ESSAYS ON ASSET PRICING AND PORTFOLIO CHOICE

Hsin-hung Jerry Tsai

A DISSERTATION

in

Finance

For the Graduate Group in Managerial Science and Applied Economics

Presented to the Faculties of the University of Pennsylvania

in

Partial Fulfillment of the Requirements for the

Degree of Doctor of Philosophy

2013

Supervisor of Dissertation

_____

Jessica A. Wachter, Professor of Finance

Graduate Group Chairperson

_____

Eric Bradlow, Professor of Marketing, Statistics and Education

Dissertation Committee

Nikolai Roussanov, Assistant Professor of Finance

Ivan Shaliastovich, Assistant Professor of Finance

Jessica A. Wachter, Professor of Finance

Amir Yaron, Robert Morris Professor of Banking, Professor of Finance

ABSTRACT

ESSAYS ON ASSET PRICING AND PORTFOLIO CHOICE

Hsin-hung Jerry Tsai

Jessica A. Wachter

The first chapter "Rare Disasters and the Term Structure of Interest Rates" offers an explanation for the properties of the nominal term structure of interest rates and time-varying bond risk premia based on a model with rare consumption disaster risk. In the model, expected inflation follows a mean reverting process but is also subject to possible large (positive) shocks when consumption disasters occur. The possibility of jumps in inflation increases nominal yields and the yield spread, while time-variation in the inflation jump probability drives time-varying bond risk premia. Predictability regressions offer independent evidence for the model's ability to generate realistic implications for both the stock and bond markets.

The second chapter "Rare booms and disasters in a multi-sector endowment economy" studies the cross-section of stock returns. Why do value stocks have higher expected returns than growth stocks, in spite of having lower risk? Why do these stocks exhibit positive abnormal performance while growth stocks exhibit negative abnormal performance? This paper offers a rare-events based explanation, that can also account for facts about the aggregate market. Patterns in time-series predictability offer independent evidence for the model's conclusions.

The third chapter "Dynamic Asset Allocation with Learning" studies an asset allocation problem. It shows that learning about the parameters of the return process induces a large negative hedging demand in an investor who is optimally rebalancing her portfolio, even after she has observed 83 years of market asset data. For example, an investor with a 5-year investment horizon decreases the percentage of wealth she allocates to the stock index by

over 20 percent when she takes learning into account. Furthermore, I show that the initial estimation sample length needs to be at least 500 years in order for the effect of learning to vanish.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF ILLUSTRATIONS

CHAPTER 1 : Rare Disasters and the Term Structure of Interst Rates

1.1. Introduction

Empirical work has documented the failure of the expectations hypothesis. The average nominal term structure of interest rates on government bonds is upward-sloping, and the excess bond returns are predictable by variables such as yield spread. This indicates that bond risk premia are on average positive and vary over time. This paper presents a representative agent asset pricing model in which the aggregate endowment is subject to large negative shocks (disasters). Earlier work has shown that models with time-varying disaster risk can account for the high equity premium, high stock market volatility and aggregate market return predictability observed in the aggregate stock market.[1] In addition to the aggregate market results shown in previous work, my model accurately captures the shape of the nominal yield curve and the time-varying bond risk premia.

This paper provides an explanation for these features of the nominal bonds in a time-varying rare disaster model. In particular, consumption disasters may co-occur with high inflation, implying that nominal bonds are risky because their real values during bad times can be very low. Table 1.1 provides evidence for the co-occurrence of consumption disaster and high inflation. In this paper, a consumption disaster is defined as a consumption decline of more than 10%, and I consider a period as having high inflation if the average annual inflation rate during the period is greater than 10%. In recorded history, 17 of the 53 consumption disasters in OECD countries, and 30 of the 89 consumption disasters among all countries, were accompanied by inflation rates greater than 10%.[2] Furthermore, in 18 of the 30 inflation disasters, inflation rates exceeded the real consumption declines.[3] Figure 1.1

---

[1]For example, Rietz (1988), Longstaff and Piazzesi (2004), and Barro (2006) obtain high equity premium, Gabaix (2012), Gourio (2008), and Wachter (2012) also obtain high volatility and predictability.

[2]One might argue that consumption disasters are accompanied by large deflation. However, only 10 of the OECD disasters, and 17 of all disasters coincide with deflation. Furthermore, none of these disasters had an abnormally large annual deflation rate; for example, the Great Depression had an annual deflation rate of 6.4%.

[3]One of the most extreme examples is the hyperinflation that occurred in Germany after World War I. Between 1922 and 1923, real consumption declines by 12.7%, but the inflation rate in the corresponding

shows that the historical distribution of annual inflation rates has a fat tail. Furthermore, these jumps in inflation rates do not happen all at once, they were gradual processes that lasted a number of years.

Motivated by this evidence, I model the aggregate endowment in the model as subject to two types of disasters. These disasters are modeled as negative jumps in the realized consumption process. When the first type of disaster occurs, aggregate endowment drops, but expected inflation is unaffected. When the second type of disaster occurs, not only does aggregate endowment drop, but expected inflation increases. There were no consumption disasters in the United States in the period following World War II. In the 1970s, however, the U.S. experienced a period of high inflation. To accommodate this possibility in the model, I allow for a third type of jump, one which affects expected inflation but not aggregate consumption growth.

Because government bonds are nominally denominated, they are subject to inflation jump risks. Investors require compensation for bearing these risks. The shape of the nominal yield curve in the model is mostly determined by the inflation jump risk since bonds with longer maturities are more sensitive to these risks. In particular, the yield spread increases in inflation jump risks, thus the model accurately predicts an upward-sloping nominal yield curve. Furthermore, the time-varying nature of disaster probability implies a time-varying bond risk premium.

This paper makes two main contributions to the existing literature. First, it provides a parsimonious model that jointly explains the stock and bond markets. Second, it can account for the time-series behavior of the bond premium and its relation to the equity premium. While the model is only calibrated to match aggregate consumption growth, inflation, and aggregate stock market moments, it generates realistic implications for the nominal term structure. Similar to the findings of Litterman and Scheinkman (1991), the first three principal components explain almost all the variations in nominal yields in

---

period is 3450%.

the model, furthermore, each of these three principal component is highly correlated with one of the three state variables in the model. This model also generates bond premium predictability because bond premium are mainly affected by the time-varying risk of the co-occurrence of a consumption disaster and high inflations. In particular, this model is able to reproduce the findings in Campbell and Shiller (1991) and Cochrane and Piazzesi (2005). Nominal bond excess returns are predictable by the yield spread and a linear combination of forward rates.

Besides the shape of the nominal term structure and the time-series behavior of the bond risk premium, this model can also account for the interaction between the stock and nominal bond markets. Duffee (2012) suggests that while term structure variables can predict the bond premium, they are not good predictors for the equity premium. In particular, I show that the price-dividend ratio predicts excess returns on the aggregate market (Campbell and Shiller (1988)) and that it has some predictive power for excess returns on the bond market. Term structure variables predict excess returns on the nominal bond market (Fama and Bliss (1987)) yet they are less effective at predicting excess returns on the aggregate market. In this model, the prices of risk have a two-factor structure, and the model is thus capable of explaining these results.

Several other papers also provide joint explanations for stock and bond market prices. Gabaix (2012) also considers a model with rare disasters. In that model, rather than time-variation in the disaster probability, it is time-variation in the expected size of an inflation jump that drives the bond premium. Furthermore, I allow fewer degrees of freedom in the calibration so that none of the parameters are chosen to match the yield curve. Wachter (2006), Bekaert et al. (2010) and Buraschi and Jiltsov (2007) consider extensions to the model with external habit formation (Campbell and Cochrane (1999)).Bakshi and Chen (1996) study monetary models in which the money supply directly enters the utility function. The economic mechanisms behind this model differ from those in the papers mentioned here. The shape of the nominal term structure is driven by the time-varying probability of the

co-occurrence of a large consumption decline and high inflations. Furthermore, this paper provides evidence of the interaction between stock and bond markets by studying cross-market predictability. It is likely that the term structure of interest rates and bond premia are affected by multiple mechanisms, and this paper provides another possible way to jointly explain the aggregate market and bond market in a single model.

This paper is also related to a stream of literature that focuses on the term structure of interest rates, but does not address equity prices. Piazzesi and Schneider (2006) focus on the negative effects of surprise inflation on future consumption growth. Bansal and Shaliastovich (2013) build on the Bansal and Yaron (2004) long-run risk framework with stochastic volatility. Similar to Piazzesi and Schneider (2006) and Bansal and Shaliastovich (2013), in this model, when the risk of the co-occurrence of a consumption disaster and high inflations is high, expected consumption growth is low and expected inflation is high. However, high inflations and low consumption growth only co-occur when this type of consumption disasters are realized. Bekaert et al. (2001) evaluate the violation of the expectations hypothesis using a Peso problem explanation. Ehling et al. (2012) study the effect of differences in beliefs about expected inflation when investors have habit-formation preferences.

