-run risk, model 1 produces a lower spread between physical and intangible capital than does model 2. In model 1, intangible capital is more exposed to long-run risk than is tangible capital. Specifically, without heterogeneous productivity of vintage capital, after a positive long-run productivity shock, physical investment increases sharply but $q_{K,t}$ remains almost flat (Fig. 2.2).

112

At the same time, the increase in the I/S ratio is associated with a drop in the option exercise threshold, $\theta^*(t)$, and a positive innovation in the return on intangible capital. As a result, as shown in Table 7, simply adding long-run productivity shocks to model 2 increases the market risk premium only slightly and eliminates most of the spread in the expected return on physical and intangible capital.

Third, compared with model 1, our benchmark model produces both a larger risk premium on physical capital and a smaller one on intangible capital, thus improving on equity and value premiums simultaneously. The heterogeneous productivity of vintage capital is responsible for both improvements because it causes the I/S ratio to drop after good long-run news. This feature of the model produces a sharp increase in the return on tangible capital, r_K, and a drop in the intangible capital return, r_S. It increases the riskiness of physical capital and makes growth options an insurance device against long-run risk. Overall, the benchmark model produces a market risk premium more than two times larger than that of model 1 and a spread between tangible and intangible capital returns larger by an order of magnitude.

We conclude our discussion on value premium by exploring the implications of our model for the cash-flow duration of book-to-market-sorted portfolios. We define the Macaulay duration, MD_t, of a stochastic cash flow process, CF_t, as:

$$MD_t = \frac{\sum_{s=1}^{\infty} s \cdot E_t \left[\Lambda_{t,t+s} CF_{t+s} \right]}{\sum_{s=1}^{\infty} E_t \left[\Lambda_{t,t+s} CF_{t+s} \right]}. \tag{2.25}$$

We provide the details of the calculation of the duration of growth options and assets in place in Appendix B.3. Here, we point out that options typically have longer duration than assets in place because they start paying cash flows only after being exercised and becoming assets in place. Under our benchmark calibration, the

113

Macaulay duration of assets in place (seventeen years) is about half of that of growth options (thirty years).

Because in our model value stocks are intensive in assets in place and growth stocks are intensive in options, our framework is consistent with the inverse relationship between cash-flow duration and book-to-market characteristics documented by Dechow, Sloan, and Soliman (2004). This feature of our model stands in contrast with previous results in the real-option literature. In Gomes, Kogan, and Zhang (2003), for example, value stocks are option intensive and therefore have longer cash-flow duration than low book-to-market stocks.

2.5.4   Additional testable implications of the model

In this section, we conduct econometric analyses on model predictions that directly link asset prices to macroeconomic fundamentals. First, we provide supporting evidence on the response of both investment growth and the spread between tangible and intangible capital returns to productivity news shocks. Second, we study the correlation of investment leads and lags with aggregate stock market returns as well as the spread between tangible and intangible capital returns.

Response to news shocks.   The key asset pricing implications of our model rely on the exposure of asset returns to long-run risk, that is, news about future productivity shocks. Our production-based general equilibrium framework links risk exposure to the response of macroeconomic quantities to these shocks. As we discuss in Section 2.5.3, a positive news shock is accompanied by a sharp increase in the spread of the returns of tangible and intangible capital. On the quantity side, it leads to an

114

immediate decrease in aggregate investment and a corresponding rise in aggregate consumption without affecting total output. We test these conditional responses in the data by jointly estimating Equation (2.1) and the following system of equations through a GMM procedure:

$$x_t = \beta_{rf}^{da} \cdot r_{f,t-1} + \beta_{pd}^{da} \cdot pd_{t-1} \tag{2.26}$$

$$HML_t = \beta_{SR}^{HML} \cdot \varepsilon_{a,t} + \beta_{LR}^{HML} \cdot \varepsilon_{x,t} + \varepsilon_t^{HML} \tag{2.27}$$

$$CY_t = \beta_{SR}^{C/Y} \cdot \varepsilon_{a,t} + \beta_{LR}^{C/Y_M} \cdot \varepsilon_{x,t} + \beta_{(-1)}^{C/Y_M} \cdot CY_{t-1} + \varepsilon_t^{C/Y_M}, \tag{2.28}$$

where $CY_t = \frac{C_t}{C_t + I_t}$ is the consumption-output ratio, or consumption share. In Equation (2.26), we follow Bansal, Kiku, and Yaron (2007a) and use the risk-free rate and the log price-dividend ratio to identify news shocks. Equation (2.27) is also consistent with our model: The spread between the value and growth portfolios, HML_t, depends on the realization of short-run and long-run productivity shocks, as well as an error term, ε_t^{HML}. In Equation (2.28), we use $\beta_{SR}^{C/Y}$ and $\beta_{LR}^{C/Y}$ to denote the sensitivity of the consumption-output ratio with respect to short-run and long-run productivity shocks, respectively. Consistent with our model, the consumption share process is very persistent in the data. Instead of estimating a full-blown DSGE model with K_t and S_t as state variables, we use the lagged value, CY_{t-1}, to control for the history dependence of the consumption share process.[7]

In Table 2.8, we report our results for three different measures of aggregate productivity. In the first row, we compute aggregate productivity according to Equation

---

[7] We have also estimated a version of Equation (2.28) that includes investment-specific shocks, an alternative determinant of the consumption-output ratio discussed in the literature. Our results are robust to this extension. We thank Dimitris Papanikolaou for sharing his data on investment-specific shocks.

(2.5). We set $\alpha = 0.3$ and assume an inelastic labor supply, $N_t = 1$, as in our benchmark model. In the second regression, we allow for changes in aggregate labor. Data on labor and physical capital are from the NIPA tables and are described in Appendix B.2. In our third specification, we use a multifactor-adjusted measure of productivity directly provided by the BLS. The first two measures of productivity can be computed starting from 1930, but the BLS productivity data are available only for the post–World War II period. Across all specifications, the identification of the long-run predictable component in productivity growth is statistically significant, as shown by the small p-value of the Wald statistics. In addition, both HML and the consumption share respond positively to the identified news shocks, as predicted by our model. The response of HML is positive and statistically significant across all specifications. The response of consumption share, in contrast, is statistically significant in regressions (1) and (2) when longer samples of the productivity data are available.

In addition, we test the implications of our model for the responses of returns and quantities to news about the differences in productivity of young and old firms. We construct the productivity difference as the ratio of the average productivity of all firms in the second to fourth age quartiles and that of the youngest 25% of firms in our data set (see Section 2.4). We estimate equations (2.1) and (2.26)–(2.28), with Δa_t replaced by the log difference of productivity. Because the productivity difference is not perfectly correlated with aggregate productivity in the data, this estimation exercise provides yet another way to empirically test the model.

We report the results of our estimation in row (4) of Table 2.8, where firm-level productivity is computed as in Equation (2.18) but assuming an inelastic labor

supply. In row (5) of the same table, we allow for variations in firm-level labor as in Section 2.4. Consistent with our model, both HML and the consumption share drop upon the arrival of good news for the relative productivity of young and old firms.

Leads and lags. The economic mechanism in the benchmark model has strong implications for the correlation of tangible investment with the market return and the spread between tangible and intangible capital returns. We plot these correlations in Fig. 2.3 for both our benchmark model and model 1 and show that our benchmark model fits the correlation patterns in the data well.

The left panel of Fig. 2.3 plots the cross-correlations between the market excess returns, $r^{ex}_{m,t+1}$, and leads ($j > 0$) and lags ($j < 0$) of tangible investment growth rates, ΔI_{t+1}. Consistent with the data, in our benchmark model the contemporaneous correlation between investment growth and excess returns is close to zero. This is the result of two offsetting effects. On the one hand, just like in the standard RBC model, a positive short-run productivity shock triggers a positive comovement of the market return and investment growth. On the other hand, a positive long-run productivity shock boosts the market return but discourages current-period investment. In contrast, in model 1 long-run productivity shocks induce positive comovements between investment growth and market returns and reinforce the effect originated from short-run shocks. As a result, model 1 produces a counterfactually large contemporaneous correlation of investment and market return.

In addition, similarly to the data, in our benchmark model the correlation reaches a peak for one-period-ahead investment and dies off at longer horizons. The correla-

tion's surge at $j = 1$ is generated by two effects that reinforce each other. First, upon the realization of a positive short-run shock at time t, both intangible investment, J_t, and therefore S_{t+1}, rise. At time $t + 1$, because tangible investment and intangible capital are complements, it is optimal to increase further tangible investment, I_{t+1}. Hence, a positive excess return at time t predicts a rise in investment at time $t + 1$. Second, upon the realization of a positive long-run shock at time t, there is an immediate spike in the market excess return and a fall in physical investment followed by sluggish investment growth. Long-run shocks reinforce the fact that positive market excess returns at time t predict future positive investment growth starting from time $t+1$. Because these quantity dynamics die off over time, their effects on long-horizon correlations taper as well. In contrast, in model 1 the correlation between current excess returns and future investment growth is too high when $j = 1$, and it quickly becomes negative at longer horizons.

In the right panel of Fig. 2.3, we plot the cross-correlations between the return on the value-minus-growth portfolio, *HML*, and leads and lags of investment growth. Consistent with the data, the correlation between investment growth and *HML* is low for $j = 0$ and increases gradually over longer horizons. Note that the returns on tangible and intangible capital move in the same direction after short-run shocks, but in opposite directions following long-run shocks; therefore, the *HML* return mainly reflects realizations of long-run productivity shocks in our model. As already noted, positive long-run shocks induce a small contemporaneous drop in physical investment followed by prolonged investment growth. Therefore, HML predicts future investment growth even though it has a negative contemporaneous correlation with investment.

118

In contrast, in model 1, the correlation between investment and HML return is too high for $j = 0, 1$ and quickly becomes negative at longer horizons. Overall, in our benchmark model, the correlations are consistently within the 95% confidence interval bands estimated from the data. We view the empirical evidence presented in this section as strongly supporting the economic mechanism emphasized by our model.

2.6   Adjustment Costs of Intangibles

As shown in Fig. 2.4, the impulse responses of both tangible and intangible investment in the benchmark model contain a periodic component. As a result, in our benchmark model the volatility of intangible investment is about two times that of tangible investment, whereas this number is about one-half in the data (Table 2.9).

Here, we introduce adjustment costs on the accumulation of intangible capital. This modification eliminates the periodic component in both tangible and intangible investments and makes their volatility consistent with the data. Specifically, we replace the law of motion of intangible capital in Equation (2.8) by the following expression,

$$S_{t+1} = (1 - \delta_S)(S_t - G(I_t, S_t)) + H(J_t, K_t), \qquad (2.29)$$

and parameterize H in the spirit of Jermann (1998),

$$H(J, K) = \left[\frac{a_1}{1 - 1/\xi} \left(\frac{J}{K} \right)^{1 - 1/\xi} + a_2 \right] K.$$

We calibrate the parameter ξ to match the volatility of intangible investment. Once ξ is chosen, the parameters $\{a_1, a_2\}$ are pinned down by the following two steady-state

119

conditions: $H(\overline{J}, \overline{K}) = \overline{J}$ and $H_J(\overline{J}, \overline{K}) = 1$, where $\overline{J}$ and $\overline{K}$ denote the steady-state levels of intangible investment and tangible capital stock, respectively. In Appendix B.3, we provide a microeconomic foundation for our specification of the adjustment cost function and prove that it arises as the result of a concave production function of new blueprints.

Given this modification of the model, the equilibrium conditions (2.14)–(2.16) are replaced by

$$q_{S,t} = 1/H_{J,t}$$

$$p_{K,t} = \alpha K_t^{\alpha-1} (A_t N_t)^{1-\alpha} + H_{K,t} q_{S,t} + (1 - \delta_K) q_{K,t}$$

$$r_{K,t+1} = \frac{\alpha K_{t+1}^{\alpha-1} (A_{t+1} N_{t+1})^{1-\alpha} + H_{K,t} q_{s,t} + (1 - \delta_K) q_{K,t+1}}{q_{K,t}},$$

$$r_{S,t+1} = \frac{G_S (I_{t+1}, S_{t+1})}{G_I (I_{t+1}, S_{t+1})} + (1 - \delta_S) q_{S,t},$$

where H_J and H_K denote the partial derivative of H with respect to J and K, respectively.

We highlight three main results. First, the impulse response functions of tangible and intangible investment in the model with adjustment costs are smooth (Fig. 2.4). Second, the incorporation of adjustment costs raises the volatility of physical investment and lowers that of intangible investment, consistent with the data (Table 2.9). Third, in the model with adjustment costs the behavior of consumption growth and returns on both tangible and intangible capital remains similar to that in the benchmark model. As a result, the implications of benchmark model for the equity premium, the value premium and the volatility of consumption growth are largely unaffected by this extension.

Not surprisingly, the conditional CAPM fails in our model because the true stochastic discount factor is a linear combination of both long-run and short-run shocks and is imperfectly correlated with the market return. Indeed, the stochastic discount factor in our benchmark model has a higher loading on long-run productivity shocks than the market return. As a result, assets with higher loadings on long-run risk yield higher alphas in our CAPM regressions. As shown in Table 2.9, physical capital not only has a higher unconditional average return than intangible capital but also a higher CAPM alpha. This pattern is more pronounced in the model with adjustment cost, where the market return is driven even more by short-run shocks and represents a worse proxy for the true stochastic discount factor.

We conclude this section with a brief explanation of the origin of the periodic component and the reason it disappears with adjustment costs. Because of the complementarity between intangible capital and physical investment, $G_{IS}(I_t, S_t) > 0$, the agent has a strong incentive to substantially increase physical investment only after more intangible capital is built up. At time t, however, the stock of intangible capital, S_t, is one-period predetermined and cannot simultaneously adjust with tangible investment, I_t. For this reason, the agent finds it optimal to proceed in alternating steps.

For the sake of simplicity, let us focus on a positive short-run productivity shock at time $t = 1$ (Fig. 2.4, left column). Upon the realization of the shock, physical investment is partially delayed and intangible investment, J_1, is immediately adjusted in order to reach a higher level of intangible stock, S_2. At time $t = 2$, physical investment is efficiently increased and intangible capital is partially depleted. At time $t = 3$, more intangible investment is needed to replenish the stock of blueprints.

As a result, in the third period physical investment is dampened, and intangible investment surges again. This pattern continues until it converges back to the steady state.

Note that the above investment policy requires large adjustments in intangible investment, J, and produces an intangible investment growth volatility of 35% per year. Even in the presence of mild adjustment costs, such large changes in intangible capital become very costly. For this reason, with adjustment costs the increase in intangible investment becomes gradual and persistent, whereas physical investment immediately spikes upon the realization of the productivity shock.

In Appendix B.3, we consider two other extensions of our model. First, we add an endogenous labor supply. Second, we consider more general specifications of the ϕ_j process that governs the heterogeneity of firms' exposure to aggregate shocks. We show that our main results are preserved and often enhanced in these more general settings. Because the incorporation of adjustment costs significantly improves our results, we keep this feature in the other two extensions as well.

2.7   Conclusion

In this study, we present a general equilibrium asset pricing model with long-run productivity shocks as in Croce (2008) and intangible capital modeled as storable investment options as in Ai (2009). We document that in the data, new investment is less exposed to aggregate productivity shocks than is capital of older vintages. We incorporate this feature in our model and show that the lower exposure of new investment is quantitatively important in accounting for (1) the high equity premium, (2) the high volatility of the stock market return, and (3) the large spread in both

expected returns and cash-flow duration across book-to-market-sorted portfolios in the data.

Several remarks are in order. First, as in Ai (2009), we have allowed idiosyncratic shocks to the quality of investment options. In our setting, we have assumed that these shocks are *i.i.d.* Although unrealistic, this assumption simplifies our aggregation results, making our model very tractable and enabling us to avoid the need to keep track of the cross-sectional distribution of option quality. Allowing for more general processes of the quality of the options is a fruitful extension that we leave for future research.

Second, considering a more general setting with heterogeneous firms will allow us to implement portfolio sorting exercises in the context of our current model, to study firms' transition among value and growth portfolios, and to confront the model with a wider set of empirical evidence at the portfolio level, as done by Ai and Kiku (2012). Based on their insights, we are optimistic that the basic intuition in this article will remain valid even with heterogeneous firms.

Finally, we believe that our model provides a valuable general equilibrium framework for the measurement of intangible capital by exploring the information from both the quantity and pricing sides of the economy. Specifically, a structural estimation of our DSGE model employing both time-series data on macroeconomic aggregates and cross-section data on equity returns may shed new light on the accumulation of intangibles in the United States.

Table 2.1: Summary statistics by firm age quantiles

Firm age quantile	Median firm age	Median capital age (T=5)	Median capital age (T=8)	Median capital age (T=15)
1	8	2.55	3.37	4.47
2	16	2.61	3.54	4.84
3	29	2.66	3.63	5.06
4	91	2.71	3.76	5.41
All Firms	24	2.64	3.60	5.01
p-Value	0.00	0.00	0.00	0.00
No. Firms	8,084	8084	6,014	2,937
No. obs.	83,089	88,283	64,871	32,239

This table reports the summary statistics of our panel. The sample ranges from 1950 to 2008 and includes approximately 8,084 different firms, for a total of 83,089 observations grouped into four firm-age quantiles. Firm age is expressed in years and is computed using founding dates from Ritter and Loughran (2004) and Jovanovic and Rousseau (2001). Capital age is computed according to Equation (2.17). We report the p-value for the null hypothesis that the average age in the fourth quartile is equal to the average age in the first quartile. The null is rejected across all of our measures at the 1% confidence level. The last two rows report the number of firms and observations available for different measures of age.

Table 2.2: Exposure to aggregate risk by firm age

Regression	$\Delta \ln \overline{A}$	AGE	$AGE * \Delta \ln \overline{A}$	B/M	Obs.	Firms
(1)	-0.042	-0.002***	0.012***	-0.005	70,909	7,335
	(0.216)	(0.000)	(0.003)	(0.004)		
(2)	0.880*	-0.003***	0.018***	-0.046***	22,432	4,023
	(0.533)	(0.001)	(0.007)	(0.007)		
(3)	0.383	-0.002***	0.011**	-0.006	59,395	7,226
	(0.319)	(0.000)	(0.004)	(0.005)		

This table reports firms' risk exposure by age. All estimates are based on the following second-stage regression: $\Delta \ln A_{i,j,t} = \xi_{0i} + \xi_1 \Delta \ln \overline{A}_t + \xi_2 AGE_{i,j,t} + \xi_3 AGE_{i,j,t} \cdot \Delta \ln \overline{A}_t + B/M_{i,j,t} + \varepsilon_{i,j,t}$. Regression (1) is obtained using the whole sample. To control for exit bias, in regression (2) we use the inverse Mills ratio (IMR) as an additional explanatory variable. In regression (3) we exclude the years with negative aggregate productivity growth. All the estimation details are reported in Appendix B.2. Numbers in parentheses are standard errors. We use $*$, $**$, and $***$ to indicate p-values smaller than 0.10, 0.05, and 0.01, respectively.

Table 2.3: Exposure to aggregate risk of young versus other firms

Regression	ϕ				Obs.	
	Young	Other	OMY		Young	Other
(1)	-0.484	0.963***	1.447***		15,030	55,879
	(0.373)	(0.100)	(0.361)			
(2)	-0.325	1.202***	1.527*		5,015	17,417
	(0.837)	(0.416)	(0.908)			
(3)	-0.727	1.501***	2.228***		12,721	46,674
	(0.618)	(0.177)	(0.597)			

Notes - This table reports risk exposure of Young and Other firms. In each sample period, a firm is classified as Young if it belongs to the set of the 25% youngest firms; otherwise it is classified in the group Other. All estimates are based on the following second-stage regression (equation (2.20)):

$$\Delta \ln A_{ijt} = \begin{cases} \xi_{0i} + \phi_Y \Delta \ln \overline{A}_t + \xi_{1i} B/M_{i,j,t} + \widetilde{\varepsilon}_{i,j,t} & i \in \text{Young} \\ \xi_{0i} + \phi_O \Delta \ln \overline{A}_t + \xi_{1i} B/M_{i,j,t} + \widetilde{\varepsilon}_{i,j,t} & \text{otherwise.} \end{cases}$$

OMY refers to $\phi_O - \phi_Y$. Regression (1) is obtained using the whole sample. To control for exit bias, in regression (2) we add the inverse Mills ratio (IMR). In regression (3) we exclude the years with negative aggregate productivity growth. All the estimation details are reported in Appendix B.2. Numbers in parentheses are standard errors. We use $*$, $**$, and $***$ to indicate p-values smaller than 0.10, 0.05, and 0.01, respectively.

Table 2.4: Main components of our economy

	Benchmark	Model 1	Model 2	Model 3
Vintage capital	Yes ($\phi_0 = 0$)	No ($\phi_0 = 1$)	No ($\phi_0 = 1$)	No ($\phi_0 = 1$)
Long-run productivity risk	Yes ($\sigma_x \neq 0$)	Yes ($\sigma_x \neq 0$)	No ($\sigma_x = 0$)	No ($\sigma_x = 0$)
Intangible capital	Yes	Yes	Yes	No
Recursive preferences	Yes	Yes	Yes	Yes

This table summarizes the main components active in each of our four models. All parameter values are reported in Table 2.5.

127

Table 2.5: Calibrated parameter values

Model:		Benchmark	Model 1	Model 2	Model 3
Preference parameters					
Discount factor	β	0.97	0.97	0.98	0.89
Risk aversion	γ	10	10	10	10
Intertemporal elasticity of substitution	ψ	2.0	2.0	2.0	2.0
Production function/Aggregator parameters					
Capital share	α	0.3	0.3	0.3	0.3
Depreciation rate of physical capital	δ_K	11%	11%	11%	11%
Depreciation rate of intangible capital	δ_S	11%	11%	11%	–
Weight on physical investment	ν	0.84	0.79	0.815	–
Elasticity of substitution	η	1.40	1.40	1.75	–
Total factor productivity parameters					
Average growth rate	μ	2.0%	2.0%	2.0%	2.0%
Volatility of short-run risk	σ_a	5.08%	6.30%	7.30%	5.00%
Volatility of long-run risk	σ_x	0.86%	0.80%	–	–
Autocorrelation of expected growth	ρ	0.925	0.925	–	–
Risk exposure of new investment	ϕ_0	0	1	1	1

This table reports the parameter values used for our calibrations. The following parameters are common across all models: risk aversion, γ; intertemporal elasticity of substitution, ψ; capital share, α; depreciation rates, δ_K and δ_S; average productivity growth rate, μ. We choose the rest of the parameters to match the moments reported in Table 2.6 whenever possible. All models are calibrated at an annual frequency.

Table 2.6: Moments used for model calibration

	Data	Benchmark	Model 1	Model 2	Model 3
$E\left[C/I\right]$	5.62	5.60	5.54	5.63	5.69
$E\left[I/J\right]$	1.00	1.01	0.95	(0.77)	—
$\sigma\left[\Delta\ln Y_M\right]$	3.49	3.49	3.49	3.49	3.52
$AC1\left[\Delta\ln Y_M\right]$	0.45	0.25	(0.60)	0.46	0.13
$E\left[r_f\right]$	0.86	0.80	0.87	0.86	(12.65)

This table reports the moments used to calibrate the parameters of the models evaluated in this paper. Our database refers to U.S. annual data from 1930 to 2003 (see Appendix B in the Supplementary Data). All moments that cannot be matched are in parentheses. In Model 1, the autocorrelation of measured output, $Y_M \equiv C + I$, is too high. In Model 2, the parameter ν is set to match the C/I ratio, even though the implied I/J ratio is lower than in the data. In Model 3, the discount factor β is chosen to match the steady-state consumption-investment ratio, even though this choice makes the risk-free interest rate too high.

Table 2.7: Quantities and prices

Moments	Data	Benchmark	Model 1	Model 2	Model 3
$\sigma\left(\Delta\ln C\right)$	02.53 (00.56)	02.60	02.85	02.69	02.14
$\sigma\left(\Delta\ln I\right)$	16.40 (03.24)	14.18	10.90	08.85	15.33
$AC_1\left(\Delta\ln C\right)$	00.49 (00.15)	00.48	00.68	00.48	00.59
$\rho\left(\Delta\ln C,\Delta\ln I\right)$	00.39 (00.29)	00.17	00.56	00.77	00.59
$\rho\left(\Delta\ln C_{10},\Delta\ln I_{10}\right)$	00.62 (00.24)	00.73	00.82	00.83	00.72
$\sigma\left[SDF\right]$		87.98	90.94	73.50	43.21
$E[r_K - r_f]$		01.95	00.80	00.82	00.31
$\sigma[r_K]$		01.99	01.26	01.27	00.92
$E[r_S - r_f]$		00.54	00.74	00.47	–
$\sigma[r_S]$		00.88	00.88	00.79	–
$\sigma[r_f]$	00.97 (00.31)	01.00	00.96	00.70	00.65
$E[r_M^L - rf]$	05.71 (02.25)	05.20	02.37	02.31	00.83
$E[r_K^L - r_S^L]$	04.32 (01.39)	04.20	00.17	01.05	–

All figures are multiplied by 100, except contemporaneous correlations (denoted by ρ) and first-order autocorrelations (denoted by AC_1). Empirical moments are computed using U.S. annual data from 1930 to 2003. Numbers in parentheses are GMM Newey-West adjusted standard errors. $\Delta\log C_{10}$ and $\Delta\log I_{10}$ denote the ten-year growth rate of consumption and investment, respectively. $E\left[r_K^L - r_S^L\right]$ measures the average difference between the levered returns of tangible and intangible capital. We use the HML Fama-French factor as an empirical counterpart of $r_K^L - r_S^L$. r_M^L indicates levered market returns. Returns are in log units, and the leverage is three (Feijo and Jorgensen 2010). All of the parameters are calibrated as in Table 2.5. The entries for the models are obtained by repetitions of small-sample simulations.

Table 2.8: Conditional responses of consumption and HML

	Productivity	Sample	Sign implied by the model				Wald Test
			$\beta_{SR}^{C/Y} < 0$	$\beta_{LR}^{C/Y} > 0$	$\beta_{SR}^{HML} > 0$	$\beta_{LR}^{HML} > 0$	$p-$value
(1)	Capital adjusted (BEA)	1930–2006	0.000***	0.034**	0.414	0.001***	0.000***
(2)	Capital and labor adjusted (BEA)	1930–2006	0.999	0.017**	0.000***	0.000***	0.000***
(3)	Multifactor adjusted (BLS)	1949–2006	0.000***	0.487	0.101*	0.000***	0.000***
(4)	O-Y capital adjusted	1951–2006	0.850	0.000***	0.565	0.000***	0.000***
(5)	O-Y capital and labor adjusted	1951–2006	0.261	0.261	0.355	0.000***	0.000***

We jointly estimate Equations (2.1) and (2.26)–(2.28) and study the sign of $\beta_{SR}^{C/Y}$, $\beta_{LR}^{C/Y}$, β_{SR}^{HML}, and β_{LR}^{HML}, that is, the contemporaneous sensitivity of consumption share and HML to both short- and long-run productivity shocks. We test the null that the sign of each coefficient is opposite that suggested by our model and report the associated p-value. For example, because the model suggests that $\beta_{LR}^{HML} > 0$, we test $H_0 : \beta_{LR}^{HML} < 0$. Small p-values indicate a sign consistent with our model. In the last column, we report the p-value of the Wald statistics that tests the null of no predictability in productivity growth. We use *, **, and *** to indicate p-values smaller than 0.10, 0.05, and 0.01, respectively. All p-values are based on GMM standard errors. The first three regressions are based on aggregate measures of productivity. The last two regressions are based on the productivity differential between Young and Other firms as defined in Section 2.5.4 to proxy for $-\varpi_t$.

Table 2.9: Adjustment costs on intangible capital

	$\sigma_{\Delta C}$	$\sigma_{\Delta I}/\sigma_{\Delta C}$	$\sigma_{\Delta J}/\sigma_{\Delta I}$	$E[r_M^L - r^f]$	$E[r_K^L - r_S^L]$	$\alpha_K - \alpha_S$
Data	02.53	05.29	00.50	05.71	04.32	04.01
	(00.56)	(00.50)	(00.07)	(02.25)	(01.39)	(01.77)
Bench.	02.60	05.40	02.50	05.20	04.20	00.40
Ext. 1	02.53	06.40	00.40	04.55	04.16	04.91

All figures are multiplied by 100. Empirical moments are computed using U.S. annual data in log units. Numbers in parentheses are GMM Newey-West adjusted standard errors. $E\left[r_K^L - r_S^L\right]$ and $E\left[r_M^L - r^f\right]$ measure the levered spread between tangible and intangible capital returns, and the market premium, respectively. The leverage coefficient is three (Feijo and Jorgensen 2010). The difference in the intercept of the CAPM regression for tangible and intangible returns is denoted by $\alpha_K - \alpha_S$. The entries for the models are obtained by repetitions of small-sample simulations. Extension 1 features adjustment costs on intangible investment. All the parameters for Extension 1 are reported in Table 2A, Appendix B.3.
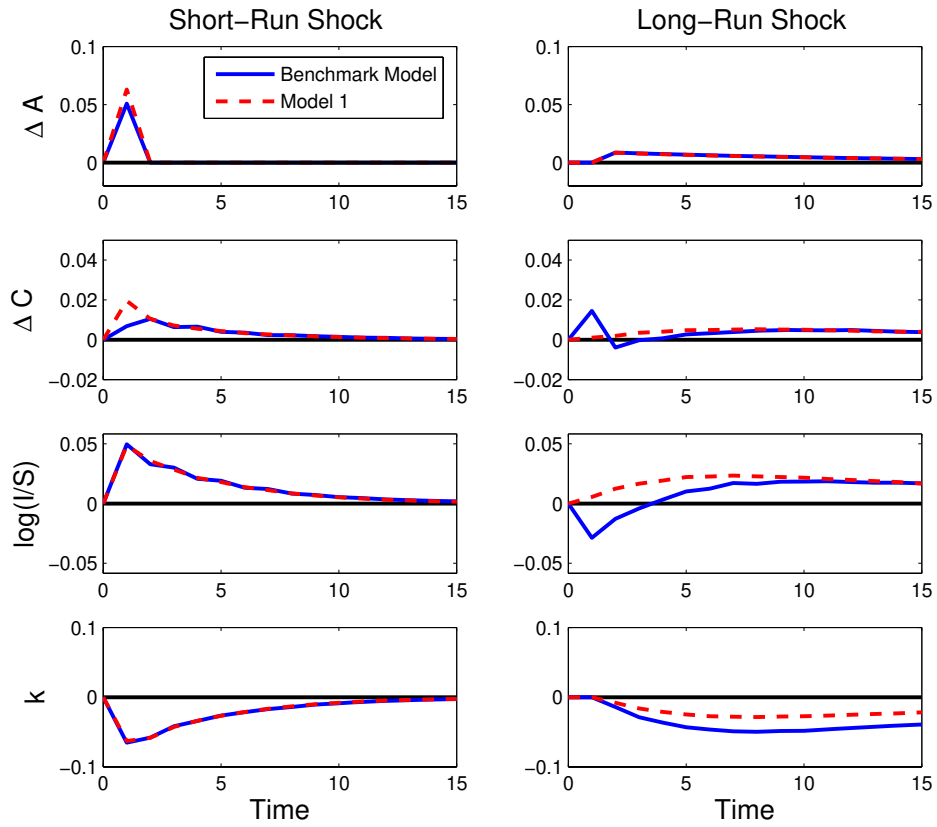
FIGURE 2.1: Impulse response functions for quantities

This figure shows annual log-deviations from the steady state. All the parameters are calibrated to the values reported in Table 2.5. The dashed lines refer to model 1; the solid lines refer to the benchmark model.
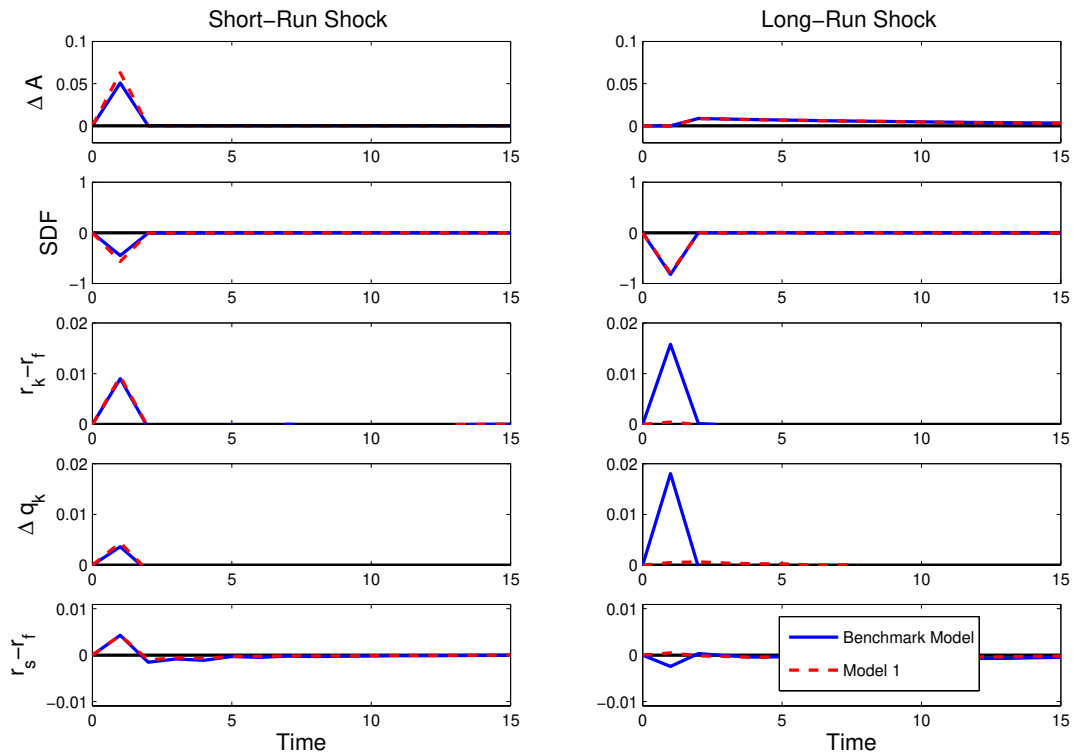
FIGURE 2.2: Impulse response functions for prices

This figure shows annual log-deviations from the steady state. All the parameters are calibrated to the values reported in Table 2.5. Returns are not levered. The dashed lines refer to model 1; the solid lines refer to the benchmark model.
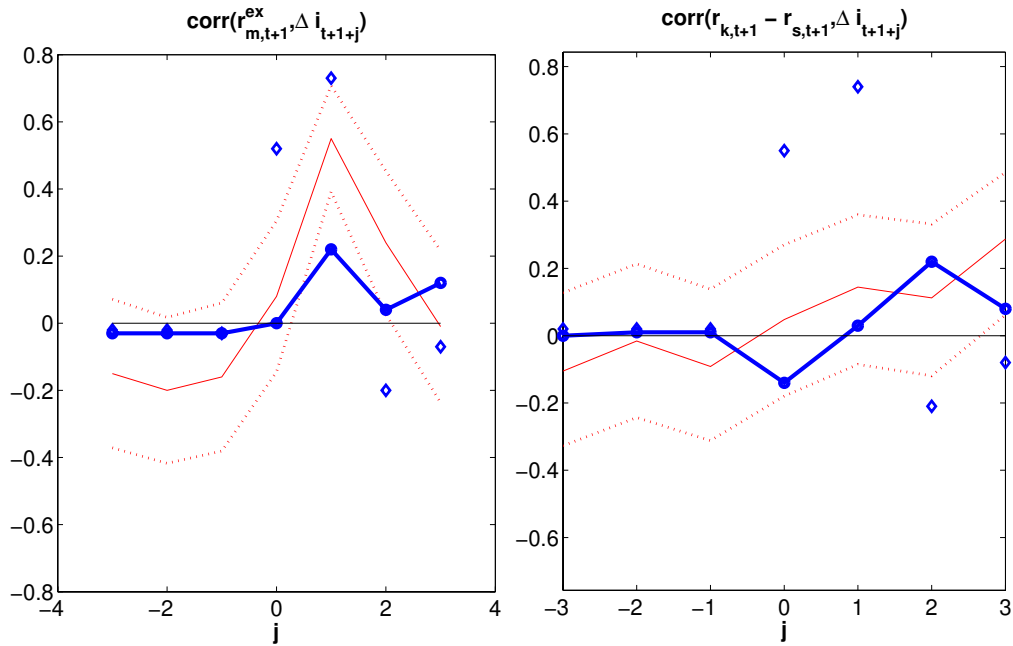
FIGURE 2.3: Returns and investment growth leads and lags

This figure shows the correlation of market excess returns (left) and the spread between the returns of tangible and intangible capital (right) with investment growth leads ($j > 0$) and lags ($j < 0$). The thin solid line represents the point estimate of the correlations computed using U.S. data from 1930 to 2003. The spread between tangible and intangible capital is proxied by the HML factor. The dotted lines mark the 95% confidence interval for the correlations. The solid line with circles represents the correlations obtained in the benchmark model. The diamond-shaped markers refer to model 1. All the parameters are calibrated to the values reported in Table 2.5. The entries from the models are obtained through repetitions of small-sample simulations.
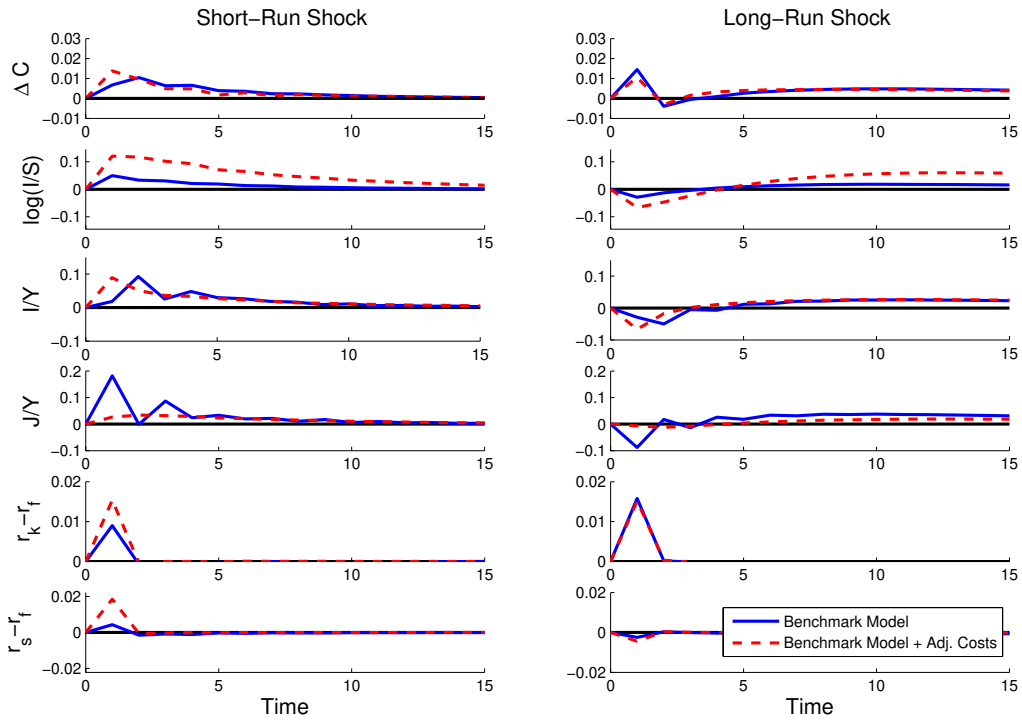
FIGURE 2.4: The Role of adjustment costs on intangible capital

This figure shows annual log-deviations from the steady state. Returns are not levered. For the model with adjustment costs, all the parameters are calibrated to the values reported in Table 2A, Appendix B.3.

3

Regime Shifts in a Long-Run Risks Model of U.S. Stock and Treasury Bond Markets

3.1 Introduction

Stocks and nominal bonds are two primary asset classes on investors' portfolio menu. It is important to have a general equilibrium model to provide a coherent explanation of the risks and returns of these two markets simultaneously. The absence of arbitrage opportunities implies that cross-market restrictions should be respected in any such models. Viceira (2010) and Campbell, Sunderam and Viceira (2010) document the empirical evidence on stochastic correlation between stock and bond returns. Fig. 3.1 plots the time-varying correlation between stock and 10-year nominal bond returns, which is calculated based on a 3-year centered moving window of monthly real returns. As in the figure, the correlation displays tremendous fluctuations, and also occasionally switches sign. Specifically, the correlation is usually positive, how-

137

ever, in periods like 1930's great depression and 2000's global financial crisis, the Treasury bond performed as hedges for stock returns. Based on the CAPM, these movements are significant enough to cause substantial changes (even switching signs) in the risk premium on Treasury bond. It has important implications for investors since the nominal bond risks are changing, rather than being constant as often assumed in traditional portfolio choice theory. Despite tremendous progress in the general equilibrium to model risks and returns for bond or stock market separately, very few has taken into account of the joint behaviors of these two asset classes.

This paper studies the joint determinants of stock and bond returns in Bansal and Yaron (2004) type of long-run risks (LRR) model framework. This framework features a recursive preference for early resolution of uncertainty, low frequency movements in both expected consumption and expected inflation, and time-varying consumption and inflation volatilities. Beyond these, an additional novel feature is that I allow for regime shifts in consumption and inflation dynamics – in particular, the means, volatilities, and the correlation structure of consumption and inflation dynamics are regime-dependent.

This rational expectations general equilibrium model framework can (1) jointly match the dynamics of consumption, inflation and cash flow; (2) generate time-varying and switching sign of stock and bond correlations, as well as switching signs of bond risk premium; (3) quantitatively reproduce another long list of salient empirical features in stock and bond markets, including time-varying equity and bond return premia, regime shifts in real and nominal yield curve across business cycles, the violation of expectations hypothesis of bond returns.

This paper broadly classifies the economy into three regimes: expansion, contrac-

tion and deep-recession. The expansion regime features a high consumption growth, a medium level of inflation, and low uncertainty (which is measured by consumption and inflation volatilities). In the contraction regime, the growth rate is lower, the uncertainty and the inflation level are both higher. One can think this regime as a stagflation regime, in which low growth and high inflation coexist. A typical sample episode is the period of late 1970's and early 1980's. The deep-recession regime features the lowest growth level and the highest uncertainty level. As opposed to the regular contraction, this regime has very low inflation, since deflation rather than inflation is more of a concern at this time. A key ingredient that is different across three regimes is the **nominal-real correlation**, in particular, it refers to the correlation between shocks to expected growth and expected inflation factors. In the first two regimes, positive news to expected inflation factor indicates a lower future expected growth; however, in the deep recession regime, the relationship is just the opposite. I provide empirical evidence to support this channel in Section 3.2. This ingredient is critical to generate tremendous movements (and potentially switching signs) of nominal bond risk premium as well as a stock-bond correlation.

I use a regime switching dynamic correlation ($RSDC$) model by Pelletier (2006) to specify the correlation structure between expected growth and inflation shocks, in particular, the correlation is constant within each regime, however, it becomes different and even switches sign across different regimes (i.e. switching signs of nominal-real correlation). This setup leads to switching sign of market price of long-run inflation risks, the magnitude of which is magnified by high persistence of expected growth and expected inflation factors. This feature, therefore, quantitatively generates switching sign of nominal bond risk premium. In the meantime, this correlation

139

structure also generates time-varying and switching sign of stock and bond correlation, consistent with the empirical evidence. Campell, Sunderam and Viceira (2010) and Vieslak and Vedolin (2010) set up a reduced form term structure model to model time-varying nominal-real correlations. Hasseltoft (2009) is closely related to my paper, which models time-varying covariance between expected consumption and expected inflation factor as an $AR(1)$ process in a long-run risks model. It can generate time-varying and switching sign of stock-bond correlation, but quantitatively not enough. Another drawback of Hasseltoft $(2009)'s$ setup is that it's hard to maintain the multivariate variance-covariance matrix to be positive-definite under such a specification. The $RSDC$ setup is a good direction to achieve a valid time-varying variance-covariance structure and a closed-form solution simultaneously. And the model is also able to explain various other salient empirical features, which are not pursued in Hasseltoft (2009) . As in the $RSDC$ model, I model the macroeconomic volatility as an autoregressive Gamma process, with regime specific means and volatility levels. Following the similar argument as in Bansal and Shaliastovich (2010) , this channel is very important to generate significant time-varying bond risk premium, and can quantitatively reproduce the violations of expectations hypothesis and Cochrane and Piazzesi (2005) single factor regressions.

Beyond the regime-specific correlation structure, I also allow for the mean levels of consumption growth and inflation to be different across regimes. In the equilibrium, the mean level acts as a "level" factor, driving the regime-shifts in levels of both real and nominal yield curves, consistent with the findings of Bansal and Zhou (2002) , which features a reduced form statistical model with regime switching.

The rest of the paper is organized as follows. In the next section I document

the empirical evidence on changing inflation risks and nominal-real correlation. In Section 3.3 I set up a long-run risks model with regime shifts. I present the solution to the model and discuss its theoretical implications in the Section 3.4. In Section 3.5, I describe the data and calibration of the model parameters. Model's implications for bond and stock markets are presented in the same section. In the last section, I conclude and discuss some further research agenda in this paper.

3.2 Empirical Evidence

In this section, I provide more empirical evidence which motivates the "nominal-real correlation" channel. Fig. 3.2 summarizes the CAPM beta of inflation, which captures the comovement of inflation shocks with stock returns. I estimate a VAR(1) model for inflation, stock returns (real), and the three-month treasury bill returns over a rolling window of 5-years' quarterly data, and then compute the CAPM beta of inflation. The figure shows that the beta of realized inflation moves tremendously and occasionally switches sign. By comparing with Fig. 3.1 one can find that the periods of positive CAPM beta of inflation lie up quite well with negative stock-bond relationship. This is intuitive since inflation is associated with high bond yields and low bond returns. This figure clearly implies that the time-varying and switching signs of stock-bond correlation is closely related to the changing inflation risks.

Fig. 3.3 provides direct evidence of the stochastic nature of nominal-real correlations. I follow a similar moving window quarterly VAR approach to compute the industrial production growth beta of inflation over the long sample (first panel), and consumption beta of inflation over post-war sample (second panel). And the last panel uses GDP and inflation expectations from the Survey of Professional Forecast-

141

ers for the period $1968Q3 - 2009Q4$ to proxy for the correlation of expected growth and expected inflation factors. All three panels display quite similar patterns, and show that the nominal-real correlation does move significantly, and is an important channel to pursue.

3.3 A Long-run Risks Model with Regime-Shifts

3.3.1 Preferences

I consider a discrete-time endowment economy. The investor's preferences over the uncertain consumption stream C_t can be described by the Kreps-Porteus, Epstein-Zin recursive utility function, (see Epstein and Zin (1989) ; Kreps and Porteus (1978)):

$$U_t = \left[(1-\delta)C_t^{\frac{1-\gamma}{\theta}} + \delta \left(E_t U_{t+1}^{1-\gamma} \right)^{\frac{1}{\theta}} \right]^{\frac{\theta}{1-\gamma}}, \tag{3.1}$$

where $\delta \in (0,1)$ is the time discount factor, γ is the risk aversion parameter, and ψ is the intertemporal elasticity of substitution (IES). Parameter θ is defined by $\theta = \frac{1-\gamma}{1-\frac{1}{\psi}}$. Its sign is determined by the magnitudes of the risk aversion and the elasticity of substitution, so that if $\psi > 1$ and $\gamma > 1$, then θ will be negative. Note that when $\theta = 1$, that is, $\gamma = \frac{1}{\psi}$, the above recursive preference collapses to the standard expected utility. As is pointed out by Epstein and Zin (1989) , in this case the agent is indifferent to the timing of the resolution of uncertainty of the consumption path. When risk aversion exceeds (is less than) the reciprocal of IES the agent prefers early (late) resolution of uncertainty of consumption path. In the long-run risks model, agents prefer early resolution of uncertainty of the consumption path.

As shown in Epstein and Zin (1989) , the logarithm of the intertemporal marginal

rate of substitution (IMRS) is given by

$$m_{t+1} = \theta \log \delta - \frac{\theta}{\psi} \Delta c_{t+1} + (\theta - 1) r_{c,t+1}, \quad (3.2)$$

where $\Delta c_{t+1} = log(C_{t+1}/C_t)$ is the log growth rate of aggregate consumption and $r_{c,t+1}$ is the log of the return (i.e. continuous return) on an asset which delivers aggregate consumption as its dividends each time period. This return is not observed in the data. It is different from the observed return on the market portfolio as the levels of market dividends and consumption are not equal: aggregate consumption is much larger than aggregate dividends. Therefore, I assume an exogenous process for consumption growth and use a standard asset-pricing restriction

$$E_t \left[\exp \left(m_{t+1} + r_{t+1} \right) \right] = 1, \quad (3.3)$$

which holds for any continuous return $r_{t+1} = \log(R_{t+1})$, including the one on the wealth portfolio, to solve for the unobserved wealth-to-consumption ratio in the model.