Finally Dai and Singleton (2002) study three-factor term structure models in the essentially affine class (Duffee (2002)) and show that a statistical model of the stochastic discount factor can resolve the expectations hypothesis puzzle. Many other recent papers also consider the role of macroeconomic variables in the term structure by introducing macroeconomic time series into the stochastic discount factor (Ang and Piazzesi (2003), Ang et al. (2007), Bikbov and Chernov (2010), Duffee (2006), and Rudebusch and Wu (2008)).

The remainder of the paper is organized as follows. Section 2 describes and solves the model. Section 3 discusses the quantitative results of the model. Section 4 concludes.

1.2. Model

1.2.1. Endowment, inflation, and preferences

The economy is populated with a representative agent. Assume that aggregate real consumption solves the following stochastic differential equation:

$$\frac{dC_t}{C_{t^-}} = \mu \, dt + \sigma_C \, dB_{Ct} + \left(e^{Z_{ct}} - 1\right) dN_{ct} + \left(e^{Z_{cq,t}} - 1\right) dN_{cq,t},$$

where B_{Ct} is a standard Brownian motion. Aggregate consumption is subject to two types of large shocks, and the arrival times of these shocks have a Poisson distribution, given by N_{ct} and $N_{cq,t}$. I will discuss the size and intensity of these Poisson jumps after I specify the inflation process.

To model nominal assets, I assume an exogenous process for the price level:

$$\frac{dP_t}{P_{t^-}} = q_t \, dt + \sigma_P \, dB_{Pt}, \tag{1.1}$$

where B_{Pt} is a standard Brownian motion, that is independent of B_{Ct}.

The expected inflation process, q_t, is time-varying. Specifically, it follows

$$dq_t = \kappa_q \left(\bar{q} - q_t\right) dt + \sigma_q \, dB_{qt} - Z_{cq,t} \, dN_{cq,t} - Z_{qt} \, dN_{qt}, \tag{1.2}$$

where B_{qt} is a standard Brownian motion, that is independent of B_{Ct} and B_{Pt}. The expected inflation process is also subject to two types of large shocks, and the arrival time of these shocks follow Poisson distributions, given by $N_{cq,t}$ and N_{qt}.

The magnitude of an N_c–type jump is determined by Z_c, the magnitude of an N_{cq}–type jump is determined by Z_{cq}, and that of an N_q–type jump is determined by Z_q. I will consider all three types of Poisson shocks to be negative, that is $Z_c < 0$, $Z_{cq} < 0$, and $Z_q < 0$; furthermore, these jump sizes are random and have time-invariant distributions ν_c,

ν_{cq}, and ν_q, respectively. In what follows, I use the notation E_{ν_j} to denote expectations taken over the distribution ν_j for $j \in \{c, cq, q\}$. The intensities of these Poisson shocks are time-varying, and each follows a square-root process as in Cox et al. (1985). In what follows, I will assume that inflation spike probability is perfectly correlated with inflation disaster probability.[4] Specifically, for $j \in \{c, cq\}$, the intensity for N_j is denoted by λ_{jt}, and it is given by

$$d\lambda_{jt} = \kappa_{\lambda_j}(\bar{\lambda}_j - \lambda_{jt})\, dt + \sigma_{\lambda_j}\sqrt{\lambda_{jt}}\, dB_{\lambda_j t}.$$

$B_{\lambda_c t}$ and $B_{\lambda_{cq,t}}$ are independent Brownian motions, and each is independent of B_{Ct}, B_{Pt}, and B_{qt}. Furthermore, assume that the Poisson shocks are independent of each other, and of the Brownian motions. Define $\lambda_t = [\lambda_{ct},\ \lambda_{cq,t}]^\top$, $\bar{\lambda} = [\bar{\lambda}_c,\ \bar{\lambda}_{cq}]^\top$, $\kappa_\lambda = [\kappa_{\lambda_c},\ \kappa_{\lambda_{cq}}]^\top$, $B_{\lambda t} = [B_{\lambda_c t},\ B_{\lambda_{cq} t}]^\top$, and $B_t = [B_{Ct},\ B_{Pt},\ B_{qt},\ B_{\lambda t}^\top]^\top$.

In what follows, a *disaster* (or *consumption disaster*) is a Poisson shock that affects realized consumption growth. In particular, I will refer to the N_c–type shock as a *non-inflation disaster* and the N_{cq}–type shock as an *inflation disaster*. The N_q–type shock only affects expected inflation and I refer to it as an *inflation spike*. Furthermore, I will refer to λ_c as the *non-inflation disaster probability* and λ_{cq} as the *inflation disaster probability*. Though the latter also governs the intensity of inflation spikes, the majority of its effects comes from inflation disasters rather than inflation spikes.

Following Duffie and Epstein (1992), I define the utility function V_t for the representative agent using the following recursion:

$$V_t = E_t \int_t^\infty f(C_s, V_s)\, ds, \tag{1.3}$$

---

[4]Inflation spikes in this model attempt to speak to the period of high inflation in the 1970s and early 1980s. During this period, consumption growth was low, and the outlook for future consumption growth was uncertain. Therefore not modeling inflation spike probability as an independent process is realistic. To simplify the model, I assume that the inflation spike probability equals inflation disaster probability.

where

$$f(C_t, V_t) = \beta(1 - \gamma)V_t \left(\log C_t - \frac{1}{1 - \gamma} \log \left((1 - \gamma)V_t \right) \right). \tag{1.4}$$

The above utility function is the continuous-time analogue of the recursive utility defined by Epstein and Zin (1989) and Weil (1990), which allows for preferences over the timing of the resolution of uncertainty. Furthermore, equation (1.4) is a special case when the elasticity of intertemporal substitution (EIS) equals one. In what follows, γ is interpreted as risk aversion and β as the rate of time preference. I assume $\gamma > 0$ and $\beta > 0$ throughout the rest of the paper.

1.2.2. The value function and risk-free rates

Let $J(W_t, \lambda_t)$ denote the value function, where W_t denotes the real wealth of the representative agent. In equilibrium $J(W_t, \lambda_t) = V_t$.

Theorem 1.1. *Assume*

$$\left(\kappa_{\lambda_c} + \beta \right)^2 > 2\sigma_{\lambda_c}^2 E_{\nu_{cq}} \left[e^{(1-\gamma)Z_c} - 1 \right] \quad and \quad \left(\kappa_{\lambda_{cq}} + \beta \right)^2 > 2\sigma_{\lambda_{cq}}^2 E_{\nu_{cq}} \left[e^{(1-\gamma)Z_{cq}} - 1 \right]. \tag{1.5}$$

The value function J takes the following form:

$$J(W_t, \lambda_t) = \frac{W_t^{1-\gamma}}{1 - \gamma} I(\lambda_t), \tag{1.6}$$

where

$$I(\lambda_t) = \exp \left\{ a + b_c \lambda_c + b_{cq} \lambda_{cq} \right\}. \tag{1.7}$$

The coefficients a and b_j for $j \in \{c, cq\}$ take the following form:

$$a = \frac{1 - \gamma}{\beta} \left(\mu - \frac{1}{2}\gamma\sigma^2 \right) + (1 - \gamma) \log \beta + \frac{1}{\beta}b^\top \left(\kappa_\lambda * \bar{\lambda} \right), \tag{1.8}$$

$$b_j = \frac{\kappa_{\lambda_j} + \beta}{\sigma_{\lambda_j}^2} - \sqrt{\left(\frac{\kappa_{\lambda_j} + \beta}{\sigma_{\lambda_j}^2} \right)^2 - 2\frac{E_{\nu_j} \left[e^{(1-\gamma)Z_j} - 1 \right]}{\sigma_{\lambda_j}^2}}, \tag{1.9}$$

Here and in what follows, we use $*$ to denote element-by-element multiplication of vectors of equal dimension. The signs of b_c and b_{cq} determine how disaster probabilities λ_c and λ_{cq} affect the investor's value function. The following corollary shows that the investor is made worse by an increase in the disaster probabilities.

Corollary 1.1. *For $j \in \{c, cq\}$, if $Z_j < 0$, then $b_j > 0$.*

The following two corollaries provide expressions for the real and nominal risk-free rates in this economy.

Corollary 1.2. *Let r_t denote the instantaneous real risk-free rate in this economy, r_t is given by*

$$r_t = \beta + \mu - \gamma\sigma^2 + \underbrace{\lambda_{ct}E_{\nu_c}\left[e^{-\gamma Z_c}(e^{Z_c} - 1)\right]}_{\text{non-inflation disaster risk}} + \underbrace{\lambda_{cq,t}E_{\nu_{cq}}\left[e^{-\gamma Z_{cq}}(e^{Z_{cq}} - 1)\right]}_{\text{inflation disaster risk}}. \qquad (1.10)$$

The terms multiplying λ_{ct} and $\lambda_{cq,t}$ in (1.10) arise from the risk of a disaster. For $Z_j < 0$, the risk-free rate falls in λ_j: Recall that both non-inflation and inflation disasters affect consumption, therefore high disaster risk increases individuals' incentive to save, and thus lowers the risk-free rate.

Corollary 1.3. *Let $r_t^\$$ denote the instantaneous nominal risk-free rate on the nominal bond in the economy, $r_t^\$$ is given by*

$$r_t^\$ = r_t + q_t - \sigma_P^2. \qquad (1.11)$$

The nominal risk-free rate is affected by expected inflation; when expected inflation is high, investors require additional compensation to hold the nominal risk-free asset.

1.2.3. Nominal government bonds

This section provides expressions for the prices, yields, and premia for nominal zero-coupon government bonds.

Prices and yields

Nominal bond prices are determined using no-arbitrage conditions and the state-price density. Duffie and Skiadas (1994) show that the real state-price density, π_t, equals

$$\pi_t = \exp\left\{\int_0^t f_V\left(C_s, V_s\right)ds\right\} f_C\left(C_t, V_t\right),$$ (1.12)

and nominal state-price density, $\pi_t^\$$, is given by[5]

$$\pi_t^\$ = \frac{\pi_t}{P_t}.$$ (1.13)

Let $L_t^{\$,(\tau)} = L^\$(q_t, \lambda_t, \tau)$ denote the time t nominal price of a nominal government bond that pays off one nominal unit at time $t + \tau$. Then

$$L^\$(q_t, \lambda_t, s - t) = E_t\left[\frac{\pi_s^\$}{\pi_t^\$}\right].$$

The price $L_t^{\$,(\tau)}$ can be solved up to four ordinary differential equations. The following corollary is a special case of Theorem A.2 in Appendix A.1.2[6].

Corollary 1.4. *The function $L^\$$ takes the following form:*

$$L^\$(q_t, \lambda_t, \tau) = \exp\left\{a_L^\$(\tau) + b_{Lq}^\$(\tau)q_t + b_{L\lambda}^\$(\tau)^\top \lambda_t\right\},$$ (1.14)

where $b_{L\lambda}^\$(\tau) = \left[b_{L\lambda_c}^\$(\tau), \, b_{L\lambda_{cq}}^\$(\tau)\right]^\top$. The function $b_{Lq}^\$$ takes the form

$$b_{Lq}^\$(\tau) = -\frac{1}{\kappa_q}\left(1 - e^{-\kappa_q \tau}\right),$$ (1.15)

---

[5]Consider a nominal asset that has nominal payoff $X_s^\$$ at time $s > t$, the time t nominal price of the asset, $X_t^\$$, can be written as $X_t^\$ = E_t[\frac{\pi_t}{\pi_s}\frac{P_s}{P_t}X_s^\$] = E_t[\frac{\pi_s^\$}{\pi_t^\$}X_s^\$]$. Therefore, $\pi_t^\$ = \frac{\pi_t}{P_t}$.

[6]This paper focuses on inflation risk, but it can be easily extended to incorporate (outright) default risk (see Appendix) and potential government default introduce another source of risk that effect the nominal yield curve through the real yield curve.

the function $b_{L\lambda_c}^{\$}$ *solves*

$$\frac{db_{L\lambda_c}^{\$}}{d\tau} = \frac{1}{2}\sigma_{\lambda_c}b_{L\lambda_c}^{\$}(\tau)^2 + \left(b_c\sigma_{\lambda_c}^2 - \kappa_{\lambda_c}\right)b_{L\lambda_c}^{\$}(\tau) + E_{\nu_c}\left[e^{-\gamma Z_{ct}}(1 - e^{Z_{ct}})\right], \qquad (1.16)$$

the function $b_{L\lambda_{cq}}^{\$}$ *solves*

$$\frac{db_{L\lambda_{cq}}^{\$}}{d\tau} = \frac{1}{2}\sigma_{\lambda_{cq}}b_{L\lambda_{cq}}^{\$}(\tau)^2 + \left(b_{cq}\sigma_{\lambda_{cq}} - \kappa_{\lambda_{cq}}\right)b_{L\lambda_{cq}}^{\$}(\tau)$$
$$+ E_{\nu_{cq}}\left[e^{-(\gamma + b_{Lq}^{\$}(\tau))Z_{cq,t}} - e^{(1-\gamma)Z_{cq,t}}\right] + E_{\nu_q}\left[e^{-b_{Lq}^{\$}(\tau)Z_{qt}} - 1\right], \quad (1.17)$$

and the function $a_L^{\$}$ *solves*

$$\frac{da_L^{\$}}{d\tau} = -\beta - \mu + \gamma\sigma^2 + \sigma_P^2 + \frac{1}{2}\sigma_q^2 b_{Lq}^{\$}(\tau)^2 + b_{Lq}^{\$}(\tau)\kappa_q\bar{q} + b_{L\lambda}^{\$}(\tau)^\top(\kappa_\lambda * \bar{\lambda}), \qquad (1.18)$$

with boundary conditions $a_L^{\$}(0) = b_{Lq}^{\$}(0) = b_{L\lambda_c}^{\$}(0) = b_{L\lambda_{cq}}^{\$}(0) = 0$.

Corollary 1.4 shows how prices respond to innovations in expected inflation and in changing disaster probabilities. Equation (1.15) shows that innovations to expected inflation lower prices for nominal bonds of all maturities. Furthermore, the effect will be larger the more persistent it is, that is, the lower is κ_q.

Higher non-inflation disaster probability has a non-negative effect on prices. Consider the ordinary differential equation (1.16); without the last term $E_{\nu_c}\left[e^{-\gamma Z_{ct}}(1 - e^{Z_{ct}})\right]$, the function $b_{L\lambda_c}^{\$}$ is identically zero. Therefore, this term determines the sign of $b_{L\lambda_c}^{\$}$. This term can be rewritten as: $E_{\nu_c}\left[e^{-\gamma Z_{ct}}(1 - e^{Z_{ct}})\right] = -E_{\nu_c}\left[e^{-\gamma Z_{ct}}(e^{Z_{ct}} - 1)\right]$, which multiplies λ_{ct} in the equation for the nominal risk-free rate (1.11). Because higher discount rates lower the price, the risk-free rate effect enters with a negative sign. With the boundary condition $b_{L\lambda_c}^{\$}(0) = 0$, this implies that $b_{L\lambda_c}^{\$}(\tau)$ is strictly positive and increasing for all τ. The intuition is straightforward: Non-inflation disaster risks only affect the nominal bonds through the underlying real bonds, and since the real bonds in this economy pay off during consumption disaster periods, they have negative premia.

Unlike non-inflation disasters, the effect of changing inflation disaster probability on bond valuation is more complicated. Recall that this process governs both the probability of an inflation disaster and the probability of an inflation spike. Similarly to the previous argument, the last two terms in ODE (1.17) determine the sign of $b_{L\lambda_{cq}}^{\$}$. The first expectation arises from inflation disasters, and it can be rewritten as:

$$E_{\nu_{cq}}\left[e^{-(\gamma+b_{Lq}^{\$}(\tau))Z_{cq,t}} - e^{(1-\gamma)Z_{cq,t}}\right] = -\underbrace{E_{\nu_{cq}}\left[e^{-\gamma Z_{cq,t}}(e^{Z_{cq,t}} - 1)\right]}_{\text{Risk-free rate effect }(-)}$$

$$-\underbrace{E_{\nu_{cq}}\left[(e^{-\gamma Z_{cq,t}} - 1)(1 - e^{-b_{Lq}^{\$}(\tau)Z_{cq,t}})\right]}_{\text{Risk premium effect }(+)} + \underbrace{E_{\nu_{cq}}\left[e^{-b_{Lq}^{\$}(\tau)Z_{cq,t}} - 1\right]}_{\text{Nominal price effect }(-)}. \quad (1.19)$$

The first component is the risk-free rate effect; as previously discussed, this term is multiplied by a negative sign. The second component is part of the bond premium: The nominal bond price drops during periods of inflation disaster, when marginal utility is high; this term captures the premium investors require for bearing these jump risks. This risk premium effect is also multiplied by a negative sign since an increase in the discount rate lowers the bond price. The last term is the nominal price effect, which represents the effect of change in λ_{cq} on expected nominal bond prices through inflation. More specifically, it is the percent change in the price of a nominal bond with maturity τ in the event of an inflation disaster. Because a higher expected bond value raises the price, this term is multiplied by a positive sign.

Given $\gamma > 0$ and $Z_{cq} < 0$, the risk-free rate effect is negative, the risk premium effect is positive and increasing in maturity τ for $\tau > 0$, and the nominal price effect is negative and decreasing in maturity τ for $\tau > 0$. The effect of changing inflation disaster probabilities on bond value depends on the sum of these three effects. Notice that when $\tau = 0$, only the risk-rate effect is non-zero. Together with the boundary condition $b_{L\lambda_q}^{\$}(0) = 0$, this implies that $b_{L\lambda_q}^{\$}(\tau) > 0$ for some small τ: An increase in inflation disaster probability raises prices on bonds with short maturity. As maturity increases, however, risk premium and nominal price effect prevail over the risk-free rate effect, implying that prices on bonds with longer

maturity decrease with inflation disaster probability.

The last term in ODE (1.17) arises from inflation spike risks. Notice that this term represents the nominal price effect, and it enters with a positive sign. Furthermore, it is negative and decreasing in maturity τ for $\tau > 0$; implying that an increase in the chance of an inflation spike lowers nominal bond prices and the effect is stronger for bonds with longer maturity.