3.3.2 Consumption and Inflation Dynamics

The key model novelty of this paper is that the consumption and inflation dynamics are subject to regime shifts. For notational brevity and expositional ease, I specify the dynamics of consumption and inflation dynamics in a rather general VAR structure with regime shifts. However, I then immediately provide a specific version of the dynamics that is my focus.

I assume there are S regimes that govern the dynamic properties of the n-dimensional state vector $Y_t \in R^n$. The regime variable s_t is a S-state Markov process,

with the probability of switching from regime $s_t = j$ to $s_t = k$ given by a constant $\pi^{jk}, 0 \leq j, k \leq (S-1)$, with $\sum_{k=0}^{S-1} \pi^{jk} = 1$, for all j. Agents are presumed to know the current and past histories of both the state vector and the regime which the economy is in. Thus, the expectation $E_t[\cdot]$ is conditioned on the information set I_t, generated by $\{Y_{t-l}, s_{t-l} : l \geq 0\}$. I use the notation $E^{(j)}[\cdot]$ to denote the unconditional mean of a random variable under the assumption of a single-regime economy governed by the parameters of regime j.

The Markov process governing regime changes is assumed to be conditionally independent of the Y process. In addition, $f(Y_{t+1}|Y_{t-l} : l \geq 0, s_t = j, s_{t+1} = k) = f(Y_{t+1}|Y_{t-l} : l \geq 0, s_t = j)$.

Given $s_t = j$, The state vector of the economy Y_{t+1} follows a VAR that is driven by both Gaussian and demeaned Gamma-distributed shocks:

$$Y_{t+1} = \mu(j) + F(j)Y_t + G_t(j)\varepsilon_{t+1} + \omega_{t+1}^j. \tag{3.4}$$

Here $\varepsilon_{t+1} \sim N(0, I)$ is the vector of Gaussian shocks, and ω_{t+1}^j is the vector of demeaned Gamma-distributed shocks. The detailed parameterization of Gamma distribution will be provided later. To put the dynamics into an affine class, I impose an affine structure on G_t^j:

$$G_t(j)G_t'(j) = h(j) + \sum_i H_i(j)Y_{t,i}, \tag{3.5}$$

where $h(j) \in R^{n \times n}$, $H_i(j) \in R^{n \times n}$, the i denotes the ith component of state vector Y_t.

In the calibration section of the paper and some of the discussions that follow, I focus on a particular specification of (3.4). This specification is a generalized LRR

144

model that incorporates regime shifts and non-neutrality of expected inflation factor. Here I give an overview of this generalized LRR model and map it into the general framework in (3.4). Further details are also provided in the calibration section.

I specify:

$$
Y_{t+1} = \begin{pmatrix} \Delta c_{t+1} \\ \pi_{t+1} \\ x_{t+1} \\ z_{t+1} \\ \sigma^2_{t+1} \\ \Delta d_{t+1} \end{pmatrix}, \mu(j) = \begin{pmatrix} \mu^j_c \\ \mu^j_\pi \\ 0 \\ 0 \\ \left(\sigma^j_c\right)^2 (1-\nu_\sigma) \\ \mu^j_d \end{pmatrix}, F(j) = \begin{pmatrix} 0 & 0 & 1 & \tau_z & 0 & 0 \\ 0 & 0 & \tau_x & 1 & 0 & 0 \\ 0 & 0 & \nu_x & 0 & 0 & 0 \\ 0 & 0 & 0 & \nu_z & 0 & 0 \\ 0 & 0 & 0 & 0 & \nu_\sigma & 0 \\ 0 & 0 & \phi & \phi\tau_z & 0 & 0 \end{pmatrix}.
$$

The vector of Gaussian shocks $\varepsilon_{t+1} = (\varepsilon_{c,t+1}, \varepsilon_{\pi t+1}, \varepsilon_{x,t+1}, \varepsilon_{z,t+1}, 0, \varepsilon_{d,t+1})' \sim N(0, I)$ and $\omega^j_{t+1} = \left(0, 0, 0, 0, \omega^j_{\sigma,t+1}, 0\right)'$, in which $\omega^j_{\sigma,t+1}$ follows demeaned Gamma distribution, i.e. $\omega^j_{\sigma,t+1} = \widetilde{\omega}^j_{\sigma,t+1} - E\left(\widetilde{\omega}^j_{\sigma,t+1}\right)$. The Gamma distribution of $\widetilde{\omega}^j_{\sigma,t+1}$ is characterized by two parameters, so we specify the mean and volatility of the volatility shocks as

$$
\begin{aligned}
E\left(\widetilde{\omega}^j_{\sigma,t+1}\right) &= \left(\sigma^j_c\right)^2 (1-\nu_\sigma), \\
Var\left(\widetilde{\omega}^j_{\sigma,t+1}\right) &= \left(\sigma^j_{c\omega}\right)^2.
\end{aligned}
$$

Therefore, we have $E\left(\omega^j_{\sigma,t+1}\right) = 0$, and $Var\left(\omega^j_{\sigma,t+1}\right) = (\sigma^j_{c\omega})^2$.

The first two components Δc_{t+1} and π_{t+1} denote the consumption growth and inflation. x_{t+1} and z_{t+1} are long-run expected growth and expected inflation factors. The term $(\mu_c + x_t + \tau_z z_t)$ is the conditional expectation of consumption growth where x_t is a small but persistent expected consumption factor that captures long run risks in consumption and dividend growth, as in standard LRR model. Similarly, the term

z_t is a small but persistent expected inflation factor that captures long-run risks in inflation. [1]

The parameter $\tau_z \neq 0$ leads to a non-neutral LRR model, that is, the expected inflation factor feeds back to the real economy, i.e. the consumption process. For a typical parameter of $\tau_z < 0$, it means that a positive expected inflation factor will lower the future expected consumption growth. In this case, the long-run inflation risk is priced and risk compensation for this risk factor is embodied in risk premium even for real assets. In Bansal and Shaliastovich (2010) , $\tau_z = 0$, that is, the inflation process does not feed back to the real economy.

The term Δd_{t+1} is logarithm dividend growth, which is defined as a leveraged process of Δc_{t+1}, with a leverage parameter $\phi > 1$. Thus, the dividend growth is more sensitive to x_t and z_t than is consumption growth.

The volatility process σ_{t+1}^2 follows an autoregressive Gamma process. I assume that the innovations in volatility process $\omega_{\sigma,t+1}^j$ follows demeaned Gamma distribution, following Barndorff-Nielson and Shephard (2001) and Bansal and Shaliastovich (2010) , in order to guarantee that the variances always stay positive. As noted in Bansal and Shaliastovich (2010) , this specification will generate very similar asset pricing results as a Gaussian volatility shock.

I set the conditional variance-covariance matrix of the Gaussian shocks to be

[1] Note that x_t and z_t are only part of the stochastic expected consumption and inflation, respectively, which is different with the standard LRR model. We call x_t and z_t expected growth and inflation factors, respectively, throughout this paper.

$G_t(j) G_t(j)' = H_\sigma(j) \sigma_t^2$, in which

$$H_\sigma(j) = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & \varphi_\pi & 0 & 0 & 0 & 0 \\ 0 & 0 & \varphi_x & \rho_{xz}^j \varphi_z & 0 & 0 \\ 0 & 0 & 0 & \varphi_z & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ \rho_{dc} & 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

The terms φ_π, φ_x, φ_z are regime-independent constants and denote the relative magnitudes of inflation volatility shock, long-run consumption shock and long-run inflation shock with respect to short-run consumption shock. This is a simplification assumption tying the dynamics of four Gaussian shocks to the same factor. It is straightforward to extend the model by allowing for multiple factors to derive different volatility shocks, though, I do not pursue this additional complication in this paper.

The parameters ρ_{xz}^j is regime specific and captures the different correlations between long-run consumption and inflation shocks in different economic regimes. This specification closely follows the regime switching dynamic correlation ($RSDC$) model in Pelletier (2006) . As a summary, I decompose the covariance into standard deviation and correlation. The standard deviation follows a continuous stochastic volatility process. And the correlation is dynamic, in particular, it follows a regime switching model; it is constant within a regime, but different across regimes. This setup can be seen as a midpoint of the constant conditional correlation (CCC) model of Bollerslev (1990) and the dynamic conditional correlation (DCC) model of Engle (2002) . An important advantage of $RSDC$ specification is that it allows for tractability of the general equilibrium model within the affine framework (after some log-linearization

147

approximation), while being able to generate time-varying correlation of risk factors.

3.4 Model Solutions and Intuitions

3.4.1 Within Regime Intuitions

Before going through the solution to the above full blown LRR model with regime switching, I lay down some within regime intuitions, i.e. I solve the model by assuming a single-regime economy governed by the parameters of regime j, for $j = 0, ...S - 1$.

To get an analytical solution of the model, I log-linearize the return on consumption claim to solve for the equilibrium discount factor and asset prices. In equilibrium, the wealth-to-consumption ratio, v_t, is linear in states,

$$v_t = A_0 + A_x x_t + A_z z_t + A_\sigma \sigma_t^2$$

Using the Euler equation (3.3) and the assumed dynamics of consumption growth and inflation, I derive the solutions coefficients A_x, A_z and A_σ :

$$A_x = \frac{1 - \frac{1}{\psi}}{1 - \kappa_1 \rho_x},$$

$$A_z = \frac{\left(1 - \frac{1}{\psi}\right) \tau_z}{1 - \kappa_1 \rho_x},$$

$$A_\sigma = \frac{(1 - \gamma)\left(1 - \frac{1}{\psi}\right)\left[1 + \kappa_1^2 \left\{ \left(\frac{1}{1 - \kappa_1 \nu_x} + \frac{\tau_z \rho_{xz}}{1 - \kappa_1 \nu_z}\right)^2 \varphi_x^2 + \left(\frac{\rho_{xz}}{1 - \kappa_1 \nu_x} + \frac{1}{1 - \kappa_1 \nu_z}\right)^2 \varphi_z^2 \right\}\right]}{2(1 - \kappa_1 \nu_\sigma)}.$$

The details for the model solution and the expression for the endogenous log-linearization coefficients are provided in Appendix C.2.

148

It follows that A_x is positive if the IES, ψ, is greater than one. In this case the intertemporal substitution effect dominates the wealth effect. In response to higher expected growth, agents buy more assets, and consequently the wealth-to-consumption ratio rises. In the standard power utility model with risk aversion larger than one, the IES is less than one, and hence A_x is negative – a rise in expected growth potentially lowers asset valuations. That is, the wealth effect dominates the substitution effect.

The coefficient A_z measures the sensitivity of wealth-to-consumption ratio to fluctuations of expected inflation factor. Consider the case that IES, ψ, is greater than one, the sign of A_z is determined by τ_z, which captures the response of expected growth factor on expected inflation. When typically $\tau_z < 0$, that is, a high expected inflation factor leads to a low expected consumption growth, the wealth-to-coumption ratio responds negatively ($A_z < 0$). In previous LRR literature, for instance, Bansal and Yaron (2004) and Bansal and Shaliastovich (2010) , τ_z is set to zero. In their setup, the fluctuations of expected inflation does not feed back to real economy (i.e. consumption growth), and thus does not affect real asset allocations and prices. In my setup ($\tau_z \neq 0$), the long-run expected inflation factor does affect the real economy, and I call it a **non-neutral model**.

The coefficient A_σ measures the sensitivity of wealth-to-consumption ratio to volatility fluctuations. If the IES and risk aversion are larger than one, the loading is negative. In this case a rise in consumption or expected growth volatility lowers asset valuations and increases the risk premium on all assets.

Using the equilibrium condition for the wealth-to-consumption ratio, I can pro-

149

vide an analytical expression for the pricing kernel:

$$m_{t+1} = m_0 + m_x x_t + m_z z_t + m_\sigma \sigma_t^2$$

$$-\lambda_c \sigma_t \varepsilon_{c,t+1} - \lambda_x \varphi_x \sigma_t \varepsilon_{x,t+1} - \lambda_z \varphi_z \sigma_t \varepsilon_{z,t+1} - \lambda_\omega \omega_{\sigma,t+1}.$$

In particular, the conditional mean of the pricing kernel is affine in state variables x_t, z_t and σ_t^2, where the loadings m_0, m_x, m_z and m_σ depend on model and preference parameters, as provided in the Appendix C.2.

The innovations in the pricing kernel are very important for thinking about risk compensation (risk premia). The magnitudes of the risk compensation depend on the market prices of short-run, long-run consumption and inflation risks, as well as the volatility risks $\lambda_c, \lambda_x, \lambda_c$ and λ_ω. The market prices of systematic risks can be expressed in terms of underlying preferences and parameters that govern the evolution of consumption growth and inflation:

$$
\begin{aligned}
\lambda_c &= \gamma, & (3.6) \\
\lambda_x &= (1-\theta)\kappa_1 A_x, \\
\lambda_z &= (1-\theta)\kappa_1 \left(A_x \rho_{xz} + A_z\right), \\
\lambda_\omega &= (1-\theta)\kappa_1 A_\sigma.
\end{aligned}
$$

The compensation for the short-run consumption risks is standard and given by the risk-aversion coefficient γ. In the special case of power utility, $\gamma = \frac{1}{\psi}$, the risk compensation parameters $\lambda_x, \lambda_z,$ and λ_ω are zero, and the intertemporal marginal rate of substitution collapses to standard power utility specification,

$$m_{t+1}^{CRRA} = \log\delta - \gamma\Delta c_{t+1}.$$

With power utility there is no separate risk compensation for long-run growth, inflation risks and volatility risks, while, with generalized preferences, these risks are priced. The pricing of long-run and volatility risks is an important feature of LRR model.

When agents have a preference for early resolution of uncertainty, $\theta < 1$, (i.e. $\gamma > \frac{1}{\psi}$), the price of long-run consumption risks λ_x is positive, and the price of volatility risks λ_ω is negative. That is, the states with low expected growth or high volatility are bad states and discounted more heavily. It is important to note that the price of long-run inflation risks λ_z is intimately related to ρ_{xz}, which captures the sensitivity of expected consumption factor to the innovations of expected inflation. In particular, when $\rho_{xz} < 0$, that is, a positive news in expected consumption factor predicts a decrease in long-run expected inflation, λ_z is negative; in contrast, when $\rho_{xz} > 0$, λ_z can switch sign and become positive. In the model, I allow for this correlation parameter ρ_{xz} to switch sign, which is the **key channel** to get switching sign of market price of long-run inflation risk, and thus switching sign of nominal bond risk premium.

The discount factor used to price nominal payoff is given by

$$m_{t+1}^{\$} = m_{t+1} - \pi_{t+1}.$$

The solution to the nominal discount factor is affine in the state variables, and nominal market prices of risks depend on the real prices of risks and inflation sensitivity to short and long-run consumption and inflation news.

Given the solutions to the real and nominal discount factor, I obtain that the

yields on real and nominal bonds satisfy

$$y_{t,n} = \frac{1}{n}\left(B_{0,n} + B_{x,n}x_t + B_{z,n}z_t + B_{\sigma,n}\sigma_t^2\right),$$

$$y_{t,n}^{\$} = \frac{1}{n}\left(B_{0,n}^{\$} + B_{x,n}^{\$}x_t + B_{z,n}^{\$}z_t + B_{\sigma,n}^{\$}\sigma_t^2\right).$$

The bond coefficients, which measure the sensitivity of bond prices to the aggregate risks in the economy, are pinned down by the preference and model parameters. As shown in Appendix C.2, the real yields respond positively to the expected growth factor, $B_{x,n} > 0$, and negatively to volatility state, $B_{\sigma,n} < 0$. When high expected inflation factor indicates lower future consumption $(\tau_z < 0)$, the real yields loads negatively on expected inflation factor, $B_{z,n} < 0$.

One-period expected excess return on a real bond with 2 periods to maturity can be written in the following form:

$$rp_{t+1}^{(2)} = E_t rx_{t+1}^{(2)} + \frac{1}{2}Var_t rx_{t+1}^{(2)} \tag{3.7}$$

$$= -\frac{1}{\psi}\left(\gamma - \frac{1}{\psi}\right)\kappa_1 \left(\begin{array}{c} \frac{\varphi_x^2 + \rho_{xz}^2\varphi_z^2}{1-\kappa_1\nu_x} + \frac{\tau_z^2\varphi_z^2}{1-\kappa_1\nu_z} \\ +\tau_z\rho_{xz}\left(\frac{1}{1-\kappa_1\nu_x} + \frac{1}{1-\kappa_1\nu_z}\right)\varphi_z^2 \end{array} \right)\sigma_t^2 - B_{3,1}\lambda_\sigma\left(\sigma_{c\omega}\right)^2.$$

The last term, $-B_{3,1}\lambda_\sigma\left(\sigma_{c\omega}\right)^2$, is the component of real bond premium attributable to volatility risks. As shown in Appendix C.2, $-B_{3,1}$ is the beta of real bond return with respect to volatility innovations, which is positive. Since the market price of volatility risks λ_σ is negative, therefore, the volatility risks always contribute negatively to real bond premium. If the correlation of expected consumption and expected inflation shocks is zero, $\rho_{xz} = 0$, the real bond risk premium from long-run consumption and inflation risks are always negative. When the expected consumption factor decreases

152

at the time of high expected inflation factor, i.e. $\tau_z < 0$, negative correlation of expected consumption and inflation shocks, i.e. $\rho_{xz} < 0$, will decrease the real bond premium even further. On the other hand, positive correlation of these shocks can increase bond risk premium. I get similar implications for real bond premium at the long end.

For nominal bonds, I obtain that nominal bond risk premium is equal to real bond risk premium plus an additional component capturing inflation shocks:

$$
rp_{t+1}^{\$(2)} - rp_{t+1}^{(2)} = -\left(\gamma - \frac{1}{\psi}\right)\kappa_1 \left(\begin{array}{c} \frac{\rho_{xz}\varphi_z^2}{1-\kappa_1\nu_x} + \frac{\tau_z\varphi_z^2}{1-\kappa_1\nu_z} \\ +\tau_x\left(\frac{\varphi_x^2+\rho_{xz}^2\varphi_z^2}{1-\kappa_1\nu_x} + \frac{\tau_x\rho_{xz}}{1-\kappa_1\nu_z}\right)\varphi_z^2 \end{array} \right) \sigma_t^2. \tag{3.8}
$$

Again, as shown in Appendix C.2, the parameters τ_z and ρ_{xz} are very important the determine the risk premium for nominal bond. And its regime-dependent feature is the driving force to lead to the switching sign of bond risk premium and stock-bond correlation.

3.4.2 Characterizations of Different Regimes

A salient feature of the model is that I allow for regime shifts in exogenous consumption and inflation dynamics. In particular, in the calibration part, I allow for 3 regimes – expansion, contraction and deep recession regimes and the following elements as shown in Table 3.1 to be regime specific. Ideally I should use a more general model as laboratory, carry out a structural estimation by both consumption, inflation and bond/equity market prices/returns, and let the data to speak for themselves about the classifications of different regimes. I leave this for further research.

I first discuss the implications of different regime-specific elements, and then give economic interpretations of three regimes.

First, I set the nominal-real correlation, ρ_{xz}, to switch signs, in particular, $\rho_{xz} < 0$ in expansion and contraction regimes, while $\rho_{xz} > 0$ in deep recession regime. According to (3.6), the market price of long-run inflation risks can switch sign, in particular, the price is negative with negative correlation, and vice versa. On the other hand, the beta of nominal bond holding period return with respect to long-run inflation risk innovation stays in the same direction. As a whole, long-run inflation risk contributes positively to the nominal bond risk premium in expansion and contraction regimes, while negatively in the deep-recession regime. The time-varying correlation also causes the equity beta to long-run inflation risks to switch sign, and therefore, generates switching sign of correlation between stock and nominal bond returns, i.e. in deep recession, the stock and bond correlation is negative, while in the other two regimes, the relationship is positive. This is consistent with the empirical evidence I highlighted in Section 3.1. These intuitions are summarized in the following Table 3.1. In sum, this time-varying nominal-real correlation structure is the key channel to generate switching signs of nominal bond risk premium as well as stock-bond correlations.

Second, I allow for different levels of consumption growth and inflation across different regimes. In equilibrium, higher μ_c, the unconditional mean of consumption growth, implies higher real and nominal bond yields at all maturities simultaneously, while μ_π, the unconditional mean of inflation, only affects nominal yield, but not real. Higher μ_π indicates higher nominal yield at all maturities. To sum up, the consumption growth and inflation levels, μ_c and μ_π, are "level" factors to term structure of interest rate. The former shifts both real and nominal yield curve, while the latter only affects the nominal side. This channel is very important to generate the shifts

154

of yield curve in different economic regimes.

Third, the macroeconomic uncertainty, captured by consumption and inflation volatility (tied to each other in the parsimonious specification of this paper) are allowed to be time-varying. This time-varying feature of uncertainty lies in twofolds. On one hand, within each regime, the volatility follows an autoregressive Gamma process. This is important to get enough variations in equity and bond risk premium, and helps to replicate the violations of expected hypothesis of bond returns. The intuition is similar to Bansal and Shaliastovich (2010). On the other hand, I allow for regime specific uncertainty levels. I specify a low volatility mean level in expansion, but higher mean level in bad states. This channel corresponds to a counter-cyclical property of stock volatility, and is very important to generate higher equity premium in contraction/deep-recession regimes, which is consistent with the empirical findings of Lustig and Verdelhan (2010) **.** Furthermore, the mean level of volatility also constitutes a "slope" factor of yield curve, as the real (nominal) yield curve slope is determined by the risk premium of long-term real (nominal) bond, which is proportional to volatility level, as shown in the Appendix B. Different levels of macroeconomic volatility will also alter the levels of real and nominal bond yields, due to a precautionary saving argument.

Table 3.1: Characterizations of Different Regimes

	Expansion	Contraction	Deep-Recession
Nominal-real correlation, ρ_{xz}	Negative	Negative	Positive
Consumpiton level, μ_c	High	Medium	Low
Inflation level, μ_π	Medium	High	Low
Uncertainty, σ_c^2	Low	Medium	High

I broadly classify the economy into three regimes: expansion, contraction and deep-recession. In the expansion regime, we have high consumption growth, a medium level of inflation, and low uncertainty (which is measured by consumption and inflation volatility). In the contraction regime, we have lower growth rate and higher uncertainty. The inflation level is also higher. One can think this regime as a stagflation regime, in which low growth and high inflation coexist. A typical sample episode is the period late 1970's and early 1980's. In the deep-recession regime, we have the lowest growth and the highest level of uncertainty. As opposed to regular contraction, this regime has very low inflation, and thus deflation rather than inflation is more of a concern at this time. Another key ingredient that is different across three regimes is the **nominal-real correlations**, in particular, in the model it refers to the correlation between shocks to expected growth and inflation factors. In the first two regimes, positive news to expected growth factor indicates a lower future expected inflation; however, in the deep recession regime, the relationship is just the opposite. I provide the empirical evidence to support this channel in the next section. This ingredient is very important to generate large movements (and potentially switch signs) of nominal bond risk premium as well as stock-bond correlation.

3.4.3 Solutions to Long-run Risks Model with Regime Shifts

I now solve for the equilibrium price process of the model economy. The solution proceeds via the representative agent's Euler equation (3.3). To price assets we must first solve for the return on consumption claim, $r_{c,t+1}$, as it appears in the pricing kernel itself. Denote the logarithm of the wealth-to-consumption ratio at given time t and state $s_t = j$ by $v_{c,t}(j)$. Since the consumption claim pays the consumption

Table 3.2: Market Price of Risks and Return Betas to Risk Innovations

Shocks	SR cons.	LR cons.	LR Infl.	Vol.	Overall Premium
Expansion Regime					
Risk Price	+	+	-	-	
Equity Return Beta	+	+	-	-	+
Nom. Bond Return Beta	**0**	+	-	-	+
Deep-Recession Regime					
Risk Price	+	+	+	-	
Equity Return Beta	+	+	+	+	+
Nom. Bond Return Beta	**0**	+/-	-	-	-

stream as its dividend, this is simply the price-dividend ratio of such a claim. Next, I use Campbell and Shiller (1988) log-linearization to linearize $r_{c,t+1}(j,k)$, which depends on two consecutive states $s_t = j$ and $s_{t+1} = k$, around the unconditional means of $v_t(j)$ and $v_{t+1}(k)$, respectively:

$$r_{c,t+1}(j,k) = \kappa_{c,0}(k) + \kappa_{c,1}(k) v_{c,t+1}(k) - v_{c,t}(j) + \Delta c_{t+1}(j). \qquad (3.9)$$

A similar approach is taken by Bansal and Yaron (2004) , and Bansal, Kiku and Yaron (2007) for a standard LRR model (without regime switching). I then conjecture that given current state $s_t = j$, the log wealth-to-consumption ratio is affine in the state vector:

$$v_t(j) = A_0(j) + A'Y_t(j), \qquad (3.10)$$

where $A(j) = (A_c(j), A_\pi(j), A_x(j), A_z(j), A_\sigma(j), A_d(j))'$ is a vector of pricing coefficients, which are regime-specific. Substituting (3.10) into (3.9) and then substituting (3.9) into the Euler equation gives the equation in terms of A, A_0 and the state variables. The expectation on the left hand side of this Euler equation can be evaluated analytically, as shown in Appendix C.3. Since any predictive information in Δc_t,

π_t and Δd_t is contained in x_t and z_t, they have no effects on $v_t(j)$ and therefore $A_c(j) = A_\pi(j) = A_d(j) = 0$.