Before moving on to discuss bond premia, the following definition and corollary provides expression for the nominal bond yield in the model:

Definition 1.1. *The yield to maturity for a nominal bond with maturity τ at time t, denoted by $y_t^{\$,(\tau)}$, is defined as:*

$$y_t^{\$,(\tau)} = \frac{1}{\tau} \log \left(\frac{1}{L_t^{\$,(\tau)}} \right). \tag{1.20}$$

Corollary 1.4 implies that the yield to maturity in this economy takes a particularly simple form:

Corollary 1.5. *The nominal yield to maturity for a nominal bond with maturity τ at time t, $y_t^{\$,(\tau)}$, is given by*

$$y_t^{\$,(\tau)} = -\frac{1}{\tau} \left(a_L^{\$}(\tau) + b_{Lq}^{\$}(\tau) q_t + b_{L\lambda}^{\$}(\tau)^{\top} \lambda_t \right), \tag{1.21}$$

where the coefficients $a_L^{\$}(\tau)$, $b_{Lq}^{\$}(\tau)$, and $b_{L\lambda}^{\$}(\tau)$ are given by (1.15) - (1.18).

The bond premium

This section provides an expression for the instantaneous bond premium and discusses its properties. For notation simplicity, I will first define the *jump operator*, which denotes how a process responds to the occurrence of a jump. Let X be a jump-diffusion process. Define the jump operator of X with respect to the jth type of jump as the following:

$$\mathcal{J}_j(X) = X_{t_j} - X_{t_{j-}} \quad j \in \{c, cq, q\},$$

12

for t_{j-} such that a type-j jump occurs. Then define

$$\bar{\mathcal{J}}_j(X) = E_{\nu_j}\left[X_{t_j} - X_{t_j-}\right] \quad j \in \{c, cq, q\}.$$

The instantaneous nominal expected return on a nominal bond with maturity τ is simply the expected percent change in nominal prices. Let $L_t^{\$,(\tau)} = L^{\$}(q_t, \lambda_t, \tau)$ be the time-t price of a τ-year nominal bond, by Ito's Lemma:

$$
\frac{dL_t^{\$,(\tau)}}{L_{t-}^{\$,(\tau)}} = \mu_{L^{\$,(\tau)},t}\, dt + \sigma_{L^{\$,(\tau)},t}\, dB_t
$$
$$
+ \frac{1}{L_t^{\$,(\tau)}}\left(\mathcal{J}_c(L_t^{\$,(\tau)})dN_{ct} + \mathcal{J}_{cq}(L_t^{\$,(\tau)})dN_{cq,t} + \mathcal{J}_q(L_t^{\$,(\tau)})dN_{qt}\right).
$$

Then the instantaneous expected return can be written as:

$$
r_t^{\$,(\tau)} = \mu_{L^{\$,(\tau)},t} + \frac{1}{L_t^{\$,(\tau)}}\left(\lambda_{ct}\bar{\mathcal{J}}_c(L_t^{\$,(\tau)}) + \lambda_{cq,t}\left(\bar{\mathcal{J}}_{cq}(L_t^{\$,(\tau)}) + \bar{\mathcal{J}}_q(L_t^{\$,(\tau)})\right)\right). \tag{1.22}
$$

Corollary 1.6. *The bond premium relative to the risk-free rate $r^{\$}$ is:*

$$
r_t^{\$,(\tau)} - r_t^{\$} = -\lambda_t^{\top}\left(b_{L\lambda}^{\$}(\tau) * b * \sigma_{\lambda}^2\right) + \lambda_{cq,t}E_{\nu_{cq}}\left[(e^{-\gamma Z_{cq,t}} - 1)(1 - e^{-b_{L\lambda_q}^{\$}(\tau)Z_{cq,t}})\right] \tag{1.23}
$$

The first term in (1.23) arises from time-varying non-inflation and inflation disaster probabilities (time-varying probability adjustment). Recall that $b_j > 0$ for $j \in \{c, cq\}$, $b_{L\lambda_c}^{\$}(\tau) > 0$ for all τ, $b_{L\lambda_{cq}}^{\$}(\tau) > 0$ for small τ and $b_{L\lambda_{cq}}^{\$}(\tau) < 0$ for larger τ. Therefore, the time-varying non-inflation disaster probability adjustment is negative because the underlying real bond provides a hedge against consumption disasters. On the other hand, the time-varying inflation disaster probability adjustment is negative for bonds with shorter maturities and positive for bonds with longer maturities. The second term arises from the co-movement in nominal bond prices and marginal utility when a disaster occurs. Notice that this term depends on $b_{Lq}^{\$}$: When an inflation disaster occurs, expected inflation rises, which pushes

future bond prices down. Given that $b_{Lq}^{\$} < 0$ and the assumption that $\gamma > 0$, $Z_{qt} < 0$, the second term is positive.

In a sample without disasters, but possibly with inflation spikes, the observed return is

$$r_{\text{nd},t}^{\$(\tau)} = \mu_{L^{\$,(\tau)},t} + \frac{1}{L_t^{\$,(\tau)}} \lambda_{cq,t} \bar{\mathcal{J}}_{cq}(L_t^{\$,(\tau)}),$$

where the subscript "nd" is used to denote expected returns in a sample without consumption disasters. The following corollary calculates these expected returns.

Corollary 1.7. *The observed expected bond excess returns in a sample without disaster is:*

$$r_{\text{nd},t}^{\$(\tau)} - r_t^{\$} = -\lambda_t^{\top} \left(b_{L\lambda}^{\$}(\tau) * b * \sigma_{\lambda}^2 \right) + \lambda_{cq,t} E_{\nu_{cq}} \left[e^{-\gamma Z_{cq,t}} (1 - e^{-b_{Lq}^{\$}(\tau) Z_{cq,t}}) \right]. \qquad (1.24)$$

1.2.4. The aggregate market

Let D_t denote the dividend on the aggregate market. Assume that total dividends in the economy evolve according to

$$\frac{dD_t}{D_t} = \mu_D \, dt + \phi\sigma \, dB_{Ct} + (e^{\phi Z_{ct}} - 1) \, dN_{ct} + (e^{\phi Z_{qt}} - 1) \, dN_{cq,t}. \qquad (1.25)$$

Under this process, aggregate dividend responds to disasters by a greater amount than aggregate consumption does (Longstaff and Piazzesi (2004)). The single parameter, ϕ, determines how aggregate dividend responds to both normal and disaster shocks. In what follows, ϕ is referred to as leverage as it is analogous to leverage in Abel (1999).

Let $H(D_t, \lambda_t, \tau)$ denote the time t price of a single future dividend payment at time $t + \tau$. Then

$$H(D_t, \lambda_t, s - t) = E_t \left[\frac{\pi_s}{\pi_t} D_s \right],$$

where π is the real state-price density defined by (1.12). The price H can be solved in closed-form up to three ordinary differential equations, and the following corollary is a

special case of Theorem A.1 in Appendix A.1.2.

Corollary 1.8. *The function* H *takes the following form:*

$$H(D_t, \lambda_t, \tau) = D_t \exp\left\{ a_\phi(\tau) + \lambda_t^\top b_{\phi\lambda}(\tau) \right\},$$ (1.26)

where $b_{\phi\lambda} = [b_{\phi\lambda_c} b_{\phi\lambda_{cq}}]^\top$. *For* $j \in \{c, cq\}$, *function* $b_{\phi j}$ *takes the following form:*

$$b_{\phi j}(\tau) = \frac{2E_{v_j}\left[e^{(1-\gamma)Z_{jt}} - e^{(\phi-\gamma)Z_{jt}}\right]\left(1 - e^{-\zeta_{b_j}\tau}\right)}{\left(\zeta_{b_j} + b_j\sigma_j^2 - \kappa_j\right)\left(1 - e^{-\zeta_{b_j}\tau}\right) - 2\zeta_{b_j}},$$ (1.27)

where

$$\zeta_{b_j} = \sqrt{\left(b_j\sigma_j^2 - \kappa_j\right)^2 + 2\sigma_j^2 E_{\nu_j}\left[e^{(1-\gamma)Z_{jt}} - e^{(\phi-\gamma)Z_{jt}}\right]}.$$ (1.28)

Function $a_\phi(\tau)$ *takes the following form:*

$$a_\phi(\tau) = \left(\mu_D - \mu - \beta + \gamma\sigma^2(1 - \phi) \right.$$
$$\left. - \left(\frac{\kappa_{\lambda_c}\bar{\lambda}_c}{\sigma_{\lambda_c}^2}(\zeta_{b_c} + b_c\sigma_{\lambda_c}^2 - \kappa_{\lambda_c}) + \frac{\kappa_{\lambda_{cq}}\bar{\lambda}_{cq}}{\sigma_{\lambda_{cq}}^2}(\zeta_{b_{cq}} + b_{cq}\sigma_{\lambda_{cq}}^2 - \kappa_{\lambda_{cq}}) \right) \right)\tau$$
$$- \left(\frac{2\kappa_{\lambda_c}\bar{\lambda}_c}{\sigma_{\lambda_c}^2} \log\left(\frac{(\zeta_{b_c} + b_c\sigma_{\lambda_c}^2 - \kappa_{\lambda_c})(e^{-\zeta_{b_c}\tau} - 1)}{2\zeta_{b_c}} \right) \right.$$
$$\left. + \frac{2\kappa_{\lambda_{cq}}\bar{\lambda}_{cq}}{\sigma_{\lambda_{cq}}^2} \log\left(\frac{(\zeta_{b_{cq}} + b_{cq}\sigma_{\lambda_{cq}}^2 - \kappa_{\lambda_{cq}})(e^{-\zeta_{b_{cq}}\tau} - 1)}{2\zeta_{b_{cq}}} \right) \right). \quad (1.29)$$