Pricing Kernel

Having solved for $r_{c,t+1}(j,k)$, I substitute it into m_{t+1}, which also depends on two consecutive states $s_t = j$ and $s_{t+1} = k$, to obtain an expression for logarithm pricing kernel at time $t + 1$:

$$m_{t+1}(j,k) = m_0(j,k) + m_1(j,k)' Y_t - \Lambda(k)' \left(G_t(j) \varepsilon_{t+1} + \omega_{t+1}^j \right). \tag{3.11}$$

The loadings $m_0(j,k)$, $m_1(j,k)$ and $\Lambda(k)$ are provided in Appendix C.3.

The Market Returns

I solve for the market return. A share in the market is modeled as a claim to a dividend with growth process given by Δd_{t+1}. To solve for the price of a market I proceed along the same lines as for the consumption claim and solve for $v_{m,t}(j)$ – the price-to-dividend ratio of the market at time t, given $s_t = j$, by using Euler equation (3.3). To do this, log-linearize the return on the market, $r_{m,t+1}$, which depends on on two consecutive states $s_t = j$ and $s_{t+1} = k$, around the unconditional means of $v_{m,t}(j)$ and $v_{m,t+1}(k)$, respectively:

$$r_{m,t+1}(j,k) = \kappa_{0,m}(k) + \kappa_{1,m}(k) v_{m,t+1}(k) - v_{m,t}(j) + \Delta d_{t+1}(j). \tag{3.12}$$

Then conjecture that $v_{m,t}(j)$ is affine in the state variables:

$$v_{m,t}(j) = A_{0,m}(j) + A_{1,m}(j)'Y_t, \tag{3.13}$$

158

where $A_{1,m}(j) = (A_{c,m}(j), A_{\pi,m}(j), A_{x,m}(j), A_{z,m}(j), A_{\sigma,m}(j), A_{d,m}(j))'$ is the vector of pricing coefficients which are regime-independent. For similar reasons as in wealth-to-consumption ratio, we have $A_{c,m}(j) = A_{\pi,m}(j) = A_{d,m}(j) = 0$.

By substituting the expression for $v_{m,t}(j)$ into the linearized return, I obtain an expression for $r_{m,t+1}(j,k)$ in terms of Y_t and its innovations:

$$r_{m,t+1}(j,k) = J_{0,m}(j,k) + J_{1,m}(j,k)'Y_t + \beta_d(k)'\left(G_t(j)\varepsilon_{t+1} + \omega_{t+1}^j\right), \qquad (3.14)$$

in which the loadings of $J_0(j,k)$, $J_1(j,k)$ and $\beta_d(k)$ are provided in Appendix C.3.

Bond Prices

The equilibrium real and nominal yield are affine in the state variables. Indeed, in Appendix C.3, I show that real and nominal yields at time t given the state $s_t = j$ satisfy

$$y_{t,n}(j) = \frac{1}{n}\left[B_{0,n}(j) + B_n(j)'Y_t\right], \qquad (3.15)$$

$$y_{t,n}^{\$}(j) = \frac{1}{n}\left[B_{0,n}^{\$}(j) + B_n^{\$}(j)'Y_t\right]. \qquad (3.16)$$

The bond coefficients, which measure the sensitivity of bond prices to the aggregate risks in the economy, are pinned down by the preference and model parameters – the expressions for the loadings are presented in Appendix C.3.

Define the holding period return of real bond as $rb_{t+1}^{(n)}(j,k) = ny_t^{(n)}(j) - (n-1)y_{t+1}^{(n-1)}(k)$, thus I get

$$rb_{t+1}^{(n)}(j,k) = G_{0,n}(j,k) + G_{1,n}(j,k)'Y_t + \beta_{b,n}(k)'\left(G_t(j)\varepsilon_{t+1} + \omega_{t+1}^j\right).$$

And similarly the holding period returns for nominal bond is

$$rb_{t+1}^{\$(n)}(j,k) = G_{0,n}^{\$}(j,k) + G_{1,n}^{\$}(j,k)'Y_t + \beta_{b,n}^{\$}(k)' \left(G_t(j)\,\varepsilon_{t+1} + \omega_{t+1}^{j} \right). \qquad (3.17)$$

3.5 Calibration and Empirical Results

3.5.1 Preference Parameters

I calibrate the subjective discount factor $\delta = 0.998$. The risk-aversion coefficient is set at $\gamma = 10$.

There is a debate in the literature about the magnitude of the IES. As in Bansal and Yaron (2004), I focus on an IES of 1.5 – an IES value larger than one is important for the quantitative results. Bansal, Kiku and Yaron (2007) document that the asset valuations fall when consumption volatility is high, which is consistent only with $\psi > 1$.

3.5.2 Calibration of Consumption and Inflation

I follow the standard LRR literature to calibrate the parameters for consumption and inflation outlined in (3.4) at a monthly frequency and time-aggregate the output from monthly simulations to match the key aspects of annual consumption growth and inflation rate is US from 1930 to 2009. I report the calibration output of the model, which is based on a very long simulation of monthly data aggregated to annual horizon, in Table 3.8. As shown in the table, the model can match very well the salient features of the consumption and inflation data.

Table 3.4 shows the detailed parameter calibrations of this model.

160

3.5.3 Calibration of Regime Switches

I use $j = 0, 1, 2$ to denote deep recession, contraction and expansion regimes, respectively. I assume $\pi^{01} = 0$ and $\pi^{10} = 0$, therefore, two-types of recessions cannot switch from each other. Using the length of NBER dated business cycles to calibrate transition probability matrix, I base the calibration on 17 recessions from $1919 - 2009$, in which I consider two events i.e. (1) 1929 August - 1933 March and (2) 2007 December to 2009 June as deep recessions, while the other 15 events as regular recessions. Table 3.6 and 3.7 shows the calibration of transition probability matrix and its implications for unconditional probability and average duration of each regime. They match the data counterparts quite well.

3.5.4 Empirical Results

Table 3.9 reports the model performance in terms of unconditional moments on the bond and stock markets. As we can see, the model, on average, generates a downward sloping real yield curve, while a positive nominal term structure. The average yields over 1 to 5 years match the data reasonably well, and the magnitude is also similar to Bansal and Shaliastovich (2010) . On the equity market, the model can match the equity premium reasonably well.

Table 3.10 reports the model implied yields conditional on different regimes. Consistent with Bansal and Zhou (2002) , we see significant level shifts across different regimes. In the deep recession regime, the nominal yield is the lowest, mainly because both consumption growth and inflation levels are low, while the macro uncertainty is high. The nominal yield curve in the contraction regime is the highest, mainly because the inflation level is the highest at this regime. The table also shows that

161

the slope of nominal term structure in contraction is flat, which is consistent with data.

In Table 3.11, I report the model implied risk premia for nominal (real) bond holding period returns as well as the equity returns, conditional on different regimes. As the intuition discussed in Section 3.4, in the deep recession regime, the nominal bond premium is negative, while it is positive in other regimes, due to the time-varying and switching signs of long-run inflation risk premium. And we also get conditional stock-bond correlation consistent with Fig. 3.1. Another feature worthy of mentioning is that conditional on the contraction regime, we have higher premia in both bond and equity than those in expansion regime. This is consistent with the findings of Lustig and Verdelhan (2010) .

In Table 3.12, I report the model implications for single factor projections of Cochrane and Piazzesi (2005). Following their approach, I regress the average of 1-year nominal excess returns of 2 to 5 years to maturity on the forward rates of 1, 3 and 5 years to maturity:

$$\frac{1}{4}\sum_{n=2}^{5} rx_{t+12,12n} = \gamma_0 + \gamma_1 f_{t,12} + \gamma_2 f_{t,36} + \gamma_3 f_{t,60} + error.$$

I extract a single bond factor $\widehat{rx}_{t,m} = \widehat{\gamma}_0 + \widehat{\gamma}_1 f_{t,12} + \widehat{\gamma}_2 f_{t,36} + \widehat{\gamma}_3 f_{t,60}$ from this regression, which is subsequently used to forecast excess bond returns at each maturity n from 2 to 5 years:

$$rx_{t+m,n} = const + b_{m,n}\widehat{rx}_{t,m} + error$$

Cochrane and Piazzesi (2005) show that the estimates $b_{m,n}$ are positive and increasing with horizon, and a single factor projection captures $15 - 18\%$ of the variation

162

in bond returns. From the model, the slope coefficients in the second stage regressions increase from 0.45 at 2-year maturity to 1.48 at 5 years, which matches very well with the data estimates. The model implied R^2 in the projections is about 11%, which matches more than two thirds of that in the data. This implies that the model is able to generate considerable time-variations of bond risk premium due to both time-varying volatility and regime switching channels. The single factor projections for the real bonds delivers a very similar pattern for the second-stage coefficients. The R^2's in real regressions are quite substantial, though, they somewhat decrease relative to their nominal counterparts. Hence, the model implies that predictability of bond returns is intrinsically a feature of the real economy.

As a summary, I use the following table to illustrate the link between regime-specific ingredients with their conditional model implied results.

Table 3.3: Model Implications Conditional on Different Regimes

	Expansion	Contraction	Deep-Recession
Nominal-Real Correlation Channel			
Bond-Stock Correlation	Positive	Positive	Negative
Nominal Bond Premium	Positive	Positive	Negative
Consumption and Inflation Level Channel			
Nominal Yield Level	Medium	High	Low
Yield Curve Slope	Upward	Upward	Flat
Volatility Channel			
Equity Premium	Low	Medium	High
Integrated Variance	Low	Medium	High

3.6   Conclusion

This paper studies the joint determinants of stock and bond returns in Bansal and Yaron (2004) type of long-run risks framework. A novel ingredient of the model is to allow for regime shifts in consumption and inflation dynamics – in particular, the means, volatilities, and the correlation structure of consumption and inflation dynamics are regime-dependent. This rational expectations general equilibrium framework can (1) jointly match the dynamics of consumption, inflation and cash flow; (2) generate time-varying and switching sign of stock-fbond correlation, as well as switching sign of bond risk premium; (3) coherently explain another long list of salient empirical features in stock and bond markets, including time-varying equity and bond return premia, regime shifts in real and nominal yield curve, the violation of expectations hypothesis of bond returns. The model also reveals that insight that term structure of interest rates and stock-bond correlations are intimately related to business cycles, while long-run risks and volatility risks play a more important role to account for high equity premium than business cycle risks.

164

Table 3.4: Model Parameters Calibration

Parameters	Notation	Value
Preference Parameters		
Subjective discount factor	δ	0.9982
IES	ψ	1.5
Risk Aversion	γ	10
Consumption and Inflation Dynamics: Regime-independent Parameters		
Persistence of x_t	ν_x	0.989
Persistence of z_t	ν_z	0.989
Persistence of σ_t^2	ν_σ	0.982
$\sigma_{\pi,t}/\sigma_t$	φ_π	1.35
$\sigma_{x,t}/\sigma_t$	φ_x	0.03
$\sigma_{z,t}/\sigma_t$	φ_z	0.05
The loadings of Δc_{t+1} on z_t	τ_z	-0.2
The loadings of π_{t+1} on x_t	τ_x	-1
Volatility of volatility	$\sigma_{c\omega}$	2.60E-06

This table reports calibrated parameter values for the baseline model. The model is calibrated at monthly frequency, hence I report monthly parameter values. $j = 0, 1, 2$ denote deep-recession, contraction and expansion regimes, respectively.

Table 3.5: Model Parameters Calibration (Continued)

Cons. and Infl.: Regime-dependent Parameters				
Regimes	j	0	1	2
Consumption mean	μ_c	0.005	0.012	0.022
Inflation mean	μ_π	0.008	0.041	0.022
Correlation btw exp. growth and inflation shocks	ρ_{xz}	0.5	-0.5	-0.8
Consumption volatility	σ_c	0.0064	0.0052	0.004

This table reports calibrated parameter values for the baseline model. The model is calibrated at monthly frequency, hence I report monthly parameter values. $j = 0, 1, 2$ denote deep-recession, contraction and expansion regimes, respectively.

Table 3.6: Transition Probability Matrix Calibration

	Deep-Recession (j'=0)	Contraction (j'=1)	Expansion (j'=2)
j = 0	0.9672	0	0.0328
j = 1	0	0.9107	0.0893
j = 2	0.0023	0.0175	0.9802

This table reports the calibrated value of transition probability matrix. The notion $j = 0, 1, 2$ denote deep-recession, contraction and expansion regimes, respectively.

Table 3.7: Model Implied Regime Switching Features

	Unconditional Probability			Duration (months)		
	j = 0	j = 1	j = 2	j = 0	j = 1	j = 2
Data	0.0561	0.1544	0.79	30.5	11.2	50.53
Model	0.0552	0.1548	0.79	30.5	11.2	50.53

The data counterparts are based on the length of 17 NBER dated recessions from $1919 - 2009$. Two events (1) $1929/08 - 1933/03$ and (2) $2007/12 - 2009/6$ are considered as deep-recessions. $j = 0, 1, 2$ denote deep-recession, contraction and expansion regimes, respectively.

Table 3.8: Model Implied Unconditional Moments

Moments	Data	Model
$E(\Delta c)$	2.03	2.04
$\sigma(\Delta c)$	2.26	2.16
$AC1(\Delta c)$	0.46	0.64
$AC2(\Delta c)$	0.13	0.43
$E(\pi)$	3.17	3.53
$\sigma(\pi)$	4.05	3.68
$AC1(\pi)$	0.63	0.7
$AC2(\pi)$	0.29	0.38
$corr(\Delta c, \pi)$	-0.24	-0.36

The data moments are based on annualized consumption and inflation data from $1930 - 2009$. The consumption are nondurable expenditure and service from BEA. The inflation are deseasonalized CPI from FRED dataset. The model implied moments are computed from a long simulation of monthly model, time aggregated to annual frequency. All the statistics reported in this table are annualized.

Table 3.9: Bond and Equity Markets

	1y	2y	3y	4y	5y
	Nominal Term Structure				
Mean (Data)	5.56	5.77	5.94	6.07	6.16
Mean (Model)	5.54	5.61	5.81	6.03	6.26
	Equity Returns				
	Data	Model			
Equity Premium	7.6	5.39			
Std. Dev.	20.45	13.6			

The equity returns are CRSP value weighted portfolio comprising the stocks traded in the NYSE, AMEX and NASDAQ, from $1926 - 2009$. The nominal yields data are from Fama-Bliss monthly Dataset from June 1952 till Dec 2009. The model implied moments are computed from a long simulation of monthly model, time aggregated to annual frequency. All the statistics reported in this table are annualized.

Table 3.10: Model Implications – Bond Yield

	1y	2y	3y	4y	5y
Conditional on expansion regime					
Real Yield	2.23	2.06	1.79	1.51	1.23
Nominal Yield	5.44	5.66	5.86	6.08	6.3
Conditional on deep-recession regime					
Real Yield	2	1.65	1.3	0.97	0.64
Nominal Yield	3.17	3.28	3.51	3.81	4.15
Conditional on contraction regime					
Real Yield	1.76	1.47	1.17	0.87	0.57
Nominal Yield	6.32	6.22	6.34	6.55	6.79

The yields are reported by averaging the monthly yields conditional on the regime state variables. The yields are annualized.

Table 3.11: Model Implications – Equity and Bond Premium

	1y	2y	3y	4y	5y
Conditional on expansion regime					
Real bond holding period return premium	-0.22	-0.43	-0.61	-0.77	-0.9
Nominal bond holding period return premium	0.42	0.81	1.14	1.44	1.69
Equity premium	5.05		$corr(r_b^{(5)}, r_d)$		0.23
Conditional on deep-recession regime					
Real bond holding period return premium	-0.27	-0.53	-0.75	-0.95	-1.11
Nominal bond holding period return premium	-0.6	-1.06	-1.41	-1.69	-1.95
Equity premium	7.36		$corr(r_b^{(5)}, r_d)$		-0.2
Conditional on contraction regime					
Real bond holding period return premium	-0.28	-0.53	-0.76	-0.95	-1.12
Nominal bond holding period return premium	0.57	1.06	1.48	1.85	2.18
Equity premium	6.43		$corr(r_b^{(5)}, r_d)$		0.06

The risk premia are reported by averaging 1-month nominal/real bond holding period excess returns and 1-month equity excess return, conditional on the regime state variables. The statistics are annualized.

Table 3.12: Model Implications, Single Factor Projection

n	2y	3y	4y	5y
U.S. data (1952-2009)				
Coeff.	0.48	0.13	1.24	1.44
Std. Dev.	0.13	0.25	0.36	0.45
R2	0.15	0.17	0.18	0.16
Model: Nominal Regression				
Coeff.	0.45	0.85	1.2	1.48
Std. Dev.	0.03	0.06	0.08	0.1
R2	0.107	0.113	0.114	0.112
Model: Real Regression				
Coeff.	0.45	0.85	1.2	1.53
Std. Dev.	0.04	0.07	0.1	0.13
R2	0.079	0.074	0.07	0.066

Nominal return predictability test in US bond market. Monthly observations of 1-5 year yields on US for June 1952 to Dec 2009 are from Fama-Bliss Dataset. Single Factor Projection reports the slope coefficients $B_{m,n}^{\$}$ and R^2 in single latent factor regression $rx_{t+m,n}^{\$} = const + b_{m,n}\widehat{rx^{\$}}_{t,m} + error$, where $rx_{t+m,n}^{\$}$ is an m-months excess return on n-period bond, and $\widehat{rx^{\$}}_{t,m}$ corresponds to a single bond factor obtained from a first-stage projection of average bond returns on three forward rates. Model implied slope coefficients and R^2 in single latent factor regression are based on a very long simulation of 3000 years' monthly observations. Standard errors are Newey-West adjusted with 10 lags, computed with GMM approach.
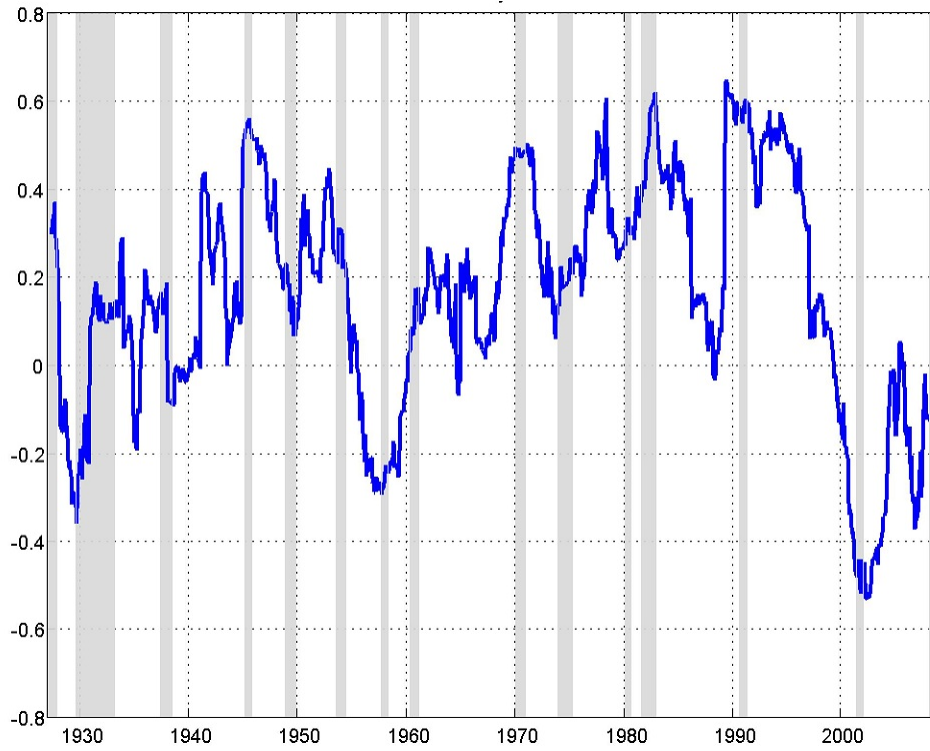
FIGURE 3.1: Correlation between Stock and Nominal Bond Return

This figure plots the time-varying correlation between stock and 10-year nominal bond returns, which is calculated based on a 3-year centered moving window of real monthly stock and bond returns. Shaded areas denote NBER recessions.