Let $F(D_t, \lambda_t)$ denote the time t price of the claim to the entire future dividend stream. Then

$$F(D_t, \lambda_t) = \int_0^\infty H(D_t, \lambda_t, \tau)\, d\tau.$$

Equation (1.27) shows that $b_{\phi j}(\tau) < 0$ for $j \in \{c, cq\}$; therefore the price-dividend ratio,

$$G(\lambda_t) = \int_0^\infty \exp\left\{ a_\phi(\tau) + \lambda_{ct} b_{\phi\lambda_c}(\tau) + \lambda_{cq,t} b_{\phi\lambda_{cq}}(\tau) \right\} d\tau,$$ (1.30)

decreases in both non-inflation and inflation disaster probability.

1.3. Quantitative results

The model is calibrated to match aggregate consumption growth, inflation, and aggregate market moments. To evaluate the quantitative implication of the model, I simulate monthly data for 60,000 years. Furthermore, I simulate 10,000 60-year samples. For each of these small-samples, the initial values of λ_{ct} and $\lambda_{cq,t}$ are drawn from their stationary distributions, and the initial value of q_t is set equal to its mean, $\bar{q}$. In each of the tables that follow, I report the data and population value for each statistic. In addition, I report the 5th-, 50th-, and 95th-percentile values from the small-sample simulations (labelled "All Simulations" in the tables), and the 5th-, 50th-, and 95th-percentile values for the subset of the small-sample simulations that do not contain disasters (labelled "No-Disaster Simulations" in the tables). Samples in this subset do not contain any jumps in consumption, but they may contain jumps in expected inflation.

In the past 60 years, the U.S. did not experience any consumption disasters; however, it experienced a period of high inflation in the late 1970s and early 1980s. The No-Disaster subset from the simulation accommodates the possibility that there was an inflation jump in the country's postwar history; statistics from this subset therefore offer the most interesting comparison for the U.S. postwar data. With this calibration, about 23% of the samples do not experience any type of consumption disaster, and about one-third of these samples contain at least one jump in expected inflation.

1.3.1. Calibration

Data

The data on bond yields are from the Center for Research in Security Prices (CRSP). Monthly data is available for the period between June 1952 and December 2011. The yield on the three-month government bills is from the Fama risk-free rate, and yields on zero-

coupon bonds with maturities between one and five year are from the Fama-Bliss discount bond dataset.

The market return is defined as the gross return on the CRSP value-weighted index. The dividend growth rate is from the dividends on the same index. To obtain real return and dividend growth, I adjust for inflation using changes in the consumer price index, which is also available from CRSP. The price-dividend ratio is constructed as the price divided by the previous 12 months of dividends. The government bill rate is the inflation-adjusted three-month Treasury Bill return. All data are annual. I use data from 1947 to 2010; using only postwar data provides a comparison between U.S. data and the simulated samples without consumption jumps.

Parameter values

Table 1.2 reports the parameter values. Mean consumption growth and the volatility of consumption growth are both about 2%, which equal their postwar data counterparts. Mean dividend growth is set to 3.48%; it is chosen to match the price dividend ratio instead of the dividend growth in the data: CRSP dividends do not include repurchases; presumably these imply that dividends are likely to be higher sometime in the future, and that the sample mean is not a good indicator of the true mean.

The leverage parameter ϕ governs both the ratio between the volatility of log dividends and the volatility of log consumption, and how dividends response to consumption disasters. In the data, the former ratio suggests leverage to be 4.66; however, I choose a smaller value, $\phi = 3$, so that dividends have a more conservative response to consumption disasters. Rate of time preference β is set to be low to obtain a realistic short-term government bill rate. Relative risk aversion γ is set equal to 3.

Mean expected inflation is set to 2.7%; with this value, the median value of the realized inflation among the simulations with no consumption disaster is 3.65%, the value in the data is 3.74%. The volatility of non-expected inflation σ_p equals 0.8% to match the realized

inflation volatility in the data; the median value among the simulations with no consumption disaster is 2.89%, and the value in the data is 3.03%. The volatility of expected inflation σ_q equals 1.3% to match the volatility of short-term bond yield; the volatility of three-month Treasury Bill yield is 3.01% in the data, and the median value among the simulations with no consumption disaster is 2.99%. The mean reversion parameter in the expected inflation process governs the persistence of the inflation process, which is highly persistent and the autocorrelation decays slowly. This parameter it is set to 0.09 to obtain a reasonable first order autocorrelation of the inflation process.

Barro and Ursua (2008) calibrated the average probability of a consumption disaster for OECD countries to be 2.86%, implying that $\bar{\lambda}_c + \bar{\lambda}_{cq} = 2.86\%$.[7] In the data, about one-third of the disasters are accompanied by high inflation (Table 1.1), therefore I set $\bar{\lambda}_c$ to equal 1.83% and $\bar{\lambda}_{cq}$ to equal 1.03%. The persistence in the price-dividend ratio is mostly determined by the persistence in the disaster probability. I therefore choose a low rate of mean reversion for both inflation and non-inflation disaster probabilities: $\kappa_{\lambda_c} = \kappa_{\lambda_{cq}} = 0.11$. With this choice, the median value of the persistence of the price-dividend ratio among the simulations with no consumption disaster is 0.73; the value in the data is 0.92. The volatilities $\sigma_{\lambda_c} = 0.112$ and $\sigma_{\lambda_{cq}} = 0.103$ lead to a reasonable volatility for the aggregate market.

The disaster distributions Z_c and Z_{cq} are chosen to match the distribution of consumption declines. I consider 10% as the smallest possible disaster magnitude and I assume that Z_c and Z_{cq} follow power law distributions. For non-inflation disasters, I set the power law parameter to equal 10, and for inflation disasters, I set the power law parameter to equal 8. Table 1.2 plots these power law distributions along with distributions of large consumption declines. In particular, I compare the power law distribution with parameter 8 to the distribution of large consumption declines that are accompanied by high inflation, and the

---

[7]In this calibration, I calibrate the disaster probability to the OECD subsample but the size of jumps to the full set of samples. This is a more conservative approach as OECD countries have disasters that are rarer but more severe.

power law distribution with parameter 10 to the distribution of large consumption declines that are not accompanied by high inflation. In addition, I will assume that Z_q follows the same distribution as Z_{cq}.

1.3.2. Yield curves and expected returns as functions of the state variables

Yield curves

It is helpful to understand how the state variables affect the nominal yield curves in order to better understand the simulation results. Equation (1.21) shows that nominal yields on nominal bonds depend on expected inflation, q; on non-inflation disaster probability, λ_c; and on inflation disaster probability, λ_{cq}. Furthermore, the coefficients on these state variables are functions of maturity τ. Figure 1.3 shows the term in the expression for the nominal bond yield (1.21). In particular, it shows the loading on expected inflation, $-b_{Lq}/\tau$; on non-inflation disaster probability, $-b_{L\lambda_c}/\tau$; and on the inflation disaster probability, $-b_{L\lambda_{cq}}/\tau$; all as functions of maturity τ.

The loading on expected inflation is positive and decreases with maturity: High expected inflation lowers bond values and raises bond yields; due to mean-reversion, the effect is larger on bonds with shorter maturities. The loading on inflation spike probability is also positive but increases with maturity: High probability of an expected inflation jump lowers bond values and raises bond yields, and the effect is stronger on bonds with longer maturities.

What is more interesting is the stark distinction between the loading on non-inflation disaster and inflation disaster probabilities. The loading on non-inflation disaster probability is negative and decreases with maturity. While the loading on the inflation disaster probability is also negative for short maturity bonds, it increases with maturity and becomes positive. Disasters in the model affect the nominal yield curves through two channels: They affect realized consumption growth and (possibly) expected inflation. Non-inflation disasters only affects consumption growth, thus high non-inflation disaster risks lower the risk-free rate, which leads to higher bond prices and lower bond yields. Therefore, the coefficient on the

19

non-inflation disaster probability is negative and decreasing with maturity. On the contrary, inflation disaster probability affects the nominal yield curve through both channels, and investors require positive compensation for bearing the risk of jumps on expected inflation. The bottom-right panel of Figure 1.3 shows that the former effect dominates for bonds with shorter maturity while the latter effect dominates for bonds with longer maturity. Notice that inflation spike risk also affects the shape of the loading on inflation disaster probability: Inflation spike risks further lower bond values and raise bond yields, and the effect is stronger on bonds with longer maturities. Thus the presence of inflation spike risks leads to an even steeper $-b_{L\lambda_{cq}}/\tau$, though the shape is mostly determined by inflation disaster risks.