FIGURE 3.2: CAPM beta of Inflation

This figure summarizes the CAPM beta of inflation. I use a rolling 5-year window of quarterly data and a first-order quarterly VAR for inflation, stock returns (real), and the three-month treasury bill returns to calculate inflation shocks. The k-period beta is defined as $\widehat{Cov}_t\left(\sum_{i=1}^{k} \pi_{t+i}, \sum_{i=1}^{k} r_{m,t+i}\right) / \widehat{Var}_t\left(\sum_{i=1}^{k} r_{m,t+i}\right)$. Shaded areas denote NBER recessions.
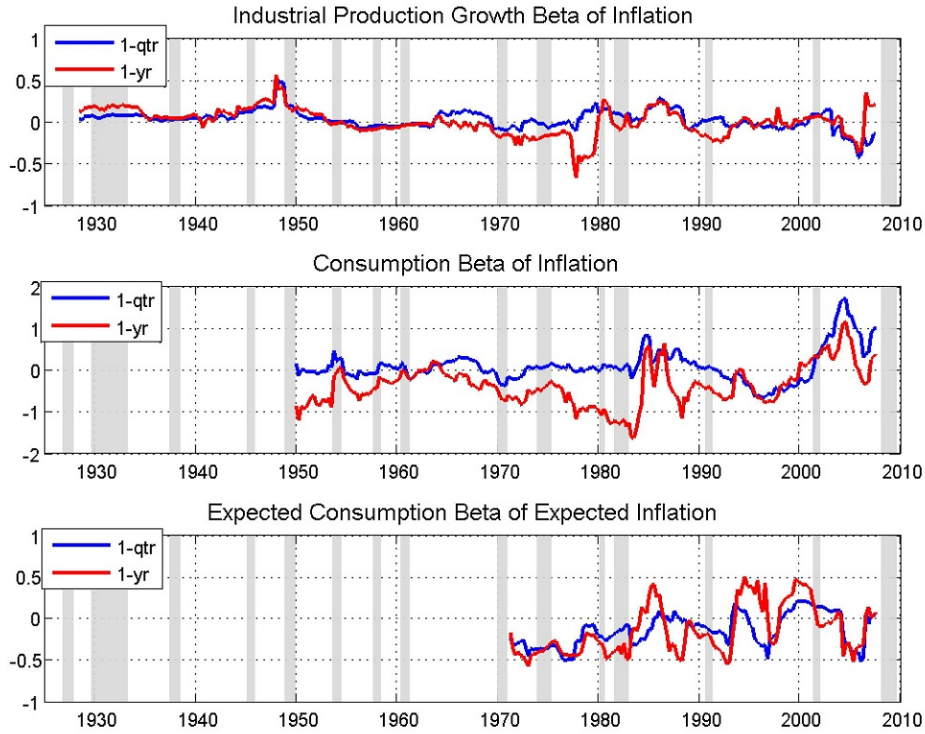
FIGURE 3.3: Consumption Beta of Inflation and Expected Inflation
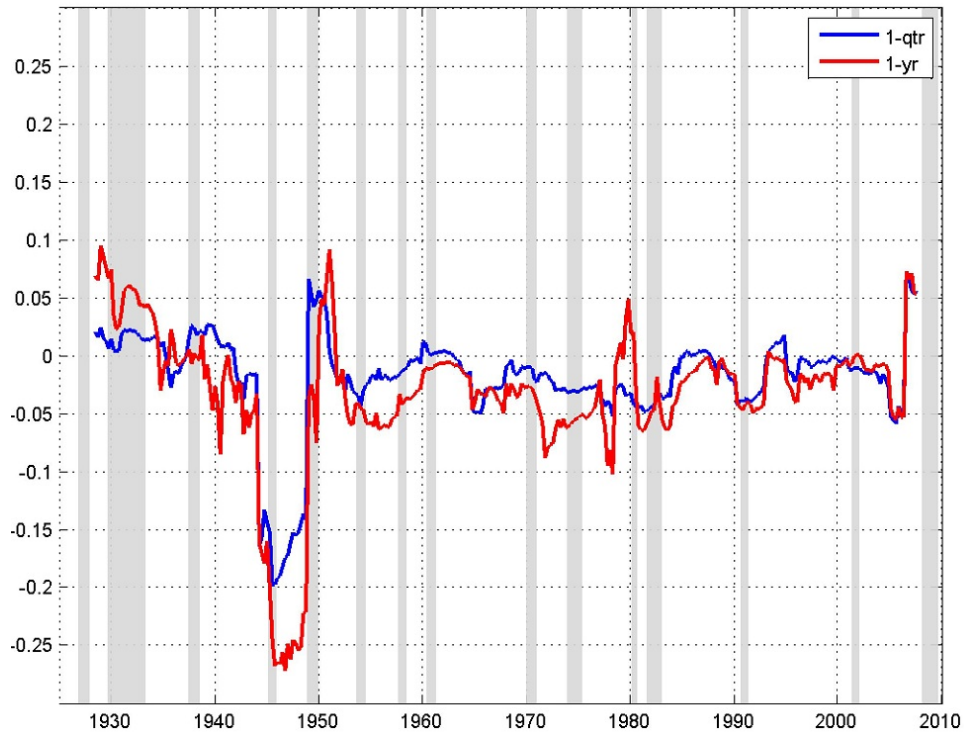
This figure summarizes the consumption beta of inflation. I use a rolling 5-year window of quarterly data and a first-order quarterly VAR for inflation, industrial production growth (or consumption growth) to calculate inflation shocks. The k-period beta is defined as

$$
\widehat{Cov_t}\left(\sum_{i=1}^{k}\pi_{t+i}, \sum_{i=1}^{k}\Delta c_{t+i}\right)\Big/\widehat{Var_t}\left(\sum_{i=1}^{k}\Delta c_{t+i}\right)
$$

. In the first panel, Δc_{t+1} stands for industrial production growth, ranging from 1926Q1-2009Q4, from FRED dataset. In the second panel, Δc_{t+1} stands for consumption growth, ranging from 1947Q1-2009Q4, from BEA. In the last panel, Δc_{t+1} and π_{t+1} stands for GDP growth and GDP deflator expectations from Surveys of Professional Forecasters for the period of 1968Q4 - 2009Q4. Shaded areas denote NBER recessions.
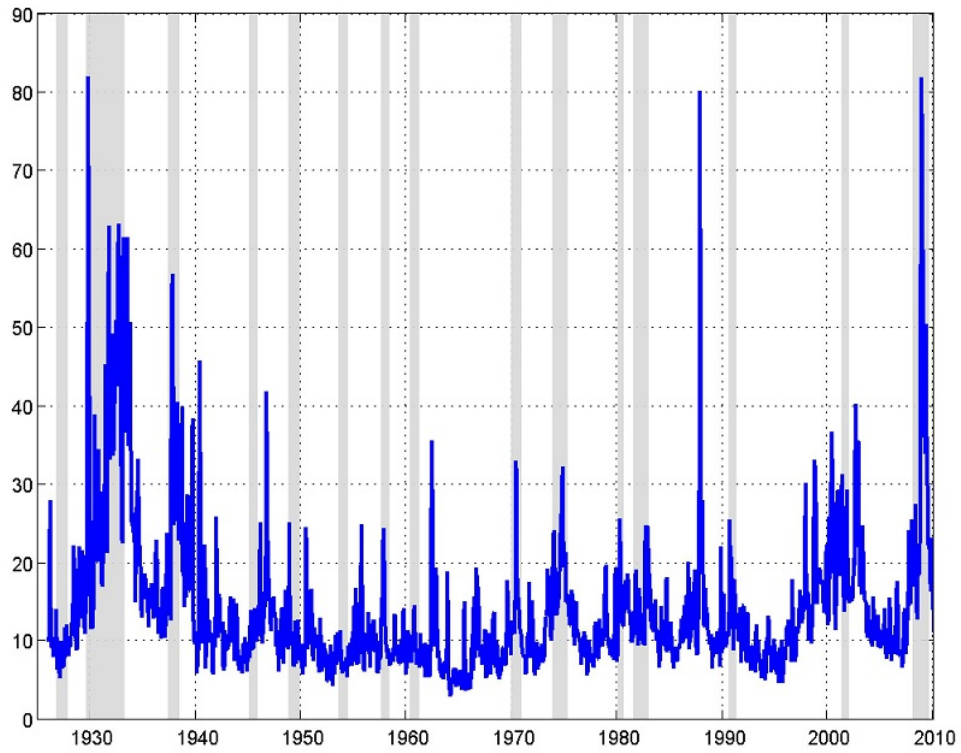
FIGURE 3.4: Integrated Volatility of Equity Return

The monthly integrated volatility is computed from squared sum of daily returns on CRSP value weighted portfolio comprising the stocks traded in the NYSE, AMEX and NASDAQ, from Jan 1926 - Dec 2009. Shaded areas denote NBER recessions.

Appendix A

Appendix to Chapter 1

A.1 Derivations of Equilibrium Conditions from Household Problem

In the benchmark model, the representative household is making optimal consumption and saving decisions by maximizing recursive preference (Kreps and Porteus, 1978; Epstein and Zin, 1989):

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

subject to the budget constraint:

$$C_t + B_t = B_{t-1} R_{f,t-1} + \pi_t.$$

The Euler equation gives:

$$E_t \left[M_{t+1} \right] R_{f,t} = 1,$$

in which the stochastic discount factor is:

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

A.2   Derivations of Equilibrium Conditions from Bank's Problem

Based on the recursive representation of a typical individual bank's optimization problem as stated in (1.21).

Use the law of the motion to substitute out n_{t+1}. Let $\eta(\mathbf{b}_t)$ denote the Lagrangian multiplier with respect to the participation constraint.

The first order condition with respect to s_{t+1} is:

$$(1 + \eta(\mathbf{b}_t)) E_t \left[M_{t+1} \{ \lambda + (1-\lambda) \mu(\mathbf{b}_{t+1}) \} \{ Q(\mathbf{b}_{t+1}) + Y_{t+1} - Q(\mathbf{b}_t) R_{f,t} \} \right] = \theta \eta(\mathbf{b}_t) Q(\mathbf{b}_t).$$

$$(A.1)$$

The envelope condition with respect to n_t is:

$$\mu(\mathbf{b}_t) = (1 + \eta(\mathbf{b}_t)) E_t \left[M_{t+1} \{ \lambda + (1-\lambda) \mu(\mathbf{b}_{t+1}) \} \right] R_{f,t}. \qquad (A.2)$$

The complementary slackness conditions are:

$$\eta(\mathbf{b}_t) \left[\mu(\mathbf{b}_t) n_t - \theta s_t Q(\mathbf{b}_t) \right] = 0, \qquad (A.3)$$

$$\eta(\mathbf{b}_t) \geq 0,$$

$$\mu(\mathbf{b}_t) n_t - \theta s_t Q(\mathbf{b}_t) \geq 0. \qquad (A.4)$$

Since all the individual banks make the same decision, it allow us to have equilibrium conditions at the aggregation level. Equations (A.1), and (A.2) stay the same, and the complementary slackness conditions become:

$$\eta(\mathbf{b}_t)\left[\mu(\mathbf{b}_t)N_t - \theta Q(\mathbf{b}_t)\right] = 0, \tag{A.5}$$

$$\eta(\mathbf{b}_t) \geq 0,$$

$$\mu(\mathbf{b}_t)N_t - \theta Q(\mathbf{b}_t) \geq 0. \tag{A.6}$$

Given $\{\mu(\mathbf{b}_{t+1}), Q(\mathbf{b}_{t+1})\}$, I define

$$v(\mathbf{b}_t) = \lambda + (1-\lambda)E_t\left[M_{t+1}\mu(\mathbf{b}_{t+1})\right]R_{f,t}, \tag{A.7}$$

$v(\mathbf{b}_t)$ is the shadow price of net worth at date t if the constraint is not binding for any bank[1]. Also, define

$$P(\mathbf{b}_t) = \frac{E_t\left[M_{t+1}\{\lambda + (1-\lambda)\mu(\mathbf{b}_{t+1})\}(Q(\mathbf{b}_{t+1}) + Y_{t+1})\right]}{v(\mathbf{b}_t)}. \tag{A.8}$$

$P(\mathbf{b}_t)$ is the equilibrium price of the Lucas tree in the case where the participation constraint does not bind for any bank. Note that $v(\mathbf{b}_t)$ and $P(\mathbf{b}_t)$ are completely determined once the functional form of $\{\mu(\mathbf{b}_{t+1}), Q(\mathbf{b}_{t+1})\}$ is known. (The prices M_{t+1} and R_f are trivially determined because it is an endowment economy.)

It is easy to show that the Lagrangian multiplier $\eta(\mathbf{b}_t)$ can be expressed as

$$\eta(\mathbf{b}_t) = \frac{\mu(\mathbf{b}_t)}{v(\mathbf{b}_t)} - 1. \tag{A.9}$$

---

[1] Note, here I adopt the following mathematical definition of a "binding" constraint. "Binding" means the Lagrangian multiplier must be strictly positive. It rules out the case where the constraint holds with equality but the Lagrangian multiplier is zero.

Use this relationship to substitute out $\eta(\mathbf{b}_t)$, it is easy to show that the equilibrium conditions are summarized by the following lammas.

Lemma 4. *(Equilibrium Price of the Lucas Tree)*

Given the equilibrium pricing functional $\{Q(\mathbf{b}_{t+1}), \mu(\mathbf{b}_{t+1})\}$, we consider the equilibrium pricing functional $Q(\mathbf{b}_t)$

1. *Suppose*

$$v(\mathbf{b}_t) N_t \geq \theta P(\mathbf{b}_t), \tag{A.10}$$

   then in equilibrium, we must have:

   - $Q(\mathbf{b}_t) = P(\mathbf{b}_t)$, *where $P(\mathbf{b}_t)$ is given in (A.8).*

   - *The constraint (A.4) is not binding for any bank in the sense that the Lagrangian multiplier on the constraint must be 0.*

2. *Suppose*

$$0 < v(\mathbf{b}_t) N_t < \theta P_t(\mathbf{b}_t), \tag{A.11}$$

   then in equilibrium, we must have:

   - *The price of the Lucas tree, $Q(\mathbf{b}_t)$, satisfies:*

$$Q(\mathbf{b}_t) = \frac{v(\mathbf{b}_t)[P(\mathbf{b}_t) + N_t]}{\theta + v(\mathbf{b}_t)} < P_t(\mathbf{b}_t). \tag{A.12}$$

   - *The constraint (A.4) is binding for all banks in the sense that the Lagrangian multiplier on the constraint must be strictly positive.*

3. *If $N_t \leq 0$, then equilibrium cannot exist.*

The three cases discussed above provide a complete characterization of the equilibrium at state $\mathbf{b}_t$ given the price and quantities at state $\mathbf{b}_{t+1}$. The first part of the lemma says that if the total net worth of the banking sector is large enough, then the participation constraint will not bind, and the equilibrium price of the Lucas tree is given by (A.8). Note, however, even if the constraint does not bind at time t, the price is still different from that in a frictionless Lucas model. This is because the possibility of a binding constraint in the future will affect today's price.

The second part of the lemma implies that if the total net worth is positive, but small, then the participation constraint will bind, and the equilibrium price has to drop (relative the price P) to lower the outside value of the bankers. The third part of the condition says total net worth can never be zero or negative in equilibrium.

Given the above lemma, we can derive the functional form of $V(\mathbf{b}_t, n_t)$. It is straightforward to show that if $V(\mathbf{b}_{t+1}, n_{t+1})$ is linear in n_{t+1} as in (1.20), then $V(\mathbf{b}_t, n_t) = \mu(\mathbf{b}_t) n_t$, and $\mu(\mathbf{b}_t)$ is given by the following lemma.

Lemma 5. *(Equilibrium Value Function of the Financial Intermediary)*

Given the equilibrium pricing functional $\{Q(\mathbf{b}_{t+1}), \mu(\mathbf{b}_{t+1})\}$, we consider the equilibrium pricing functional $\mu(\mathbf{b}_t)$:

1. *Under condition (A.10),*

$$\mu(\mathbf{b}_t) = v(\mathbf{b}_t).$$

2. *Under condition (A.11),*

$$\mu(\mathbf{b}_t) = v(\mathbf{b}_t) \times \frac{\theta\{P(\mathbf{b}_t) + N(\mathbf{b}_t)\}}{N(\mathbf{b}_t)[\theta + v(\mathbf{b}_t)]}. \tag{A.13}$$

To summarize the above two lemmas, under condition (A.10), the constraint does not bind, and $\{\mu(\mathbf{b}_t), Q(\mathbf{b}_t)\}$ can be constructed recursively from $\{\mu(\mathbf{b}_{t+1}), Q_{t+1}(\mathbf{b}_{t+1})\}$:

$$\mu(\mathbf{b}_t) = \lambda + (1 - \lambda) E_t [M_{t+1} \mu(\mathbf{b}_{t+1})] R_{f,t}, \tag{A.14}$$

and

$$Q_t(\mathbf{b}_t) = \frac{E_t [M_{t+1} \{\lambda + (1 - \lambda) \mu(\mathbf{b}_{t+1})\} \{Q(\mathbf{b}_{t+1}) + Y_{t+1}\}]}{\mu(\mathbf{b}_t)}. \tag{A.15}$$

Note that

$$\frac{E_t [M_{t+1} \{\lambda + (1 - \lambda) \mu(\mathbf{b}_{t+1})\}]}{\mu(\mathbf{b}_t)} = \frac{1}{R_{f,t}}.$$

Note that on the right hand side of equations (A.14) and (A.15), all quantities are known except $\{\mu(\mathbf{b}_{t+1}), Q(\mathbf{b}_{t+1})\}$. So the system (A.14) and (A.15) defines a mapping

$$\{\mu(\mathbf{b}_t), Q(\mathbf{b}_t)\} = T\{\mu(\mathbf{b}_{t+1}), Q(\mathbf{b}_{t+1})\}.$$

Under condition (A.11), I similarly define the mapping $\{\mu(\mathbf{b}_{t+1}), Q(\mathbf{b}_{t+1})\} \implies \{\mu(\mathbf{b}_t), Q(\mathbf{b}_t)\}$. To save notation, we can summarize the two case with a compact notation. Using (A.13),

$$Q(\mathbf{b}_t) = \frac{v(\mathbf{b}_t) P(\mathbf{b}_t) + v(\mathbf{b}_t) N(\mathbf{b}_t) \wedge \theta P(\mathbf{b}_t)}{v(\mathbf{b}_t) + \theta}. \tag{A.16}$$

Also,

$$\mu(\mathbf{b}_t) = v(\mathbf{b}_t) \vee \frac{\theta Q(\mathbf{b}_t)}{N_t}. \tag{A.17}$$

Here I used the short-hand notation $x \wedge y \equiv \min\{x, y\}$ and $x \vee y = \max\{x, y\}$. Obviously, $Q_t(\mathbf{b}_t) \leq P(\mathbf{b}_t)$ and $\mu(\mathbf{b}_t) \geq v(\mathbf{b}_t)$, and strict inequality holds if and only if (A.11) is true, in which case the participation constraint is binding.

183

A.3   Data Sources

Consumption: Per capita consumption data are from the National Income and Product Accounts (NIPA) annual data reported by the Bureau of Economic Analysis (BEA). The data are constructed as the sum of consumption expenditures on nondurable goods and services (Table 1.1.5, lines 5 and 6) deflated by corresponding price deflators (Table 1.1.9, lines 5 and 6).

Dividend: The dividend process is constructed from VWRETD and VWRETX, i.e. the value weighted return on NYSE/AMEX including and excluding dividends, taken from CRSP. The construction of price-dividend ratio follows the data appendix in Bansal, Khatchatrian and Yaron (2005).

Earnings: Corporate earnings data are from corporate profits (earnings) after tax (in billions of dollars) from National Income and Product Accounts (NIPA) data reported by the Bureau of Economic Analysis (BEA) (Table 1.14, line 29). The construction of price-earnings ratio follows the data appendix in Bansal, Khatchatrian and Yaron (2005).

Market Return: Nominal market return is the value weighted return on NYSE/AMEX including dividends taken from CRSP. The real market return is computed by deflating the nominal return by corresponding price deflators (Table 1.1.9, lines 5 and 6).

Risk-free Rate: The nominal risk-free rate is measured by the annual 3-month T-Bill return. The real risk-free rate is computed by subtracting the nominal risk-free rate by expected inflation, a procedure detailed in Beeler and Campbell (2012).

TED Spread: Computed by the difference between annualized 3-month LIBOR

184

rate and 3-month T-bill rate. Both series are from FRED dataset.

Leverage Ratio: I follow Adrian, Moench and Shin (2011)'s composition of the aggregate financial intermediary sector. From Flow of Funds Table in U.S. I aggregate the assets and liabilities of each component, and then compute the aggregate leverage ratio based on:

$$\text{Leverage}_t = \frac{\text{Aggregate Financial Assets}_t}{\text{Aggregate Financial Assets}_t - \text{Aggregate Liabilities}_t}$$

Integrated Volatility: Integrated variance is the sum of squared daily stock returns on NYSE/AMEX. Integrated volatility is the square root of integrated variance. The daily value weighted return data on NYSE/AMEX including dividends are taken from CRSP.

Table A.1: Composition of Aggregate Financial Intermediary Sector

Symbol		Descriptions
FINBANK		**Banks**
	CBSI	Charted depository institutions, excluding credit unions
	CU	Credit unions
FINPI		**Pension Funds and Insurances**
	PCIC	Property-casualty insurance companies
	LIC	Life insurance companies
	PPF*	Private pension funds
	SLGERF*	State & local government employee retirement funds
	FGRF*	Federal government retirement funds
FINMF		**Mutual Funds**
	MMMF*	Money market mutual funds
	MF*	Mutual funds
	CEF*	Closed-end funds and exchange-traded funds
SHADBANK		**Shadow Banks**
	MORTPOOL*	Agency- and GSE-backed mortgage pools
	ABS	Issuers of asset-backed securities
	FINCO	Finance companies
	FUNDCORP	Funding corporations
SBRDLR		**Security brokers and dealers**

This Table is based on the definitions in Adrian, Moench and Shin (2010). The component intermediaries denoted by "*" means they are only financed by equity.

A.4 Additional Details of the Numerical Solutions

I approximate the i.i.d. consumption shock $\varepsilon_{y,t}$ by a finite-state Markov chain. I fix 5 realizations evenly spaced on the bounded interval $[-2 \times \sigma, 2 \times \sigma]$, in which σ denotes the consumption volatility, and I confirm that the lowest realization of the consumption shock satisfies the parameter requirement as emphasized in Section 1.3.3. The probability vector for these 5 states are pinned down by the following five conditions: (1) matching the first four moments in the demeaned consumption growth process; (2) The probabilities in the vector sum up to 1.

I specify 500 grids of the state variable $\mathbf{b}$, evenly spanned on the the state space $[0, \overline{\mathbf{b}}]$, in which $\overline{\mathbf{b}}$ denotes the highest possible debt level supported by the equilibrium. It is endogenously determined and updated in each iteration. I start with a large enough $\overline{\mathbf{b}}$ in the initial iteration, and update $\overline{\mathbf{b}}$ in each iteration to make sure that $\widehat{n}$ is strictly positive. The supportable state space of $\mathbf{b}$ converges in the iterative procedure.

I use constant price-dividend ratio and unit shadow price of net worth, suggested by the standard Lucas economy without frictions, as the initial guesses to start the iteration. In the subsequent iterations, I use point-wise linear spline to approximate the new price functions. The critical value for determining the constrained region corresponds to the last grid where the Lagrangian multiplier associated with the constraint is strictly positive.

In the computation, I use extensively the approximation toolkit in the CompEcon Toolbox of Miranda and Fackler (2002).

A.5 Additional Details of Constructing Simulation Accuracy Test Statistic

Based on equations (A.7) and (A.8) and the Euler equation from household problem, the prediction errors corresponding to the first-order conditions are given by

$$w_{t+1} = \begin{bmatrix} (1 + \eta_t) \left[\widetilde{M}_{t+1} \left(R_{m,t+1} - R_{f,t} \right) \right] - \theta \eta_t \\ (1 + \eta_t) \widetilde{M}_{t+1} R_{f,t} - \mu_t \\ M_{t+1} R_{f,t} - 1 \end{bmatrix}.$$

I use the following vector of five instrument variables:

$$h_t = [1, g_t, \widehat{n}_t, \widehat{n}_{t-1}, \widehat{n}_{t-2}].$$

Following Den Haan and Marcet (1994), the accuracy test consists of obtaining long simulations of the process and calculating

$$W_T = \frac{\sum\limits_{t=1}^{T} \overline{w}_{t+1} \otimes \overline{h}_t}{T},$$

where $\overline{w}_t$ and $\overline{h}_t$ are calculated with simulated data, and T denotes the simulation length. The accuracy test statistic is constructed as $T W_T' A_T^{-1} W_T$, where A_T denotes a consistent estimator of covariance matrix of W_T. In the implementation, I simulate the model 500 times, each with 1000 annual observations. I use Newey-West estimator A_T of the covariance matrix. By proposition 1 in Den Haan and Marcet (1994), under null hypothesis that the numerical solution is accurate, the simulation accuracy test statistic has a χ^2 distribution, with degree of freedom of 15. The simulation accuracy test results are insensitive to the number of instrument variable I choose.

188

Appendix B

Appendix to Chapter 2

B.1   Aggregation of Production Units

Lemma 6. *Suppose there are m types of firms. For $i = 1, 2, 3, \cdots m$, the productivity of the type i firm is denoted by $A(i)$, and the total measure of the type i firm is denoted by $K(i)$. The production technology of the type i firm is given by*

$$y(i) = [A(i) \, n(i)]^{1-\alpha},$$

where $n(i)$ denotes the labor hired at firm i. The total labor supply in the economy is N. Then total output is given by

$$Y = \left[\sum_{i=1}^{m} K(i) \left[\frac{A(i)}{A(1)} \right]^{\frac{1-\alpha}{\alpha}} \right]^{\alpha} [A(1) \, N]^{1-\alpha}.$$

Proof. Without loss of generality, we assume that firms of the same type employ the

same amount of labor. In this case, the total output in the economy is given by

$$Y = \max \sum_{i=1}^{m} K(i) A(i)^{1-\alpha} n(i)^{1-\alpha} \tag{B.1}$$

$$subject\ to\ \sum_{i=1}^{m} K(i) n(i) = N$$

The first-order condition of the above optimization problem implies that for all i,

$$\frac{n(i)}{n(1)} = \left(\frac{A(i)}{A(1)}\right)^{\frac{1-\alpha}{\alpha}}$$

Using the resource constraint, we determine the labor employed in firm 1:

$$\sum_{i=1}^{m} K(i) \left(\frac{A(i)}{A(1)}\right)^{\frac{1-\alpha}{\alpha}} n(1) = N$$

This implies that

$$n(1) = \left[\sum_{i=1}^{m} K(i) \left[\frac{A(i)}{A(1)}\right]^{\frac{1-\alpha}{\alpha}}\right]^{-1} N \tag{B.2}$$

Therefore, the total production is given by

$$
\begin{aligned}
Y &= \sum_{i=1}^{m} K(i) A(i)^{1-\alpha} \left[\left(\frac{A(i)}{A(1)}\right)^{\frac{1-\alpha}{\alpha}} n(1)\right]^{1-\alpha} \\
&= \left[A(1)^{-\frac{1-\alpha}{\alpha}} n(1)\right]^{1-\alpha} \sum_{i=1}^{m} K(i) A(i)^{\frac{1-\alpha}{\alpha}} \\
&= \left[\sum_{i=1}^{m} K(i) A(i)^{\frac{1-\alpha}{\alpha}}\right]^{\alpha} N^{1-\alpha} \\
&= A(1) \left[\sum_{i=1}^{m} K(i) \left(\frac{A(i)}{A(1)}\right)^{\frac{1}{\alpha}}\right]^{\alpha} N^{1-\alpha}
\end{aligned}
$$

190

Plugging in the expression for $n(1)$ in equation (B.2), we have

$$
\begin{aligned}
Y &= \left[A(1)^{-\frac{1-\alpha}{\alpha}} \left[\sum_{i=1}^{m} K(i) \left[\frac{A(i)}{A(1)} \right]^{\frac{1-\alpha}{\alpha}} \right]^{-1} N \right]^{1-\alpha} \sum_{i=1}^{m} K(i) A(i)^{\frac{1-\alpha}{\alpha}} \\
&= \left[\sum_{i=1}^{m} K(i) A(i)^{\frac{1-\alpha}{\alpha}} \right]^{\alpha} N^{1-\alpha} \\
&= \left[\sum_{i=1}^{m} K(i) \left[\frac{A(i)}{A(1)} \right]^{\frac{1-\alpha}{\alpha}} \right]^{\alpha} A(1)^{1-\alpha} N^{1-\alpha}
\end{aligned}
$$

as needed. $\qquad\square$

At time t, there are $t+1$ types of operating production units in the economy, namely, production units of generation -1, 0, 1, $\cdots$, $t-1$. The measures of these production units are $(1-\delta_K)^t K_0$, $(1-\delta_K)^{t-1} M_0$, $(1-\delta)^{t-2} M_1$, $\cdots$, M_{t-1}. Using the above lemma, at date t, the total production in the economy is given by

$$
Y_t = A_t \left[(1-\delta_K)^t K_0 + \sum_{j=0}^{t-1} (1-\delta_K)^{t-j-1} E_j \left(\frac{A_t^j}{A_t} \right)^{\frac{1-1}{\alpha}} \right]^{\alpha} N_t^{1-\alpha}.
$$

Clearly, if we define the $\{K_t\}_{t=0}^{\infty}$ according to (9), the aggregate production function can be summarized as in (5).