Figure 1.4 shows how the yield curve responds to changes in each of the three state variables. In each of the panels, the dashed line represents the yield curve when all state variables are at their means. I then increase the value of one state variable at a time and plot the resulting yield curve. The solid line in the top-left panel shows the yield curve when expected inflation is increased by σ_q: High expected inflation shifts the nominal yield curve up, and the effect is slightly stronger for bonds with short maturities. The solid line in the top-right panel shows the yield curve when non-inflation disaster probability is increased by one standard deviation: High non-inflation disaster probability shifts the nominal yield curve down (the risk-free rate effect), and the effect is slightly stronger for bonds with long maturities. The solid line in the bottom-left panel shows the yield curve when inflation disaster probability is increased by one standard deviation: High inflation disaster probability changes the shape of the nominal yield curve. The yields for short maturity bonds become lower (risk-free rate effect) and the yields for long maturity bonds become higher (risk-premium effect and nominal price effect). The risk of inflation spikes further increases the nominal bond yield (nominal price effect).

From this figure, one can also observe that the primary effect of expected inflation and non-inflation disaster probability is on the level of the yield curve, while non-inflation disaster

probability also affects the slope of the yield curve. Furthermore, most of the variations in the term structure variables such as the yield spreads and forward rates come from variations in the probabilities of inflation disasters and inflation spikes.

Risk premia

Figures 1.5 and 1.6 plots the bond risk premia as functions of non-inflation disaster probability, λ_c, and inflation disaster probability, λ_q, using Equation (1.23). Expected inflation q is set equal to 2.8% in all cases. To illustrate the impact of changes in disaster probabilities on bonds with different maturities, I compare the risk premia on one- and five-year bonds.

Figure 1.5 shows that risk premia decrease with non-inflation disaster probability, and also that bonds with longer maturities are more sensitive to these changes. Equation (1.23) shows that non-inflation disaster probability implies a negative premium, and that the absolute magnitude of this premium increases with maturity. Figure 1.6 shows that risk premia increase as a function of the inflation disaster probability and that bonds with longer maturities are more sensitive to these changes. The co-movement of marginal utility and bond prices in inflation disaster periods generates a positive premium for all nominal bonds, and this premium increases with maturity. Time-varying inflation disaster risks generate a small negative premium for short maturity bonds, and this premium increases with maturity and becomes positive when the maturity is longer. Comparing Figures 1.5 and 1.6, one can see that bond risk premia are more sensitive to inflation disaster risks than to non-inflation disaster risks, furthermore, one can also see that long-term bonds are more sensitive to these risks than short-term bonds.

Figure 1.4 – 1.6 provide evidence of predictable bond premia in the model. Figure 1.5 and 1.6 imply that bond excess returns are high when inflation disaster risk is high, or when non-inflation disaster risk is low. Figure 1.4 shows that yield spread is also high when inflation disaster risk is high, or when non-inflation disaster risk is low. Therefore, one should expect yield spread to have some predictive power on bond excess returns. Furthermore, since

excess returns on long-term bonds are more sensitive to these disaster probabilities than excess returns on short-term bonds are, long-term bond excess returns should be more sensitive to changes in yield spreads than excess returns of short-term bonds are.

1.3.3. Simulation results

Nominal yields

Figures 1.7 and 1.8 show the first two moments of yields for nominal bonds with different maturities. Figure 1.7 plots the data and model-implied average nominal bond yields, and Figure 1.8 plot the data and model-implied volatility of nominal bond yields, both as functions of time to maturity. In each figure, I plot the median, the 25th-, and 75th-percentile values drawn from the subset of small-sample simulations that do not contain any consumption disasters. Table 1.5 reports the data and all model-implied statistics of mean and volatility of yields for nominal bonds with one, two, three, four, and five years to maturity.

The model is capable of explaining the average nominal yield curve. The median values among the simulations with no consumption disasters are close to their data counterparts. Furthermore, the median values increase with time to maturity, implying an upward-sloping average yield curve in the model. The median small-sample value of the mean increases from 5.67% for one-year bonds to 6.03% for five-year bonds; in the data, the average bond yields increase from 5.20% for one-year bonds to 5.82% for five-year bonds. In addition to the first moment, the model also generates realistic implications for the volatility of bond yields. The median values among the simulations having no consumption disasters decreases from 2.79% for one-year bond to 2.61% for five-year bond; in the data, it decreases from 3.02% for one-year bond to 2.78% for five-year bond. Notice that the nominal yields are on average higher and more volatile in the full set of simulations and in population. This is because more jumps in expected inflation (inflation disasters) are realized, and expected inflation are on average higher in these samples.

This model is able to match both the first two moments of the nominal yield curve, while previous literature successfully capture the upward-sloping shape of the nominal yield curve, they do not generate realistic implication for the second moment. In both Piazzesi and Schneider (2006) and Bansal and Shaliastovich (2013), short-term bond yields are also more volatile than long-term bond yields, but the levels are much lower than the data counterparts. The habit formation model in Wachter (2006) implies that short-term yields are more volatile than long-term yields, which is counterfactual. Comparing with the three models, this model also impose a potentially more reasonable requirement on the utility function of the representative agent. In the current calibration, relative risk aversion is set equal to 3. In contrast, Piazzesi and Schneider (2006) set it equal to 43 and Bansal and Shaliastovich (2013) estimate it to be 20.90. The habit formation model in Wachter (2006) assumes a time-varying risk aversion, which is greater than 30 when the state variable is at its long-run mean.

One concern is that if consumption disasters co-occur with deflations instead of inflations, then it will change the implication of the model. However, deflation disasters in the data are less significant, for example, the average annual deflation rate during the Great Depression was 6.4%. Furthermore, comparing the right panel of Table 1.1 to the bottom-right panel of Table 1.2, one can see that using a power law distribution instead of the empirical distribution for jumps in inflation rates truncate the right tail of the inflation rates distribution significantly. In fact, the results do not change much if one calibrate the inflation disasters using the empirical distributions of consumption decline and inflation rates and take into account events that are associated with deflation.

Principal component analysis

Litterman and Scheinkman (1991) find that most of the variations in yield curve can be explained by a three-factor model. Specifically, the first factor affects the level of the yield curve, the second factor affects the slope, and the third factor affects the curvature. To

evaluate whether the model also exhibits this feature, I perform a principal component analysis on the data and model-simulated yields. Figure 1.11 reports the results. For the model, I only report the median values drawn from the subset of small-sample simulations that does not contain any disasters. I plot the loadings on yields with different maturities on each of the first three principal components. Similar to Litterman and Scheinkman (1991), a shock to the first principal component has similar effects across yields of different maturities (*level factor*); a shock to the second principal component raises yields on short-term bonds and reduces yields on long-term bonds (*slope factor*); and a shock to the third principal component raises yields on bonds with median maturity, but lowers yields on short- and long-term bonds (*curvature factor*). In addition, the bottom-right panel also shows that almost all the variations in yield curve are explained by the first three principal components, both in the data and in the model.

Given the three-factor structure of the model, it is natural to ask how these three factors relate to the three state variables in the model. Table 1.6 reports the correlation between each of the three state variables in the model and each of the three principal components. The level factor is mostly correlated with expected inflation; consistent with Figure 1.4, an increase in expected inflation or inflation disaster probability raises the yield curve, while an increase in non-inflation disaster probability lowers it. The slope factor is highly negatively correlated with the inflation disaster probability, and slightly positively correlated with expected inflation and non-inflation disaster probability.[8] The curvature factor is mostly correlated with non-inflation disaster risks; a shock to non-inflation disaster risks (also expected inflation and inflation disaster risks) increase the curvature of the yield curve.

Collin-Dufresne and Goldstein (2002) provide evidence of unspanned volatility using data on fixed income derivative. Their findings suggest that interest rate volatility risk cannot be hedged by bonds. Following Collin-Dufresne et al. (2009), I simulate the model to obtain 13-year samples at daily frequency. I then regress realized volatility of 6-month yields,

---

[8]Note that the loadings on the second principal component decrease with maturity, so a positive shock to this factor reduces the slope.

constructed using five daily data, on the first three principal components at the beginning of the period. Similar to Collin-Dufresne et al. (2009), these regressions yield low R^2-statistics. For example, the median value in the subset of small-samples with out disaster is around 0.03, with the 95th percentile value around 0.19, and Collin-Dufresne et al. (2009) find the R^2 to be 0.155 using data from 1990 to 2002. This suggests that the first three principal components do not forecast volatility in the model and in the data.

Time-varying bond risk premia

First I consider the "long-rate" regression in Campbell and Shiller (1991):

$$y_{t+h}^{\$,(n-h)} - y_t^{\$,(n)} = \text{constant} + \beta_n \frac{1}{n-h} \left(y_t^{\$,(n)} - y_t^{\$,(h)} \right) + \text{error}, \qquad (1.31)$$

where n denotes bond maturity and h denotes the holding period. In what follows, I will consider this regression at quarterly frequency, $h = 0.25$, from June 1952 to December 2011.