B.2 Estimation Details and Aggregate Data Sources

B.2.1 Robustness analysis for firm-level regressions

Endogeneity and the dynamic error component model. We follow Blundell and Bond (2000) and write the firm-level production function as follows:

$$\ln y_{i,j,t} \quad = \quad z_{ij,} + w_t + \alpha_1 \ln k_{i,j,t} + \alpha_2 \ln n_{i,j,t} + v_{i,j,t} + u_{i,j,t} \tag{B.1}$$

$$v_{i,j,t} \quad = \quad \rho v_{i,j,t-1} + e_{i,j,t},$$

where $z_{i,j}$, w_t and $v_{i,j,t}$ indicate a firm fixed effect, a time-specific intercept, and a possibly autoregressive productivity shock, respectively. The residuals from the regression are denoted by $u_{i,j,t}$ and $e_{i,j,t}$ and are assumed to be white noise processes. The model has the following dynamic representation:

$$\Delta \ln y_{i,j,t} \quad = \quad \rho \Delta \ln y_{i,j,t-1} + \alpha_{1,j} \Delta \ln k_{i,j,t} - \rho \alpha_{1,j} \Delta \ln k_{i,j,t-1} \tag{B.2}$$

$$+ \alpha_2 \Delta \ln n_{i,j,t} - \rho \alpha_2 \Delta \ln n_{i,j,t-1} + (\Delta w_{j,t} - \rho \Delta w_{j,t-1}) + \Delta \kappa_{i,j,t},$$

where $\kappa_{i,j,t} = e_{i,j,t} + u_{i,j,t} - \rho u_{i,j,t-1}$. Let $x_{i,j,t} = \{\ln (k_{i,j,t}), \ln (n_{i,j,t}), \ln (y_{i,j,t})\}$. Assuming that $E[x_{i,j,t-l} e_{i,t}] = E[x_{i,j,t-l} u_{i,t}] = 0$ for $l > 0$ yields the following moment conditions:

$$E[x_{i,j,t-l} \Delta \kappa_{i,t}] \quad = \quad 0 \text{ for } l \geq 3$$

$$E[\Delta x_{i,j,t-l} \Delta \kappa_{i,t}] \quad = \quad 0 \text{ for } l \geq 3.$$

that are used to conduct a consistent GMM estimation of (B.2). Given the estimates $\widehat{\alpha}_{1,j}$ and $\widehat{\alpha}_{2,j}$, log productivity of firm i is computed as:

$$\ln \widehat{a}_{i,j,t} = y_{i,j,t} - \widehat{\alpha}_{1,j} k_{i,j,t} - \widehat{\alpha}_{2,j} n_{i,j,t}. \tag{B.3}$$

192

We apply this method for regressions (5)–(6) and (9)–(10) of Table B.1. In all specifications, the correlation between firm and aggregate productivity is increasing in capital age, consistent with our main results reported in Table 2.2.

Endogeneity and fixed effects. An alternative way to estimate the production function avoiding endogeneity issues is to work with the following regression:

$$\ln y_{i,j,t} = v_j + z_{i,j} + w_{j,t} + \alpha_{1,j} \ln k_{i,j,t} + \alpha_{2,j} \ln n_{i,j,t} + u_{i,j,t}. \tag{B.4}$$

The parameters v_j, $z_{i,j}$, and $w_{j,t}$ indicate an industry dummy, a firm fixed effect, and an industry-specific time dummy, respectively. The residual from the regression is denoted by $u_{i,j,t}$. Given our point estimate of $\widehat{\alpha}_{1,j}$ and $\widehat{\alpha}_{2,j}$, we can use equation (B.3) to estimate $\Delta a_{i,j,t}$. Given this estimation of firms' productivity, we proceed as before in estimating equation (19). The results are summarized in regression (1)–(4) and (7)–(8) of Table B.1 and are consistent with those reported in Table 2.2.

Sample Selection Bias. If exits caused by exposure to negative aggregate productivity shocks are correlated with firm age, they can induce an upward bias in our estimate of ξ_3, the coefficient that measures variation in productivity exposure due to age. We correct for sample selection bias in Regressions (2) and (3) of Tables 2.2 and 2.3. Our result that firms' exposure to aggregate productivity shocks is increasing in age is robust even after we control for potential sample selection bias.

We implement the Heckman (1979) two-stage procedure in regression (2). First, we project an indicator variable of firms' exit on a vector of observables, including the Altman (2000) Score, size (measured by total book value of assets in real terms),

193

size squared, R&D expenditure-sales ratio, earnings over sales, capital-labor ratio, and aggregate productivity growth. Second, we compute the implied Inverse Mills Ratio (IMR) (see Greene (2002)) and include it as an additional independent variable in equation (19). In regression (3), we include observations only in years with positive aggregate productivity shocks. Overall, our point estimate for ξ_3 is positive and significant across all specifications.

B.2.2 Aggregate Data Sources

Per capita consumption (C_t) data are from the National Income and Product Accounts (NIPA) annual data reported by the Bureau of Economic Analysis (BEA). The data are constructed as the sum of consumption expenditures on nondurable goods and services (Table 1.1.5, lines 5 and 6) deflated by corresponding price deflators (Table 1.1.9, lines 5 and 6).

Physical investment (I_t) data are also from the NIPA tables. We measure physical investment by fixed investment (Table 1.1.5, line 8) minus information processing equipment and software (Table 5.5.5, line 3) deflated by its price deflator (Table 1.1.9, line 8). Information processing equipment and software is interpreted as investment in intangible capital and is therefore subtracted from fixed investment.

Measured output ($Y_{M,t}$) is the sum of total consumption and physical investment, that is, $C_t + I_t$. We exclude government expenditure and net export because not

explicitly modelled in our economy. Notice also that the NIPA tables do not account for J_t over the sample 1929–2003.

Intangible investment (J_t) is constructed as in Corrado et al. (2006) from 1953 to 2003. We include expenses in brand equity, firm-specific resources, R&D and computerized information. Prior to 1953 data are not available. Data and notes on our sources are available upon request.

Labor (N_t) is measured as the total number of full-time and part-time employees as reported in the NIPA table 6.4. Data are annual from 1929 to 2003. In table B.3, we divide N by total population.

We follow Hall (2001) and construct physical capital (K_t) according to the following recursion: $K_t = (1 - \delta_K)K_{t-1} + I_t$, where $\delta_K = .11$ as in our calibration. This recursion is initialized in 1929 by imposing $K_{1929} = I_{1929}/\delta_K$ as in Hall (2001).

Both the market returns (r_M^L) and the HML factor $(r_k^L - r_S^L)$ are from the Fama-French dataset available online on K. French's webpage:
http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/ftp/
F-F_Research_Data_Factors.zip.

The nominal risk-free rate is measured by the annual 3-month T-Bill return. The real risk-free rate (r^f) is computed by subtracting realized inflation from the nominal risk-free rate.

B.3 Calculation of the Macaulay duration and model extensions

B.3.1 Duration.

In this section, we derive a recursive relation that can be used to calculate the Macaulay duration of growth options and assets in place defined in equation (2.25). We first prove the following lemma.

Lemma 7. *Let MD_t and P_t be the time-t Macaulay duration and present value of the cash flow process $\{CF_t\}_{t=1}^{\infty}$, respectively; then*

$$MD_t \cdot P_t = E_t\left[\Lambda_{t+1} \cdot CF_{t+1}\right] + E_t\left[\Lambda_{t+1}(1 + MD_{t+1})P_{t+1}\right]. \tag{B.1}$$

Furthermore, if $CF_t = CF_{1,t} + CF_{2,t}$ for all t, then

$$\begin{aligned}
MD_t \cdot P_t = E_t\left[\Lambda_{t+1}\left(CF_{1,t+1} + CF_{2,t+1}\right)\right] \quad &+ \quad E_t\left[\Lambda_{t+1}(1 + MD_{1,t+1})P_{1,t+1}\right] \dots \tag{B.2} \\
&+ \quad E_t\left[\Lambda_{t+1}(1 + MD_{2,t+1})P_{2,t+1}\right],
\end{aligned}$$

where $P_{i,t}$ and $MD_{i,t}$ denote the price and Macaulay duration of cash flow $CF_{i,t}$ for $i = 1, 2$.

Proof. By the definition of Macaulay duration,

$$\begin{aligned}
MD_t \cdot P_t &= E_t\left[\sum_{j=1}^{\infty} j\Lambda_{t,t+j}CF_{t+j}\right] \\
&= E_t\left[\Lambda_{t,t+1}CF_{t+1}\right] + E_t\left[E_{t+1}\left[\sum_{j=2}^{\infty} j\Lambda_{t,t+j}CF_{t+j}\right]\right] \\
&= E_t\left[\Lambda_{t,t+1}CF_{t+1}\right] \\
&\quad + E_t\left[\Lambda_{t,t+1}\left(E_{t+1}\left[\sum_{j=2}^{\infty}(j-1)\Lambda_{t+1,t+j}CF_{t+j}\right] + E_{t+1}\left[\sum_{j=2}^{\infty}\Lambda_{t+1,t+j}CF_{t+j}\right]\right)\right] \\
&= E_t\left[\Lambda_{t,t+1}\left(CF_{t+1} + MD_{t+1}P_{t+1} + P_{t+1}\right)\right],
\end{aligned}$$

196

as needed.

Equation (B.1) can then be proved by applying the definition of present value and duration separately for cash flow $\{CF_{1,t}\}_{t=1}^{\infty}$ and $\{CF_{2,t}\}_{t=1}^{\infty}$. □

Let $MD_{S,t}$ denote the Macaulay duration of a growth option created at the end of period t. Let $MD_{K,t}$ refer to the Macaulay duration of a generation-0 production unit survived until the end of period t. We show that $MD_{S,t}$ and $MD_{K,t}$ satisfy the following recursive relation:

$$MD_{S,t} \cdot q_{S,t} = -E_t \left[\Lambda_{t+1} \frac{I_{t+1}}{S_{t+1}} \right] \quad + \quad E_t \left[\Lambda_{t+1} \frac{G\left(I_{t+1}, S_{t+1}\right)}{S_{t+1}} (1 + MD_{K,t+1}) q_{K,t+1} \varpi_{t+1} \right] \dots \tag{B.3}$$

$$+ \quad E_t \left[\Lambda_{t+1} \left(1 - \frac{G\left(I_{t+1}, S_{t+1}\right)}{S_{t+1}} \right) (1 + MD_{S,t+1}) q_{S,t+1} (1 - \delta_S) \right],$$

$$MD_{K,t} \cdot q_{K,t} = E_t \left[\Lambda_{t+1} \left(\alpha y_{t+1} - \frac{J_{t+1}}{K_{t+1}} \right) \right] \quad + \quad E_t \left[\Lambda_{t+1} (1 - \delta_K)(1 + MD_{K,t+1}) q_{K,t+1} \right] \dots \tag{B.4}$$

$$+ \quad E_t \left[\Lambda_{t+1} \frac{J_{t+1}}{K_{t+1}} (1 + MD_{S,t+1}) q_{S,t+1} \right].$$

Consider a growth option at the end of period t. In period $t+1$ after the quality of the option, θ, is revealed, there are two possibilities. If $\theta \geq \theta_{t+1}^* = G_I\left(I_t, S_t\right)$, then the option is exercised. In this case, the cash flow in period $t+1$ is $-\frac{1}{\theta}$, and the option becomes a generation $t+1$ production unit, which generates cash flow equivalent to ϖ_{t+1} generation-0 production units. In the case $\theta < G_I\left(I_t, S_t\right)$, the option is not exercised and survives to the next period with probability $1 - \delta_S$, in which case it generate the cash flow of a period $t+1$ growth option. Note that the above argument provides a cash flow decomposition of a growth option at period t.

197

By lemma 7, we have

$$
\begin{aligned}
MD_{S,t} \cdot q_{S,t} &= E_t \left[\Lambda_{t,t+1} \cdot \int_{\theta \geq G_I(I_t, S_t)} \left[-\frac{1}{\theta} + (1 + MD_{K,t+1}) \, \varpi_{t+1} q_{K,t+1} \right] f(\theta) \, d\theta \right] \\
&\quad + E_t \left[\Lambda_{t,t+1} \cdot \int_{\theta < G_I(I_t, S_t)} \left[(1 - \delta_S)(1 + MD_{S,t+1}) \, q_{S,t+1} \right] f(\theta) \, d\theta \right].
\end{aligned}
$$

Using the results in Ai (2009), the integrals can be written as functions of the aggregate quantities:

$$
\begin{aligned}
MD_{S,t} \cdot q_{S,t} &= E_t \left[\Lambda_{t,t+1} \cdot \left[-\frac{I_{t+1}}{S_{t+1}} + \frac{G(I_{t+1}, S_{t+1})}{S_{t+1}} (1 + MD_{K,t+1}) \, \varpi_{t+1} q_{K,t+1} \right] \right] \\
&\quad + E_t \left[\Lambda_{t,t+1} \cdot \frac{G(I_{t+1}, S_{t+1})}{S_{t+1}} \left[(1 - \delta_S)(1 + MD_{S,t+1}) \, q_{S,t+1} \right] \right],
\end{aligned}
$$

as needed.

We can decompose the cash flow of a production unit as well. The total output of a period-t production unit is used for three purposes: consumption; tangible investment that produces new-generation production units; and intangible investment that creates new blueprints, which are associated with three difference sources of future cash flows. Equation (B.4) can then be established by applying lemma 7 to the above cash flow decomposition. By solving the system of recursive equations (B.3)–(B.4) we obtain the pair of Macaulay durations, $(MD_{K,t}, MD_{S,t})$.

B.3.2 *Microeconomic foundation of the adjustment cost function H*

Here we show that the law of motion of intangible capital in equation (2.29) arises as the result of a concave production function of new blueprints. Suppose consumption goods, new blueprints, and new investment goods are produced in production

units. Let c, i and j denote the amount of general output used to produce consumption goods, investment goods, and blueprints, respectively. Normalize the price of consumption goods to 1, and denote q_S and q_I the price of blueprints and investment goods, respectively. In this case, the profit maximization problem of a typical production unit can be written as:

$$\max_{c,i,j,n} \left[c + q_I i + q_S h(j) - wn \right]$$

$$c + i + j = (An)^{1-\alpha}.$$

In equilibrium we must have $q_I = 1$. In addition, optimality requires $q_S = 1/h'(j)$, which implies that all production units produce the same amount of blueprints. We continue to use K to denote the total measure of production units. The total amount of resources used to produce blueprints is therefore $J = j \cdot K$, and the total amount of blueprints produced is $K \cdot h(J/K)$ in this economy. After denoting $H(J,K) = h(J/K)K$, the law of motion of intangible capital can be written as in equation (2.29).

The function H, which resembles adjustment costs in neoclassical models, is homogenous of degree one and concave in (J,K). Accordingly, we specify H in the spirit of Jermann (1998) as follows:

$$H(J,K) = \left[\frac{a_1}{1 - 1/\xi} \left(\frac{J}{K} \right)^{1-1/\xi} + a_2 \right] K,$$

where $1/\xi$ determines the concavity of H and the parameters $\{a_1, a_2\}$ satisfy the following two steady state conditions: $H(\overline{J}, \overline{K}) = \overline{J}$ and $H_J(\overline{J}, \overline{K}) = 1$.

199

B.3.3 Further extensions

In this section we retain adjustment costs in the production of intangibles and consider two further extensions of our model. First, we add an endogenous labor supply. Second, we further enrich the model and consider more general specifications of the ϕ_j process that governs the heterogeneity of firms' exposure to aggregate shocks.

We show that our main results are preserved and often enhanced in these more general settings. For the sake of brevity, we focus our discussion only on moments that significantly change across the new extensions.

Endogenous labor supply

In this section, we allow for an endogenous labor supply and explore the ability of the model to account for the joint dynamics of aggregate consumption, investment, and hours worked. We report conventional business cycle statistics generated by the model in Table B.3 and illustrate the response of macroeconomic quantities to productivity shocks in Fig. B.2.

To allow for endogenous labor supply, we adopt a Cobb-Douglas aggregator between consumption goods, C_t, and leisure, $1 - N_t$:

$$u_t = C_t^o (1 - N_t)^{1-o}.$$

We set the parameter o so that the average number of hours worked is equal to one-third of the total number of workable hours. The intratemporal first-order condition for labor is

$$\frac{1-o}{o} \cdot \frac{C_t}{1 - N_t} = (1 - \alpha)\frac{Y_t}{N_t},$$

and the stochastic discount factor becomes

$$\Lambda_{t+1} = \beta \left(\frac{C_{t+1}}{C_t} \right)^{-1} \left(\frac{u_{t+1}}{u_t} \right)^{1-\frac{1}{\psi}} \left[\frac{V_{t+1}}{E_t \left[V_{t+1}^{1-\gamma} \right]^{\frac{1}{1-\gamma}}} \right]^{\frac{1}{\psi}-\gamma} .$$

All other equations that characterize the equilibrium remain unchanged.

We now turn our attention to the implied quantitative results. First, we note that in Fig. B.2 the impulse response of consumption, investments, and returns are qualitatively similar to that obtained with an inelastic labor supply and adjustment costs on intangible capital. As a result, the model's implications for asset prices and macroeconomic quantities discussed so far remain largely unchanged (Table B.3).

Second, in Fig. B.2, short-run shocks (contemporaneous productivity shocks) induce co-movement among consumption, investment, and hours worked, as in standard RBC models. Upon the realization of positive long-run shocks (news about future productivity shocks), however, both investment and labor drop while consumption increases. The negative response of labor with respect to news is due to the income effect: good news about future productivity does not raise the current-period marginal product of labor but does increase the wealth of the agent. As a result, the agent works less, consumes more, and lowers investment. This feature of our model is consistent with the empirical evidence reported in Barsky and Sims (2011) and Kurmann and Otrok (2010) and enables us to produce a low contemporaneous correlation between consumption and labor growth. In this extension, $corr(\Delta C_t, \Delta n_t)$ is slightly lower than in the data, but it increases in the two model specifications to be discussed in the next subsection.

The reduction in the labor supply in response to news shocks lowers the marginal

201

product of physical capital and enhances the decline in investment observed in the benchmark model. Consequently, intangible capital provides even more insurance against news shocks. This explains why the model with an elastic labor supply generates a slightly higher value premium, as reported in Table B.3.

Overall, the inclusion of an endogenous labor supply preserves the success of previous versions of our model and generates plausible cyclical dynamics of hours worked, similar to the standard RBC model.

More general productivity processes

We consider an additional extension that allows for a more flexible specification of the heterogeneity in the productivity processes of production units. As we show in Section 3 of the main article and Appendix B, the exposure to aggregate productivity shocks is increasing in capital vintage age. Allowing for a gradual transition from low to high exposure requires more general specifications of the ϕ_j process and renders the aggregation results in equation (2.9) invalid. Intuitively, multiple transition periods introduce heterogeneity and history dependence of the productivity exposures and require us to keep track of the age distribution as a state variable.

To avoid computational complexity, we restrict our attention to the following simple form of the ϕ_j function:

$$\phi_j = \begin{cases} 0 & j = 0 \\ \phi_1 & j = 1 \\ 1 & j = 2, ... \end{cases},$$

which allows the transition to happen in two periods. Previous versions of the model correspond to the special case with $\phi_1 = 1$.

In this case, the social planner's problem can be made recursive by including last-period physical investment as an additional state variable. To this aim, we use X_t to denote the total measure of production units constructed at time $t-1$ and continue to use K_t to denote the productivity-adjusted stock of production units at time t. The laws of motion of K_t and X_t can be written as follows:

$$K_{t+1} = (1-\delta_K)K_t + X_t[\pi_{t+1}-1] + \varpi_{t+1}G(I_t, S_t), \qquad (B.5)$$

$$X_{t+1} = (1-\delta_K)\varpi_{t+1}G(I_t, S_t),$$

$$\pi_{t+1} = e^{-\frac{1-\alpha}{\alpha}(1-\phi_1)(x_t+\sigma_a\varepsilon_{a,t+1})} \quad \forall t.$$

The social planner's problem is the same as before except that we replace equation (2.9) with equation (B.5), and the value function, $V(K_t, X_t, S_t, x_t, A_t)$, now contains the additional state variable X_t.

We consider two calibrations of the ϕ_j function. In the first, we set $\phi_1 = 0$ so that new production units have zero exposure to aggregate productivity shocks for two periods. In the second calibration, we set $\phi_1 = 0.5$ so that the correlation with aggregate productivity shocks of new production units is 0 in its first period, 0.5 in its second period, and 1 from the third period on. We report our results in the last two rows of Table B.3. Multiple transition periods enhance the positive exposure of the return on tangible capital and the negative exposure of return on intangible capital to long-run productivity shocks. Consequently, these extensions allow us to reduce the volatility of long-run productivity shocks relative to previous calibrations, while still maintaining high equity and value premiums. As a result, the negative correlation between consumption and hours worked induced by long-run shocks is dampened, and our model produces a strong co-movement of consumption and labor,

203

closely matching this moment in the data. All other quantitative implications of the model remain largely unchanged.

B.3.4 Linking the aggregator G to microeconomic evidence

In Section 2.5.1 of the paper, we calibrate the CES aggregator, $G(I, S)$, to match macroeconomic moments. Here we describe a procedure that links the functional form of $G(I, S)$ to microeconomic evidence on the market-to-book ratio of new Initial Public Offering (IPO) firms.

Using the results in Ai (2009), the CES aggregator, $G(I, S)$ implies that $\theta_{i,t}$ is an i.i.d. draw from the following distribution:

$$P\left(\theta_{i,t} \leq \theta\right) = \int_0^\theta \frac{\nu^{-\xi} x^{\xi-1}}{\left[\nu^{-\xi} x^{\xi-1} - 1\right]^{1+\frac{1}{1-1/\xi}}} dx, \quad \theta \in [0, +\infty). \tag{B.6}$$

In our model, the time-t market value of a newly created production unit is $\varpi_t p_{K,t}$. Its book value is the value of investment goods used to implement the blueprint in the last period, $\frac{1}{\theta_{i,t-1}}$. Therefore the market-to-book ratio of a new production unit at time t is $\theta_{i,t-1} \varpi_t p_{K,t}$. Note that not all blueprints are implemented. By Proposition 1, a blueprint i is implemented in period $t-1$ if and only if $\theta_{i,t-1} \geq \theta^*_{t-1} = G_I(I_{t-1}, S_{t-1})$.

The above argument links the truncated density f to the distribution of market-to-book ratios of newly exercised options across firms at a given time. The market-to-book ratios of the firms in our COMPUSTAT-CRSP data set, in contrast, reflect the market-to-book ratio of options exercised by the same firm at different points in times. For this reason, we consider the market-to-book ratio of new IPO firms a better proxy for that of newly exercised options. We think of implementation of blueprints as initial public offering, and we compare the cross-section distribution of

the market-to-book ratio of newly exercised options in our simulation with that of the new IPO firms in the SDC Platinum data set.

We now provide more details about our simulation procedure. In our simulation, we allow the productivity of production units to have an idiosyncratic component ε_t^i:

$$A_{i,t}^t = A_t^t \cdot \varepsilon_t^i.$$

We set $E\left[\varepsilon_t^i\right] = 1$ and choose $Var\left[\varepsilon_t^i\right]$ to match the cross-sectional dispersion of productivity in our COMPUSTAT data. In this case, all aggregation results in our model remain unchanged. The only difference is that the market-to-book ratio of firms i becomes $\varepsilon_t^i \theta_{i,t-1} \varpi_t p_{K,t}$.

We simulate the time series of macroeconomic quantities from our model. Note that S_t measures the amount of blueprints in period t; therefore in each period we sample from the distribution (B.6) a number of independent draws of $\theta_{i,t}$ proportional to S_t. We compute the market-to-book ratio for all new production units which are constructed from the implemented blueprints. This procedure gives a panel of market-to-book ratios of newly established production units which can be used to plot the empirical density. In Fig. B.1, we compare this density (denoted by circles) to the observed empirical distribution (denoted by dots) of the market-to-book ratio of new IPO firms in the SDC Platinum data set. Our sample ranges from 1970 to 2009 and includes 44,922 firms. Fig. B.1 suggests that our choice of the $G\left(I,S\right)$ function conforms well with the microeconomic evidence on the cross-sectional distribution of market-to-book ratio of new IPO firms.

Table B.1: Exposure to Aggregate Risk by Firm Age

Regr.	Age	Sample	E.M.	ξ_3	ξ_4	Obs.	Firms
(1)	F	All	FE	0.011***	-0.003	70899	7333
(2)	F	$\Delta a_t > 0$	FE	0.013***	-0.004	59385	7224
(3)	K-5	All	FE	0.426**	-0.120***	82791	8012
(4)	K-5	$\Delta a_t > 0$	FE	0.880**	-0.133***	63717	7828
(5)	K-5	All	ECM	0.494**	-0.146***	82787	8010
(6)	K-5	$\Delta a_t > 0$	ECM	0.862**	-0.159***	63713	7826
(7)	K-15	All	FE	0.383***	-0.014***	32127	2928
(8)	K-15	$\Delta a_t > 0$	FE	0.511***	-0.016***	25955	2876
(9)	K-15	All	ECM	0.425***	-0.019***	32130	2929
(10)	K-15	$\Delta a_t > 0$	ECM	0.510***	-0.022***	25958	2877

This table reports firms' risk exposure by age. All estimates are based on the following second-stage regression: $\Delta \ln a_{i,j,t} = \xi_{0i} + \xi_1 \Delta \ln A_t + \xi_2 AGE_{i,j,t} + \xi_3 AGE_{i,j,t} \cdot \Delta \ln A_t + \xi_4 B/M_{i,j,t} + \varepsilon_{ijt}$. In col. "Age", F, K-5 and K-15 denote firm age, capital age with T=5, and capital age with T=15, respectively. All regressions denoted by an even number are computed on a sub-sample including only years of positive productivity growth to control for firms exit. "E.M." stands for Estimation Method used in the first stage to estimate $\Delta a_{i,j,t}$. We use FE to denote our Fix Effect procedure described in (B.4) and ECM for the dynamic Error Component Model of Blundell and Bond (2000) described in (B.2). We use *, **, and *** to indicate p-values smaller than 0.10, 0.05, and 0.01, respectively. Standard errors available upon request.

Table B.2: Calibration for Model Extensions

Extension		1	2	3	4
Preference parameters					
Discount factor	β	0.97	0.98	0.98	0.98
Risk aversion	γ	10	29	29	29
Labor adjusted Risk aversion		10	10	10	10
Intertemporal elasticity of substitution	ψ	2.0	2.0	2.0	2.0
Production function/Aggregator parameters					
Capital share	α	0.3	0.3	0.3	0.3
Depreciation rate of physical capital	δ_K	11%	11%	11%	11%
Depreciation rate of intangible capital	δ_S	11%	11%	11%	11%
Weight on physical investment	ν	0.88	0.93	0.85	0.85
Elasticity of substitution	η	2.50	3.8	1.60	1.60
Adjustment Costs	ξ	5	5	5	5
TFP parameters					
Average growth rate	μ	2.0%	2.0%	2.0%	2.0%
Volatility of short-run risk	σ_a	5.08%	4.40%	5.38%	4.98%
Volatility of long-run risk	σ_x	0.86%	0.75%	0.64	0.75
Autocorrelation of expected growth	ρ	0.925	0.925	0.925	0.925
Time-0 Risk exposure of new investment	ϕ_0	0	0	0	0
Time-1 Risk exposure of new investment	ϕ_1	1	1	0	0.5

This table reports the parameter values used for our calibrations referring to the model extensions studied in section 5 of the main article and Appendix C. All models are calibrated at an annual frequency. The labor adjusted risk aversion is computed as γ/o (see Swanson (2012)). Extension 1 features adjustment costs on intangible investment. In Extension 2 we also add endogenous labor. In Extension 3 and 4 we retain adjustment costs and endogenous labor and set ϕ_1 to 0 and 0.5, respectively, to study 2-period transitions of productivity exposure.

Table B.3: Key Moments across Several Model Extensions

	$\sigma_{\Delta C}$	$\sigma_{\Delta I}/\sigma_{\Delta C}$	$\sigma_{\Delta J}/\sigma_{\Delta I}$	$\sigma_{\Delta n}$	$\rho_{\Delta C, \Delta n}$	$E[r_M^{L,ex}]$	$E[r_K^L - r_S^L]$	$\alpha_K - \alpha_S$
Data	02.53	05.29	00.50	02.07	00.28	05.71	04.32	04.01
	(00.56)	(00.50)	(00.07)	(00.21)	(00.07)	(02.25)	(01.39)	(01.77)
Bench.	02.60	05.40	02.50	–	–	05.20	04.20	00.40
Ext. 1	02.53	06.40	00.40	–	–	04.55	04.16	04.91
Ext. 2	02.56	06.00	00.60	01.92	00.08	04.00	04.76	05.98
Ext. 3	02.59	06.23	00.50	01.84	00.28	05.41	05.68	07.70
Ext. 4	02.66	05.40	00.58	01.70	00.19	04.45	05.92	08.00

All figures are multiplied by 100, except contemporaneous correlations (denoted by ρ). Empirical moments are computed using US annual data in log units. Numbers in parentheses are GMM Newey-West adjusted standard errors. $E\left[r_K^L - r_S^L\right]$ and $E\left[r_M^{L,ex}\right]$ measure the levered spread between tangible and intangible capital returns, and the market premium, respectively. The leverage coefficient is 3 (Feijo and Jorgensen (2010)). All the parameters are calibrated as in Table B.2 in Appendix C. The difference in the intercept of the CAPM regression for tangible and intangible returns is denoted by $\alpha_K - \alpha_S$. The entries for the models are obtained by repetitions of small-sample simulations. Extension 1 features adjustment costs on intangible investment. In Extension 2 we also add endogenous labor. In Extensions 3 and 4 we retain adjustment costs and endogenous labor and set ϕ_1 to 0 and 0.5, respectively, to study two-period transitions of productivity exposure.
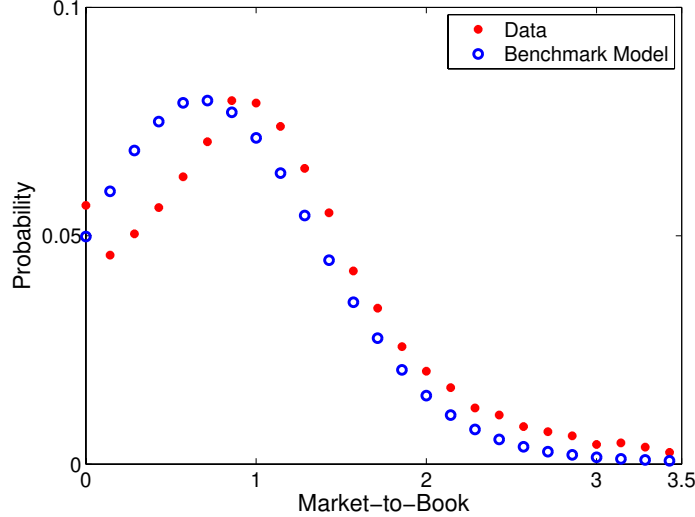

FIGURE B.1: Book-to-Market Distribution for New Production Units.

The circles refer to our Benchmark Model, while the stars refer to IPO.s data from the CDS data set. Our annual sample starts in 1970 and ends in 2009 and includes 22,116 observations.
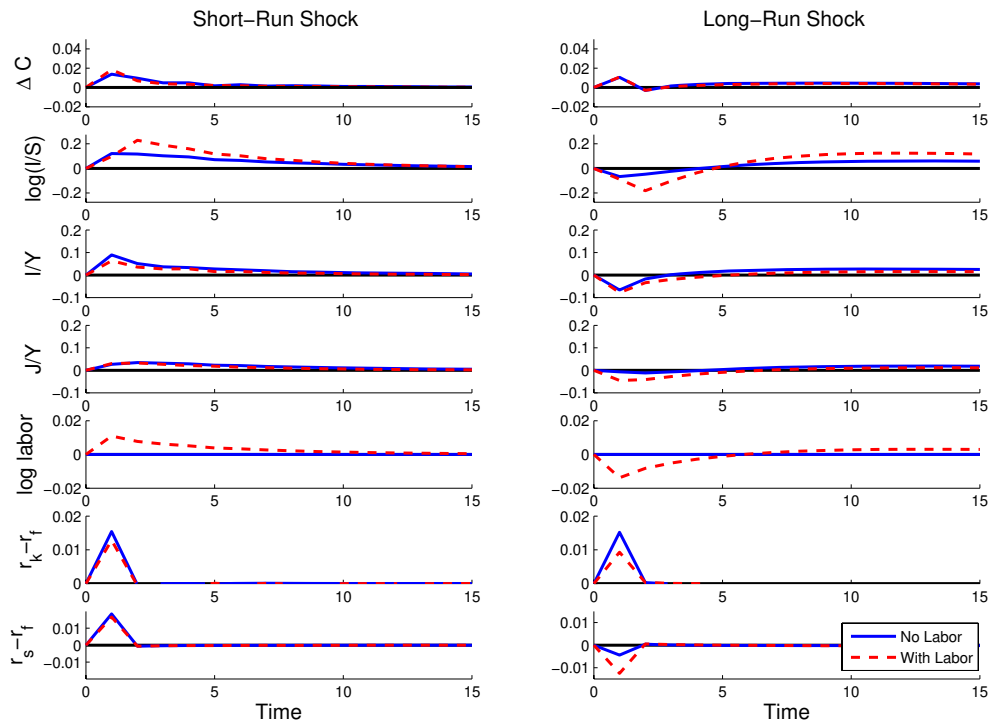
FIGURE B.2: The Role of Endogenous Labor

This figure shows annual log-deviations from the steady state. Returns are not levered. All the parameters are calibrated to the values reported in Table B.2.

Appendix C

Appendix to Chapter 3

C.1   Moment Generation Function of Gamma Distribution

Denote $\Psi(u)$ the moment-generating functions for the Gamma distributed shocks in volatility states:

$$\Psi(u) = E\left(e^{u\omega_{\sigma,t+1}}\right).$$

For my parameterization of Gamma distribution, the expression for the moment-generating functions are given by,

$$\Psi(u) = \left(1 - \widetilde{\theta}u\right)^{-\widetilde{k}}, \text{ for } u < \frac{1}{\theta},$$

in which

$$\widetilde{k} = \left[\frac{\sigma_c^2\left(1-\nu\right)}{\sigma_{c\omega}}\right]^2,$$

$$\widetilde{\theta} = \frac{\sigma_{c\omega}^2}{\sigma_c^2\left(1-\nu_\sigma\right)}.$$

It is important to note that even though the volatility shocks are non-Guassian, the model specification belongs to the exponentially affine class. Indeed, the expectations of the exponential of the state variable is exponentially linear in the current states, which generally facilitates the solution of the model.

C.2   Model Solutions of Within-Regime LRR Model

The dynamics of consumption growth and inflation is characterized as below:

$$
\begin{aligned}
\Delta c_{t+1} &= \mu_c + x_t + \tau_z z_t + \sigma_t \varepsilon_{c,t+1}, \\
\pi_{t+1} &= \mu_\pi + \tau_x x_t + z_t + \varphi_\pi \sigma_t \varepsilon_{\pi,t+1}, \\
x_{t+1} &= \nu_x x_t + \varphi_x \sigma_t \varepsilon_{x,t+1} + \rho_{xz} \varphi_z \sigma_t \varepsilon_{z,t+1}, \\
z_{t+1} &= \nu_z z_t + \rho_{zx} \varphi_x \sigma_t \varepsilon_{x,t+1} + \varphi_z \sigma_t \varepsilon_{z,t+1}, \\
\sigma_{t+1}^2 &= \sigma_c^2 (1 - \nu_\sigma) + \nu_\sigma \sigma_t^2 + \omega_{\sigma,t+1}.
\end{aligned}
$$

The short-run consumption and inflation shocks $\varepsilon_{c,t+1}$ and $\varepsilon_{\pi,t+1}$, the long-run consumption and inflation shocks $\varepsilon_{x,t+1}$ and $\varepsilon_{z,t+1}$ are standard Normal. To ensure the positivity of volatility process, I assume that the volatility shock $\omega_{\sigma,t+1}$ follows demeaned Gamma distribution, i.e. $\omega_{\sigma,t+1} = \widetilde{\omega}_{\sigma,t+1} - E\left(\widetilde{\omega}_{\sigma,t+1}\right)$. The Gamma distribution of $\widetilde{\omega}_{\sigma,t+1}$ is characterized by two parameters, so I specify the mean and volatility of the volatility shocks as

$$
\begin{aligned}
E\left(\widetilde{\omega}_{\sigma,t+1}\right) &= \sigma_c^2 \left(1 - \nu_\sigma\right), \\
Var\left(\widetilde{\omega}_{\sigma,t+1}\right) &= \sigma_{c\omega}^2.
\end{aligned}
$$

Using the Euler equation for the consumption asset, I obtain that the equilibrium

log wealth-to-consumption ratio v_t is linear in the states of the economy:

$$v_t = A_0 + A_x x_t + A_z z_t + A_\sigma \sigma_t^2.$$

Using the Euler equation (3.3) and the assumed dynamics of consumption growth and inflation, I derive the solutions coefficients A_x, A_z and A_σ :

$$A_x = \frac{1 - \frac{1}{\psi}}{1 - \kappa_1 \nu_x},$$

$$A_z = \frac{\left(1 - \frac{1}{\psi}\right) \tau_z}{1 - \kappa_1 \nu_z},$$

$$A_\sigma = \frac{\theta \left[\left(1 - \frac{1}{\psi}\right)^2 + \kappa_1^2 \left\{ (A_x + A_z \rho_{zx})^2 \varphi_x^2 + (A_x \rho_{xz} + A_z)^2 \varphi_z^2 \right\} \right]}{2(1 - \kappa_1 \nu_\sigma)},$$

$$A_0 = \frac{\theta \log \delta + \theta \left(1 - \frac{1}{\psi}\right) \mu_c + \theta \kappa_0 + \log \Psi \left(\theta \kappa_1 A_\sigma\right)}{\theta(1 - \kappa_1)}.$$

Using the equilibrium solution to the wealth-to-consumption ratio, I can write down the expression for the real discount factor in the following way:

$$
\begin{aligned}
m_{t+1} = & \; m_0 + m_x x_t + m_z z_t + m_\sigma \sigma_t^2 \\
& - \lambda_c \sigma_t \varepsilon_{c,t+1} - \lambda_x \varphi_x \sigma_t \varepsilon_{c,t+1} - \lambda_z \varphi_z \sigma_t \varepsilon_{z,t+1} - \lambda_\omega \omega_{\sigma,t+1}.
\end{aligned}
$$

The solution to the discount factor loadings are given by

$$m_0 = \theta \log \delta - \gamma \mu_c + (\theta - 1) \kappa_0 + (\theta - 1) A_0 (\kappa_1 - 1) + (\theta - 1) \kappa_1 A_\sigma \sigma_c^2 (1 - \nu_\sigma),$$

$$m_x = -\frac{1}{\psi},$$

$$m_z = -\frac{\tau_z}{\psi},$$

$$m_\sigma = (\theta - 1) A_\sigma (\kappa_1 \nu_\sigma - 1).$$

The market prices of systematic risks can be expressed in terms of underlying preferences and parameters that govern the evolution of consumption growth and inflation:

$$\lambda_c = \gamma,$$

$$\lambda_x = (1 - \theta)\kappa_1 (A_z \rho_{zx} + A_x),$$

$$\lambda_z = (1 - \theta)\kappa_1 (A_x \rho_{xz} + A_z),$$

$$\lambda_\omega = (1 - \theta)\kappa_1 A_\sigma.$$

Denote $q_{t,n}$ and $q_{t,n}^\$$ the equilibrium solution to the real and nominal n-period bond prices, respectively. Using the Euler equation, the equilibrium solutions to the real bond prices are affine in state variables:

$$q_{t,n} = -B_{0,n} - B_{x,n} x_t - B_{z,n} z_t - B_{\sigma,n} \sigma_t^2.$$

where the loadings satisfy the recursions

$$B_{0,n} = B_{0,n-1} - m_0 - \log \Psi \left(-\lambda_\omega - B_{\sigma,n} \right) - \lambda_\omega \sigma_c^2 \left(1 - \nu \right),$$

$$B_{x,n} = B_{x,n} \nu_x - m_x,$$

$$B_{z,n} = B_{z,n} \nu_z - m_z,$$

$$B_{\sigma,n} = B_{\sigma,n} \nu_\sigma - m_\sigma - \frac{1}{2} \left(\lambda_x + B_{x,n-1} + B_{z,n-1} \rho_{zx} \right)^2 \varphi_x^2$$

$$- \frac{1}{2} \left(\lambda_z + B_{x,n-1} \rho_{xz} + B_{z,n-1} \right)^2 \varphi_z^2 - \frac{1}{2} \lambda_c^2.$$

Define the holding period return of bond as $rb_{t+1}^{(n)} = q_{t+1}^{(n-1)} - q_t^{(n)}$, thus I get

$$rb_{t+1}^{(n)} = G_{0,n} + G_{x,n} x_t + G_{z,n} z_t + G_{\sigma,n} \sigma_t^2$$

$$+ \beta_{x,n}^b \varphi_x \sigma_t \varepsilon_{c,t+1} + \beta_{z,n}^b \varphi_z \sigma_t \varepsilon_{z,t+1} + \beta_{\sigma,n}^b \omega_{\sigma,t+1},$$

in which

$$G_{0,n} = B_{0,n} - B_{0,n-1} - B_{\sigma,n-1} \sigma_c^2 \left(1 - \nu \right),$$

$$G_{x,n} = B_{x,n} - B_{x,n-1} \nu_x,$$

$$G_{z,n} = B_{z,n} - B_{z,n-1} \nu_z,$$

$$G_{\sigma,n} = B_{\sigma,n} - B_{\sigma,n-1} \nu_\sigma,$$

and beta's are

$$\beta_{x,n}^b = - \left(B_{x,n-1} + B_{z,n-1} \rho_{zx} \right),$$

$$\beta_{z,n}^b = - \left(B_{x,n-1} \rho_{xz} + B_{z,n-1} \right),$$

$$\beta_{\sigma,n}^b = -B_{\sigma,n-1}.$$

Risk premium is

$$rp_c = -Cov\left(m_{t+1}, rb_{t+1}^{(n)}|I_t\right),$$

$$= -\left[(B_{x,n-1} + B_{z,n-1}\rho_{zx})\lambda_x\varphi_x^2 + (B_{x,n-1}\rho_{xz} + B_{z,n-1})\lambda_z\varphi_z^2\right]\sigma_t^2 - B_{\sigma,n-1}\lambda_\sigma\sigma_{c\omega}^2.$$

The discounts factor used to price nominal payoff is given by

$$m_{t+1}^{\$} = m_{t+1} - \pi_{t+1}.$$

Similarly, using the equilibrium solution to the nominal discount factor and the Euler equation, the nominal bond prices are affine in state variables:

$$q_{t,n}^{\$} = -B_{0,n}^{\$} - B_{x,n}^{\$}x_t - B_{z,n}^{\$}z_t - B_{\sigma,n}^{\$}\sigma_t^2,$$

where the nominal bond loadings satisfy the recursions:

$$B_{0,n}^{\$} = B_{0,n-1}^{\$} - m_0 - \log\Psi\left(-\lambda_\omega - B_{\sigma,n}\right),$$

$$B_{x,n}^{\$} = B_{x,n}^{\$}\rho_x - m_x,$$

$$B_{z,n}^{\$} = B_{z,n}^{\$}\rho_z - m_z,$$

$$B_{\sigma,n}^{\$} = B_{\sigma,n}^{\$}\nu - m_\sigma - \frac{1}{2}\left(\lambda_x + B_{x,n-1}\varphi_x + B_{z,n-1}\rho_{xz}\right)^2$$

$$-\frac{1}{2}\left(\lambda_z + B_{x,n-1}\rho_{zx} + B_{z,n-1}\varphi_x\right)^2 - \frac{1}{2}\lambda_c^2.$$

C.3   Solutions to LRR Model with Regime Switching

C.3.1   Price-Consumption Ratio

Given $s_t = j$ and $s_{t+1} = k$,

$$m_{t+1} + r_{c,t+1} = B_0(j,k) + B_1(j,k)'Y_t + B_2(k)'\left(G_t(j)\varepsilon_{t+1} + \omega_{t+1}^j\right),$$

in which

$$B_0(j,k) = \theta \log \delta + \theta \left(1 - \frac{1}{\varphi}\right) e_1' \mu(j) + \theta k_0(k) + \theta \kappa_1(k) \left[A_0(k) + A_1(k)' \mu(j)\right] - \theta A_0(j),$$

$$B_1(j,k) = \theta \left[\left(1 - \frac{1}{\varphi}\right) F(j)' e_1 + \kappa_1(k) F(j)' A_1(k) - A_1(j)\right],$$

$$B_2(k) = \theta \left[\left(1 - \frac{1}{\varphi}\right) e_1 + \kappa_1(k) A_1(k)\right].$$

Conjecture log price-to-consumption ratio $v_t(j)$ is linear in the states of the economy:

$$v_t(j) = A_0(j) + A_1(j)' Y_t, \tag{C.1}$$

in which $A_0(j)$ and $A_1(j)$ are jointly determined by the following equation systems:

$$1 = E_t \left[\exp(m_{t+1} + r_{c,t+1})\right],$$

$$= \sum_{k=1}^{S} \pi^{jk} E \left[\exp\left\{B_0(j,k) + B_1(j,k)' Y_t + B_2(k)' \left(G_t(j) \varepsilon_{t+1} + \omega_{t+1}^j\right)\right\} | I_t, s_{t+1}\right],$$

$$= \sum_{k=1}^{S} \pi^{jk} \left[\exp\left(\begin{array}{c} B_0(j,k) + \log \Psi \left[B_2(k)' e_5\right] - B_2(k)' e_5 e_5' \mu(j) \\ + \left\{B_1(j,k) + \frac{1}{2} B_2(k)' H_\sigma(j)' B_2(k)\right\}' Y_t \end{array}\right)\right].$$

for $j = 1, 2, ... S$.

Loglinearize the above equation $(C.1)$ around the unconditional mean $\mu(j)$ as:

$$0 = g(j) + g_1(j)' Y_t, \tag{C.2}$$

216

where

$$g(j) = \log\left\{\sum_{k=1}^{S} \pi^{jk}\left[\exp\left(\begin{array}{c} B_0(j,k) + \log \Psi\left[B_2(k)'e_5\right] - B_2(k)'e_5 e_5'\mu(j) \\ + \left\{B_1(j,k) + \frac{1}{2}B_2(k)'H_\sigma(j)'B_2(k)\right\}'\mu(j) \end{array}\right)\right]\right\},$$

$$g_1(j) = \sum_{k=1}^{S} \pi^{jk}\left\{\exp\left(\begin{array}{c} B_0(j,k) + \log \Psi\left[B_2(k)'e_5\right] - B_2(k)'e_5 e_5'\mu(j) \\ + \left\{B_1(j,k) + \frac{1}{2}B_2(k)'H_\sigma(j)'B_2(k)\right\}'\mu(j) \\ \times\left[B_1(j,k) + \frac{1}{2}B_2(k)'H_\sigma(j)B_2(k)e_5\right] \end{array}\right)\right\}.$$

Since the equation $(C.7)$ holds for all j $(j = 1, 2, ...S)$, I should have

$$g(j) = 0,$$

$$g_1(j) = 0_{6\times 1},$$

for $j = 1, 2, ...S$. With $(7 \times S)$ equations, I can determine $(7 \times S)$ unknowns jointly, i.e. $A_0(j)$ and $A_1(j)$ for $j = 1, 2, ...S$.