Table 1.7 reports the results for regression (1.31). I consider long-term bonds with maturities of one, two, three, four, and five years, and report the data and model-simulated coefficient of the above regression. Under the expectations hypothesis, excess returns on long-term bonds are unpredictable, which implies that β_n should equal one for all n. As in Campbell and Shiller, the coefficient β_n's are negative and decreasing in maturity n, implying that bond excess returns are predictable by yield spread, and a high yield spread predicts a higher excess return for bonds with longer maturity. The model is capable of capturing this feature. The median value of these coefficients among the simulations that contain no consumption disasters is also negative and decreasing with maturity n, furthermore, the data values are all above the 5th percentile of the values drawn from the model. In what follows, I will discuss how the mechanism drives the model's ability to explain the failure of the expectations hypothesis.

Bond risk premia are not constant in this model; (1.23) and (1.24) show that higher inflation

disaster risks lead to a higher bond risk premium, and that this premium increases with maturity. Furthermore, Figure 1.4 shows that variations in inflation disaster risk have a large effect on yield spread, and higher inflation disaster risks lead to higher yield spreads. These imply that bond premia are expected to be high when yield spread is large.[9] However, higher inflation disaster risks also lead to a higher probability of expected inflation jumps, and once these jumps are realized, bond prices drop and realized excess returns also fall. In summary, when the variations in yield spread arise from variations in non-inflation and inflation disaster probabilities – and conditional on inflation jumps are not being realized – one should expect a high yield spread to be followed by high bond premia. Furthermore, a high yield spread predicts a larger premium for long-term bonds than it does for short-term bonds. Variations in expected inflation, however, have the opposite effect on the coefficient β_n's. An increase in expected inflation leads to a lower yield spread (Figure 1.4); furthermore, it leads to a higher bond premium. Therefore, if the variations in yield spread arise from variations in expected inflation, it will have a positive effect on these coefficient β_n's.

In Table 1.7, one can see that while the median values drawn from the subset of small-sample simulations containing no consumption disasters are negative, the median values among the full set of simulations are positive. This is because there are substantially more inflation jumps among all the small-samples. While the effects of variations in λ_c and λ_{cq} dominate in the subset without disasters, the realizations of inflation jumps and variations in expected inflation dominate among the full set of small-samples.

In addition to the long-rate regressions, I also consider the forward rate regressions performed by Cochrane and Piazzesi (2005) to evaluate the model's success in capturing time-varying bond risk premia. In what follows, I consider the annual forward rate. I denote the

---

[9]Non-inflation disaster risks decrease both yield spread and bond premium, which also implies that bond premia will be high when yield spread is high.

n-year forward rate at time t for a loan from time $t+n$ to time $t+n+1$ by f_t^n, defined as:

$$f_t^n = \log L_t^{\$,(n-1)} - \log L_t^{\$,(n)}.$$

As in Cochrane and Piazzesi, these forward rate regressions are done in two steps. First I regress the average annual excess returns on two-, three-, four-, and five-year nominal bonds on all available forward rates:

$$\frac{1}{4} \sum_{n=2}^{5} r_{t+1}^{e,(n)} = \theta^\top \mathbf{f_t} + \text{error}, \tag{1.32}$$

where $r_{t+1}^{e,(n)} = r_{t+1}^{\$,(n)} - r_{t+1}^{\$,(1)}$ is the excess return of a bond with maturity n and $\mathbf{f_t}$ denotes the vector of all forward rates available at time t.

The second step is to form a single factor $\widehat{\mathrm{cp}}_{t+1} = \widehat{\theta}^\top \mathbf{f_t}$ and regress the excess returns of bonds with different maturities on this single factor:

$$r_{t+1}^{e,(n)} = \text{constant} + \rho_n \widehat{\mathrm{cp}}_{t+1} + \text{error}. \tag{1.33}$$

I consider monthly overlapping annual observations from June 1952 to December 2011. In the data, I construct one-, two-, three-, four-, and five-year forward rates. In the model, however, I can only construct three independent forward rates since the model only has three factors. Therefore, I will use all five forward rates in the data, but only one-, three-, and five-year forward rates in the model.

Table 1.8 reports the results from the second stage regression, (1.33). In the data the single forward rate factor predicts bond excess returns with an economically significant R^2, furthermore, the coefficient on this factor increases with bond maturity. The model successfully generates these findings: The median values of the R^2 drawn from the subset of the small-sample simulations containing no consumption disasters are slightly smaller than those in the data, but still economically significant. For example, the single forward rate

factor predicts excess returns on five-year nominal bonds with $R^2 = 0.18$, and the median values drawn from the subset of samples containing no disasters is 0.17. The median value of the coefficients in these samples increases from 0.40 for one-year bonds to 1.59 for five-year bonds, in the data it increases from 0.44 to 1.47. In the full set of small-samples, the R^2 are lower, but still economically meaningful. The small-sample bias in these regressions, however, is significant: The R^2-statistics are almost zero in population.

One other finding of Cochrane and Piazzesi (see also Stambaugh (1988)) is that in the first stage regression (1.32), the coefficients exhibit a tent-shaped pattern as a function of maturity. This model is also able to generate these tent-shaped patterns. In about 35% of the subset that contain no consumption disasters, the coefficients from the first stage regression (1.32) exhibit a tent-shaped pattern. Figure 1.10 reports the average of these coefficients.

The aggregate market

This model is also successful in matching moments in the aggregate market. Table 1.9 reports the simulation results. The model is able to explain most of the equity premium, which is 7.25% in the data; the median value from the small-sample containing no consumption disasters is 5.06%, and the data is below the 95th percentile of the values drawn from the model.

To calculate the real three-month Treasury Bill returns, I calculate the realized returns on the nominal three-month Treasury Bill, then adjust them by realized inflation. This model generates reasonable values for the short-term interest rate; this value in the data is 1.25%, and the median value from the small-sample containing no consumption disasters is 2.03%. Furthermore, the data value is above the 5th percentile of the values drawn from the model, indicating that we cannot reject the model at the 10% level.

The model, however, only has limited ability to explain the volatility of the price-dividend ratio. As discussed in Bansal et al. (2012) and Beeler and Campbell (2012), this is a

limitation shared by models that explain aggregate prices using time-varying moments but parsimonious preferences. Time-varying moments imply cash flow, risk-free rate, and risk premium effects, and one of these generally acts as an offset to the other two, thus limiting the effect time-varying moments have on prices.

Interactions between the aggregate and bond market

In this section, I study the model's implications for the interaction between the aggregate market and the term structure of interest rates. Previous works have shown that variables that predict excess returns in one asset class often fail in another. For example, Duffee (2012) showed that while term structure variables predict bond excess returns, they do not predict stock market excess returns. In this section, I consider two predictor variables, the price-dividend ratio and the linear combination of forward rates that best predicts bond returns. I also consider two excess returns, the aggregate market returns over short-term bonds, and the average long-term bond returns over short-term bonds. The average long-term bond return is defined as the average of the returns on one-, two-, three-, four- and five-year nominal bonds. I calculate the predictive regressions of each excess returns on each predictor variable. Data are annual from 1953 to 2010. Tables 1.10 – 1.13 report the results from these predictive regressions.

Tables 1.10 and 1.12 show the results of regressing aggregate and bond market excess returns on the price-dividend ratio. It is well known that price-dividend ratio predicts aggregate market excess returns in the data (e.g. Campbell and Shiller (1988), Cochrane (1992), Fama and French (1989) and Keim and Stambaugh (1986)). Equation (1.30) shows that the price-dividend ratio in the model is governed by both non-inflation and inflation disaster probabilities. In particular, a high disaster probability lowers the price-dividend ratio. Furthermore, investors require a higher-than-average premium when the total disaster risk is high, implying that on average, a high total disaster probability is followed by high returns. Notice that predictability still exists in the full set of simulations, though the

29

R^2-statistics are smaller. This is because realized returns are much lower when disasters actually occur. In population, the predictability is even smaller, reflecting the well-known small-sample bias in predictive regressions.

In the data, the price-dividend ratio also has some predictive power on long-term bond excess returns, though the t-statistics are not significant and R^2 values are low, as shown in Tables 1.10 and 1.12. The model generates similar implications. An increase in either non-inflation or inflation disaster probability leads to a low price-dividend ratio; bond excess return, however, is governed mainly by inflation disaster probability. Therefore, investors require a higher-than-average bond premium only when inflation disaster risk is high, implying that on average, high inflation disaster probability is followed by high bond returns. Furthermore, a high non-inflation disaster probability lowers the expected bond excess returns; nonetheless, this effect is substantially smaller. Therefore, if the variation in the price-dividend ratio comes from the inflation disaster probability, then the price-dividend ratio predicts long-term bond excess returns with a negative sign. On the other hand, if the movement in the price-dividend ratio comes from the non-inflation disaster probability, then the price-dividend ratio predicts long-term bond excess returns with a small but positive sign. Notice that predictability still exists in the full set of simulations, but disappears in population.