C.3.2 Discount Factor

The equilibrium discount factor can be written in the following way:

$$m_{t+1}(j,k) = m_0(j,k) + m_1(j,k)'Y_t - \Lambda(k)'\left(G_t(j)\varepsilon_{t+1} + \omega_{t+1}^j\right), \qquad (C.3)$$

in which

$$m_0(j,k) = \theta\log\delta - \gamma e_1'\mu(j) + (\theta - 1)\left[\kappa_0(k) + \kappa_1(k)\left\{A_0(k) + A_1(k)'\mu(j)\right\} - A_0(j)\right],$$

$$m_1(j,k) = -\gamma F(j)'e_1 + (\theta - 1)\left[\kappa_1(k)F(j)'A_1(k) - A_1(j)\right],$$

and the market prices of risks are:

$$\Lambda(k) = \gamma e_1 + (1 - \theta)\kappa_1(k)A_1(k).$$

217

C.3.3 Return to Consumption Claims

From Campbell-Shiller Decomposition,

$$r_{c,t+1}(j,k) = \kappa_{c,0}(k) + \kappa_{c,1}(k) v_{c,t+1}(k) - v_{c,t}(j) + \Delta c_{t+1}(j). \qquad \text{(C.4)}$$

We have

$$r_{c,t+1}(j,k) = J_0(j,k) + J_1'(j,k)x_t + \beta_c(k)' \left(G_t(j) \varepsilon_{t+1} + \omega_{t+1}^j \right), \qquad \text{(C.5)}$$

in which

$$
\begin{aligned}
J_0(j,k) &= e_1' \mu(j) + \kappa_0(k) + \kappa_1(k) \left[A_0(k) + A_1(k)' \mu(j) \right] - A_0(j), \\
J_1(j,k) &= F(j)' e_1 + \kappa_1(k) F(j)' A_1(k) - A_1(j), \\
\beta_c(k) &= e_1 + \kappa_1(k) A_1(k).
\end{aligned}
$$

C.3.4 Return on Equity

Given $s_t = j$ and $s_t = k$,

$$m_{t+1} + r_{m,t+1} = B_{0,m}(j,k) + B_{1,m}(j,k)'Y_t + B_{2,m}(k)' \left(G_t(j) \varepsilon_{t+1} + \omega_{t+1}^j \right),$$

in which

$$
\begin{aligned}
B_{0,m}(j,k) &= m_0(j,k) + \kappa_{0,m}(k) + \kappa_{1,m}(k) \left[A_{0,m}(k) + A_{1,m}(k)' \mu(j) \right] - A_{0,m}(j) + e_6' \mu(s_t), \\
B_{1,m}(j,k) &= F(j)' e_6 + m_1(j,k) + \kappa_{1,m}(k) F(j)' A_{1,m}(k) - A_{1,m}(j), \\
B_{2,m}(k) &= e_6 + k_{1,m}(k) A_{1,m}(k) - \Lambda(k).
\end{aligned}
$$

Conjecture log price-to-dividend ratio $v_{m,t}(j)$ is linear in the states of the economy:

$$v_{m,t}(j) = A_{0,m}(j) + A_{1,m}(j)'Y_t, \qquad \text{(C.6)}$$

in which $A_{0,m}(j)$ and $A_{1,m}(j)$ are jointly determined by the following equation systems:

$$1 = E_t \left[\exp(m_{t+1} + r_{m,t+1}) \right],$$

$$= \sum_{k=1}^{S} \pi^{jk} E \left[\exp \left\{ B_{0,m}(j,k) + B_{1,m}(j,k)'Y_t + B_{2,m}(k)' \left(G_t(j) \varepsilon_{t+1} + \omega_{t+1}^j \right) \right\} | I_t, s_{t+1} \right],$$

$$= \sum_{k=1}^{S} \pi^{jk} \left[\exp \left(\begin{array}{c} B_{0,m}(j,k) + \log \Psi \left[B_{2,m}(k)'e_5 \right] - B_{2,m}(k)'e_5 e_5'\mu(j) \\ + \left\{ B_{1,m}(j,k) + \frac{1}{2}B_{2,m}(k)'H_\sigma(j) B_{2,m}(k) \right\}' Y_t \end{array} \right) \right],$$

for $j = 1, 2, ...S$.

Loglinearize the above equation $(C.1)$ around the unconditional mean $\mu(j)$ as:

$$0 = g_{0,m}(j) + g_{1,m}(j)'Y_t. \tag{C.7}$$

where

$$g_{0,m}(j) = \log \left\{ \sum_{k=1}^{S} \pi^{jk} \left[\exp \left(\begin{array}{c} B_{0,m}(j,k) + \log \Psi \left[B_{2,m}(k)'e_5 \right] - B_{2,m}(k)'e_5 e_5'\mu(j) \\ + \left\{ B_{1,m}(j,k) + \frac{1}{2}B_{2,m}(k)'H_\sigma(j)' B_{2,m}(k) \right\}' \mu(j) \end{array} \right) \right] \right\},$$

$$g_{1,m}(j) = \sum_{k=1}^{S} \pi^{jk} \left\{ \exp \left(\begin{array}{c} B_{0,m}(j,k) + \log \Psi \left[B_{2,m}(k)'e_5 \right] - B_{2,m}(k)'e_5 e_5'\mu(j) \\ + \left\{ B_{1,m}(j,k) + \frac{1}{2}B_{2,m}(k)'H_\sigma(j) B_{2,m}(k) \right\}' \mu(j) \\ \times \left[B_{1,m}(j,k) + \frac{1}{2}B_{2,m}(k)'H_\sigma(j) B_{2,m}(k)e_5 \right] \end{array} \right) \right\}.$$

Since the equation $(C.7)$ holds for all j $(j = 1, 2, ...S)$, we should have

$$g_{0,m}(j) = 0,$$

$$g_{1,m}(j) = 0_{6\times 1},$$

for $j = 1, 2, ...S$. With $(7 \times S)$ equations, we can determine $(7 \times S)$ unknowns jointly, i.e. $A_{0,m}(j)$ and $A_{1,m}(j)$ for $j = 1, 2, ...S$.

From Campbell-Shiller Decomposition,

$$r_{m,t+1} = \kappa_{0,m} + \kappa_{1,m}v_{m,t+1} - v_{m,t} + \Delta d_{t+1},$$

we have

$$r_{m,t+1}(j,k) = J_{0,m}(j,k) + J_{1,m}(j,k)'Y_t + \beta_d(k)'\left(G_t(j)\varepsilon_{t+1} + \omega_{t+1}^j\right), \qquad (C.8)$$

in which

$$
\begin{aligned}
J_{0,m}(j,k) &= e_6'\mu(j) + \kappa_{0,m}(k) + \kappa_{1,m}(k)\left[A_{0,m}(k) + A_{1,m}(k)'\mu(j)\right] - A_{0,m}(j), \\
J_{1,m}(j,k) &= F(j)'e_6 + \kappa_{1,m}(k)F(j)'A_{1,m}(k) - A_{1,m}(j), \\
\beta_d(k) &= e_6 + \kappa_{1,m}(k)A_{1,m}(k).
\end{aligned}
$$

C.3.5 Real Bond Prices

The log prices at period t of real discount bonds with n periods to maturity, $q_{t,n}$, satisfies the Euler condition

$$\exp(q_{t,n}) = E_t \exp(m_{t+1} + q_{t+1,n-1}),$$

for $q_{t+n,0} = 0$.

The log bond price can be derived as

$$q_{t,n}(j) = B_{0,n}(j) + B_{1,n}(j)'Y_t, \qquad (C.9)$$

an affine structure of the states of the economy. The coefficients $B_{0,n}$ and $B_{1,n}$ are state s_t dependent, and follow the recursive relations:

$$
B_{0,n}(j) = \log\left\{\sum_{k=1}^{S}\pi^{jk}\left(\begin{array}{c} D_0(j,k) + \log\Psi\left[D_2(k)'e_5\right] - D_2(k)'e_5 e_5'\mu(j) \\ + \left\{D_1(j,k) + \frac{1}{2}D_2(k)'H_\sigma(j)D_2(k)\right\}'\mu(j) \end{array}\right)\right\}, \qquad (C.10)
$$

$$
B_{1,n}(j) = \frac{1}{\exp(B_{0,n}(j))}\left\{\sum_{k=1}^{S}\pi^{jk}\left\{\exp\left(\begin{array}{c} D_0(j,k) + \log\Psi\left[D_2(k)'e_5\right] - D_2(k)'e_5 e_5'\mu(j) \\ + \left\{D_1(j,k) + \frac{1}{2}D_2(k)'H_\sigma(j)D_2(k)\right\}'\mu(j) \\ \times\left[D_1(j,k) + \frac{1}{2}D_2(k)'H_\sigma(j)D_2(k)e_5\right] \end{array}\right)\right\}\right\},
$$

220

where

$$D_0(j,k) = m_0(j,k) - B_{0,n-1}(k) - B_{1,n-1}(k)'\mu(j),$$

$$D_1(j,k) = m_1(j,k) - F(j)' B_{1,n-1}(k),$$

$$D_2(k) = -\left[\Lambda(k) + B_{0,n-1}(k)\right].$$

Define the holding period return of bond as $rb_{t+1}^{(n)} = q_{t+1}^{(n-1)} - q_t^{(n)}$, thus

$$rb_{t+1}^{(n)}(j,k) = G_{0,n}(j,k) + G_{1,n}(j,k)'Y_t + \beta_{b,n}(j,k)' \left(G_t(j)\varepsilon_{t+1} + \omega_{t+1}^j\right),$$

where

$$G_{0,n}(j,k) = B_{0,n}(j) - B_{0,n-1}(k) - B_{1,n-1}(k)'\mu(j),$$

$$G_{1,n}(j,k) = B_{1,n}(j) - F(j)'B_{1,n-1}(ks_{t+1}),$$

$$\beta_{b,n}(j,k) = -B_{1,n-1}(k).$$

C.3.6 Nominal Discount Factor

The equilibrium stochastic discount factor can be written in the following way:

$$m_{t+1}^{\$} = m_{t+1} - \pi_{t+1} = m_0^{\$}(j,k) + m_1^{\$}(j,k)'x_t - \Lambda^{\$}(j,k)' \left(G_t(j)\varepsilon_{t+1} + \omega_{t+1}^j\right), \quad \text{(C.11)}$$

in which

$$m_0^{\$}(j,k) = m_0(j,k) - e_2'\mu(j),$$

$$m_1^{\$}(j,k) = m_1(j,k) - F(j)'e_2,$$

$$\Lambda^{\$}(j,k) = \Lambda(k) + e_2.$$

The nominal bond pricing and nominal bond risk premium works exactly the same as the real bond case.

Bibliography

Abel, A. February 1999. Risk premia and term premia in general equilibrium. *Journal of Monetary Economics* 43(1):3–33.

Adrian, T., and N. boyarchenko. Intermediary leverage cycles and financial stability. New York Fed Working paper 2012.

Adrian, T., and H. S. Shin. July 2010. Liquidity and leverage. *Journal of Financial Intermediation* 19(3):418–437.

Adrian, T., E. Moench, and H. S. Shin. Financial intermediation, asset prices and macroeconomic dynamics. Working paper 2011.

Ai, H. 2009. Intangible capital and the value premium. Working paper, Fuqua School of Business.

Ai, H., and D. Kiku. 2012. Growth to value: Option exercise and the cross-section of equity returns. *Journal of Financial Economics*. Advance Access published August 10, 2012, 10.1016/j.jfineco.2012.08.009.

Ai, H., R. Bansal, and K. Li. Financial intermediary, asset prices and the real economy. Working paper November 2012.

Altman, E. I. 2000. Predicting fiancial distress of companies: Revisiting the z−score and zeta models. Working paper, NYU.

Bansal, R., and I. Shaliastovich. A long-run risks explanation of predictability puzzles in bond and currency markets. Working paper. Feburary 2010.

Bansal, R., and A. Yaron. 2004. Risk for the long run: A potential resolution of asset pricing puzzles. *The Journal of Finance* 59(4):1481–1509.

Bansal, R., and H. Zhou. 2002. Term structure of interest rates with regime shifts. *Journal of Finance* pages 1997–2043.

Bansal, R., V. Khatchatrian, and A. Yaron. Interpretable asset markets? *European Economic Review, year = 2005, volume = 49, number = 3, pages = 531-560.*

Bansal, R., R. Dittmar, and C. Lundblad. 2005. Consumption, Dividends, and the Cross Section of Equity Returns. *The Journal of Finance* 60(4):1639–1672.

Bansal, R., D. Kiku, and A. Yaron. 2007a. Risks for the long run: Estimation and inference. Working paper, Wharton School.

Bansal, R., D. Kiku, and A. Yaron. Risks For the Long Run: Estimation and Inference. Working paper, Wharton School. 2007b.

Bansal, R., D. Kiku, and A. Yaron. 2012. An empirical evaluation of the long-run risks model for asset prices. *Critical Finance Review* (1):183–221.

Barndorff-Nielsen, O. E., and N. Shephard. 2001. Non-gaussian ornstein-uhlenbeck-based model and some of their uses in financial economics. *Journal of Royal Statistical Society* pages 167–241.

Barro, R. August 2006. Rare disasters and asset markets in the twentieth century. *Quarterly Journal of Economics* 121(3):823–866.

Barsky, B., and E. Sims. 2011. News shocks and business cycles. *Journal of Monetary Economics* 58:273–89.

Basak, S., and D. Cuoco. 1998. An equilibrium model with restricted risky asset market participation. *Review of Financial Studies* 11(2):309–341.

Beeler, J., and J. Y. Campbell. 2012. The long-run risks model and aggregate asset prices: An empirical assessment. *Critical Finance Review* (1):141–182.

Berk, J., R. C. Green, and V. Naik. 1999. Optimal investment, growth options, and security returns. *Journal of Finance* 54:1553–608.

Blundell, R. W., and S. R. Bond. 2000. GMM estimation with persistent panel data: An application to production functions. *Econometric Reviews* 19:321–340.

Bollerslev, T. 1990. Modelling the coherence in short-run nominal exchange rate: A multivariate generalized arch model. *Review of Economics and Statistics* pages 498–505.

Borovička, J., and L. Hansen. 2011. Examining macroeconomic models through the lens of asset pricing. Working Paper, Chicago University.

Borovička, J., L. Hansen, M. Hendricks, and J. Scheinkman. 2011. Risk-price dynamics. *Journal of Financial Econometrics* 9:3–65.

Brunnemeier, M., and L. Pedersen. 2009. Market liquidity and funding liquidity. *Review of Financial Studies* 22(6):2201–2238.

Brunnemeier, M., and Y. Sannikov. A macroeconomic model with a financial sector. Working paper, Princeton University April 2012.

Campbell, J. 1999. *Asset Prices, Consumption, and the Business Cycle.* North-Holland, Amsterdam.

Campbell, J. 2000. Asset pricing at the millennium. *Journal of Finance* 55:1515–1567.

Campbell, J., and J. Cochrane. April 1999. By force of habit: A consumption-based explanation of aggregate stock market behavior. *Journal of Political Economy* 107: 205–251.

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

Campbell, J. Y., A. Sunderam, and L. M. Viceira. Inflation bets or deflation hedges? the changing risks of nominal bonds. Working paper. July 2010.

Carlson, M., A. Fisher, and R. Giammarino. 2004. Corporate investment and asset price dynamics: Implications for the cross-section of returns. *Journal of Finance* 59: 2577–603.

Carlstrom, C., and T. Fuerst. 1997. Agency costs, net worth, and business fluctuations: A computable general equilibrium analysis. *American Economic Review* 87 (2):893–910.

Cieslak, A., and A. Vedolin. Real-nominal uncertainty and asset prices. Working paper. Feburary 2009.

Cochrane, J. H., and M. Piazzesi. 2005. Bond risk premia. *American Economic Review* pages 138–160.

Cooper, I. 2006. Asset pricing implications of nonconvex adjustment costs and irreversibility of investment. *Journal of Finance* 61:139–70.

Corrado, C., C. Hulten, and D. Sichel. 2006. Intangible capital and economic growth. Working Paper, Federal Reserve Board.

Croce, M. M. 2008. Long-run productivity risk: A new hope for production-based asset pricing. Working paper, Kenan-Flagler Business School, UNC.

Da, Z. 2006. Cash flow, consumption risk, and the cross-section of stock returns. *The Journal of Finance*.

Da, Z., R.-J. Guo, and R. Jagannathan. 2012. CAPM for estimating the cost of equity capital: Interpreting the empirical evidence. *Journal of Financial Economics* 103:204–20.

Danielsson, J., H. S. Shin, and J.-P. Zigrand. Balance sheet capacity and endogenous risk. Working paper 2011.

Dechow, P., R. Sloan, and M. Soliman. 2004. Implied equity duration: A new measure of equity risk. *Review of Accounting Studies* 9:197–228.

Den-Haan, W. J., and A. Marcet. January 1994. Accuracy in simulations. *The Review of Economic Studies* 61(1):3–17.

Eisfeldt, A., and D. Papanikolaou. 2012. Organization capital and the cross section of expected returns. *Journal of Finance*. Forthcoming.

Engle, R. 2002. Dynamic conditional correlation: A simple class of multivariate generalized autoregressive conditional heteroskedasticity modelsl. *Journal of Business and Economic Statistics* 20.

Epstein, L., and S. E. Zin. 1989. Substitution, Risk Aversion, and the Temporal Behavior of Consumption and Asset Returns: A Theoretical Framework. *Econometrica* 57(4):937–969.

Fama, E., and K. R. French. 1992. The cross section of expected stock returns. *The Journal of Finance* XLVII(2):427–465.

Fama, E., and K. R. French. 1995. Size and book-to-market factors in earnings and returns. *The Journal of Finance* 50(1):135–155.

Feijo, L. G., and R. D. Jorgensen. 2010. Can operating leverage be the cause of the value premium? *Financial Management* 39:1127–54.

Gala, V. 2005. Investment and Returns. Working paper, London Business School.

225

Geanakoplos, J. The leverage cycle. Working paper 2012.

Geanakoplos, J., and A. Fostel. 2008. Leverage cycles and the anxious economy. *American Economic Review* 98(4):1211–1244.

Gertler, M., and P. Karadi. 2011. A model of unconventional monetary policy. *Journal of Monetary Economics* 58(1):17–34.

Gertler, M., and P. Karadi. A framework for analyzing large scale asset purchases as a monetary policy tool. Working paper May 2012.

Gertler, M., and N. Kiyotaki. 2010. *Financial Intermediation and Credit Policy in Business Cycle Analysis.* Elsevier, Handbook of Monetary Economics, Chapter 11.

Gertler, M., N. Kiyotaki, and A. Queralto. Financial crises, bank risk exposure and government financial policy. Working paper May 2011.

Gomes, J., L. Kogan, and L. Zhang. 2003. Equilibrium Cross Section of Returns. *Journal of Political Economy* 111(4).

Gourio, F., and L. Rudanko. 2010. Customer capital. Working Paper, Boston University.

Greene, W. H. 2002. *Econometric Analysis.* Prentice Hall.

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

Hand, J., and B. Lev. 2003. *Intangible Assets.* Oxford Management Readers.

Hansen, L., J. Heaton, and N. Li. 2005. *Intangible Risk?* NBER Books.

Hansen, L. P., J. Heaton, and N. Li. 2006. Consumption strikes back. Working paper, University of Chicago.

Hansen, L. P., J. C. Heaton, and N. Li. 2008. Consumption strikes back? measuring long-run risk. *Journal of Political Economy* 116(2).

Hasseltoft, H. The 'fed model' and the changing correlation of stock and bond returns: An equilibrium approach. January 2009.

He, Z., and A. Krishnamurthy. 2012b. Intermediary asset pricing. *American Economic Review (forthcoming).*

Heckman, J. 1979. Sample selection bias as a specification error. *Econometrica* 47: 153–62.

Jermann, U. J. 1998. Asset Pricing in Production Economies. *Journal of Monetary Economics* 41(2):257–275.

Jovanovic, B. 2008. Investment options and the business cycle. *Journal of Economic Theory.*

Jovanovic, B., and P. Rousseau. 2001. Why wait? A century of life before IPO. *American Economic Review, P&P* 91:336–41.

Kaltenbrunner, G., and L. Lochstoer. 2010. Long-run risk through consumption smoothing. *Review of Financial Studies.* Forthcoming.

Kiku, D. 2006. Long run risks and the value premium puzzle. Working paper, Duke University.

Kiyotaki, N., and J. Moore. 1997. Credit cycles. *The Journal of Political Economy* 105(2):211–248.

Kogan, L., and D. Papanikolaou. 2009. Growth opportunities, technology shocks and asset prices. Working paper, MIT Sloan School.

Kogan, L., and D. Papanikolaou. 2010. Growth opportunities and technology shocks. *American Economic Review, P&P* 100:532:36.

Kogan, L., and D. Papanikolaou. 2011. Economic activity of firms and asset prices. Working paper, MIT Sloan School.

Kreps, D. M., and E. L. Porteus. 1978. Temporal Resolution of Uncertainty and Dynamic Choice Theory. *Econometrica* 46(1):185–200.

Kurmann, A., and C. Otrok. 2010. News shocks and the slope of the term structure of interest rates. Working paper, University of Missouri-Columbia.

Kydland, F., and E. Prescott. 1982. Time to build and aggregate fluctuations. *Econometrica* 50:1345–70.

Li, E., and L. Liu. 2010. Intangible assets and cross sectional stock returns: Evidence from structural estimation. Working paper, Mitchgan University Ross School.

Li, N. 2009. Intangible capital and stock prices. Working Paper, Chicago University.

Lin, X. 2010. Endogenous technological progress and the cross section of stock returns. *Journal of Financial Economics* 103:411–27.

Lucas, R. E. November 1978. Asset prices in an exchange economy. *Econometrica* 46(6):1429–1445.

Lustig, H., and A. Verdelhan. Business cycle variation in the risk-return trade-off. Working paper. September 2010.

Maggiori, M. Financial intermediation, international risk sharing, and reverse currencies. Working paper November 2012.

Mankiw, G., and S. Zeldes. 1991. The consumption of stockholders and nonstockholders. *Journal of Financial Economics* 29(1):97–112.

McGrattan, E. R., and E. C. Prescott. 2009a. Unmeasured investment and the puzzling u.s. boom in the 1990s. *American Economic Journal: Macroeconomics*. Forthcoming.

McGrattan, E. R., and E. C. Prescott. 2009b. Technology capital and the u.s. current account. *American Economic Review*. Forthcoming.

Mehra, R., and E. Prescott. March 1985. The equity premium: A puzzle. *Journal of Monetary Economics* 15(2):145–161.

Papanikolaou, D. 2011. Investment-specific shocks and asset prices. *Journal of Political Economy*. Forthcoming.

Pastor, L., and P. Veronesi. 2009. Technological revolutions and stock prices. *American Economic Review* 99:1451–83.

Pelletier, D. 2006. Regime switching for dynamic correlations. *Journal of Econometrics* 131:445–473.

Rampini, A., and S. Viswanathan. Financial intermediary capital. Working paper 2012.

Ritter, J., and T. Loughran. 2004. Why has IPO underpricing changed over time? *Financial Management* 33:5–37.

Santos, T., and P. Veronesi. 2010. Habit formation, the cross section of stock returns and the cash flow risk puzzle. *Journal of Financial Economics*.

Swanson, E. 2012. Risk aversion and the labor margin in dynamic equilibium models. *American Economic Review* 102:1663–91.

Tallarini, T. D. 2000. Risk Sensitive Real Business Cycles. *Journal of Monetary Economics* 45(3):507–532.

Viceira, L. M. 2010. Bond risks, bond return volatility, and the term structure of interest rates. *International Journal of Forecasting, forthcoming.*

Vissing-Jorgensen, A. 2002. Limited asset market participation and the elasticity of intertemporal substitution. *Journal of Political Economy* 110(4):825–853.

Weil, P. January 1989. The equity premium puzzle and the riskfree rate puzzle. *Journal of Monetary Economics* 24(3):401–422.

Zhang, L. 2005. The value premium. *Journal of Finance* 60:67–103.

Biography

Kai Li was born in Wuxi, Jiangsu Province, China on January 5, 1982. He earned his B.A. degree in Finance from Shanghai Jiao Tong University in May 2004, and his M.A. degree in Economics from China Center for Economic Research, Peking University in May 2007. After obtaining his Ph.D. degree from Duke University, Kai will become an Assistant Professor of Finance at the Hong Kong University of Science and Technology.