As shown in previous section, long-term bond excess returns can be predicted using a linear combination of forward rates. Tables 1.11 and 1.13 report the results of the long-horizon regression. Unsurprisingly, the model successfully generates the long-term bond excess return predictability found in the data. In the model, both the shape of the term structure and bond excess returns are largely determined by the inflation disaster probability: A high inflation disaster probability leads to a steeper term structure and also a higher bond premium.

On the other hand, the linear combination of forward rates has less predictive power on the aggregate market excess returns (Duffee (2012)). In the full sample from 1953 to 2010, the

linear combination of forward rates appears to have no predictive power.[10] In the model, forward rates depend on inflation disaster probability, and high inflation disaster probability is, on average, followed by high returns. Therefore, the linear combination of forward rates still predicts aggregate market excess returns. However, comparing Panel A and Panel B of Tables 1.11 and 1.13, one can see that the linear combination of forward rates predicts the long-term bond excess returns with a much higher R^2 value than for the aggregate market excess returns, implying that the forward rate factor has a stronger predictive power on bond excess returns.

Lettau and Wachter (2011) also consider these regressions; the single forward rate factor in their model predicts bond excess returns and aggregate market excess returns with similar R^2 values. In the data, even though the R^2 depends on the sample period, the forward rate factor has a stronger predictive power on bond excess returns. Wachter (2006) and Gabaix (2012) also study both the stock and bond markets. The model of Wachter (2006), however, implies that the risk premia on stocks and bonds move together. In Gabaix (2012), the time-varying risks in stock and bond market are unrelated, where in this paper, the underlying risks are the same, but they have different effect on the premia. The model in this paper is able to generate more realistic implications for these predictive regressions because the prices of risks in the model have a two-factor structure, and these factors have differential effects on the stock and bond markets.

1.4. Conclusion

Why is the average term structure upward-sloping? Why are excess returns on nominal bonds predictable? This paper provides an explanation for these questions using a model with time-varying rare disaster risks. Previous research has shown that a model that includes time-varying disaster risks can generate high equity premium and excess returns

---

[10]The magnitude of the R^2-statistics depends on the subsample. For example, Cochrane and Piazzesi (2005) find that the linear combination of forward rates predicts one-year aggregate market excess returns with an $R^2 = 0.07$ in the sample from 1964 through 2003. In the corresponding period, the R^2 is 0.36 for one-year nominal bond excess returns.

volatility. Motivated by historical data, disasters in this model affect not only aggregate consumption, but also expected inflation. A jump in expected inflation pushes down the real value of nominal bonds, and investors require compensation for bearing these inflation disaster risks. Furthermore, this premium increases with bond maturity, which leads to an upward-sloping nominal term structure. Time-varying bond risk premia arise naturally from time-varying disaster probabilities, and prices of risk in this model follow a two-factor structure.

The model is calibrated to match the aggregate consumption, inflation, and equity market moments, and the quantitative results show that this model produces realistic means and volatilities of nominal bond yields. The three state variables in the model are highly correlated with the first three principal components, which explain almost all of the variations in the nominal yield curve both in the model and in the data. This model can also account for the violation of the expectations hypothesis. In particular, I show that the yield spread and a linear combination of forward rates can predict long-term bond excess returns. Furthermore, the model is capable of capturing the joint predictive properties of the aggregate market returns and of the bond returns. Aggregate market variables have higher predictive powers for equity excess returns while the term structure variables have higher predictive powers for bond excess returns.

Figure 1.1: Inflation disasters: Distribution of consumption declines and inflation rates
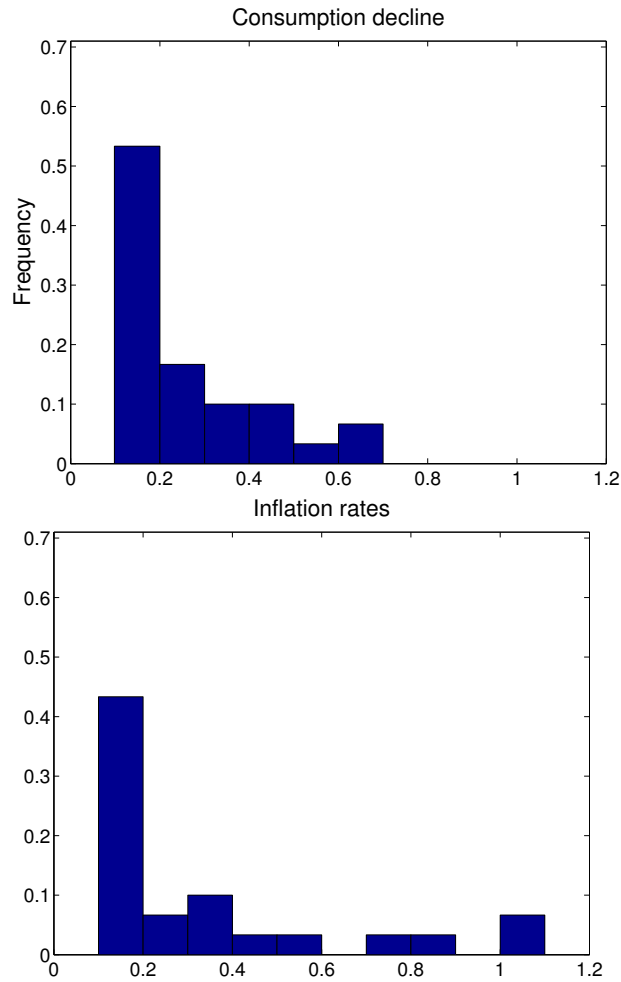
Notes: Histograms show the distribution of large consumption declines (peak-to-trough measure) and high inflation (average annual inflation rate) in periods where large consumption declines and high inflation co-occur. These figures exclude eight events in which average annual inflation rates exceeded 100%. Data from Barro and Ursua (2008).

Figure 1.2: Data vs. model consumption declines
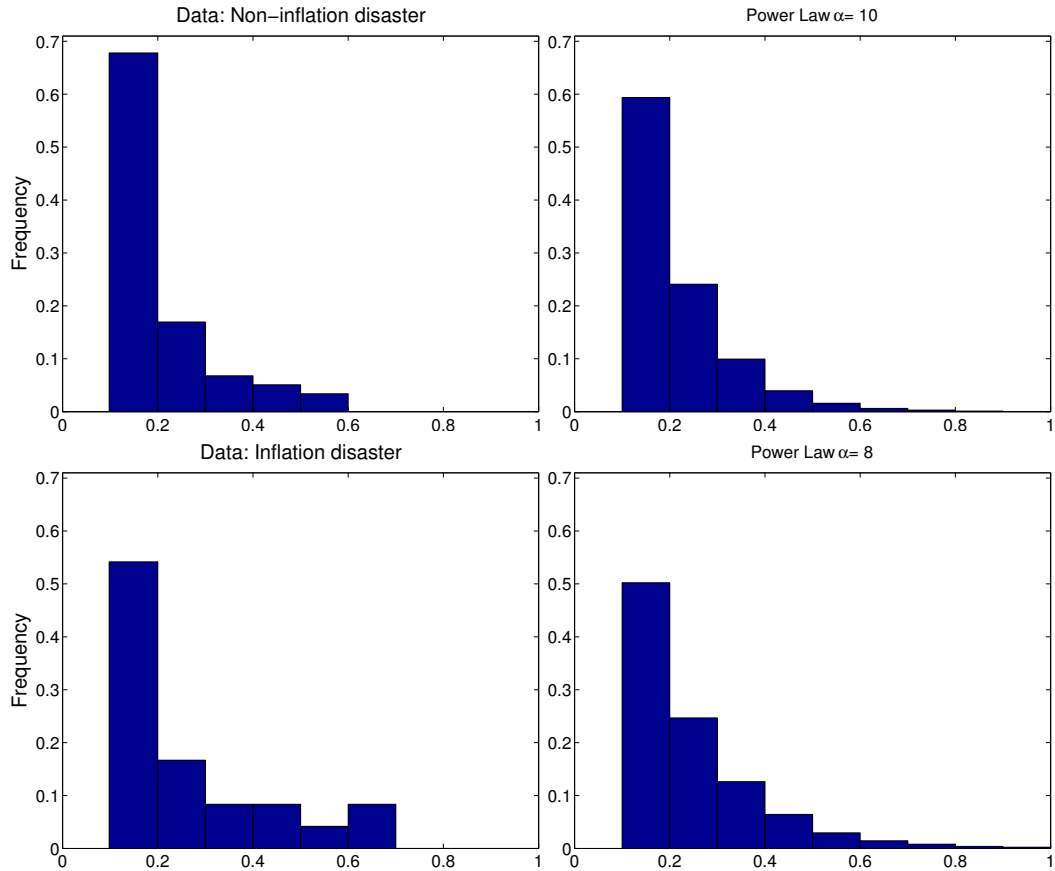
Notes: This figure plots the distributions of large consumption declines in the data and the power law distribution used in the model. The top-left panel plots the distributions of large consumption declines that do not co-occur with high inflation and the top-right panel plots the power law distribution with parameter 10. The bottom-left panel plots the distributions of large consumption declines that co-occur with high inflation and the bottom-right panel plots the power law distribution with parameter 8. Data from Barro and Ursua (2008).

Figure 1.3: Solution for the nominal bond yield
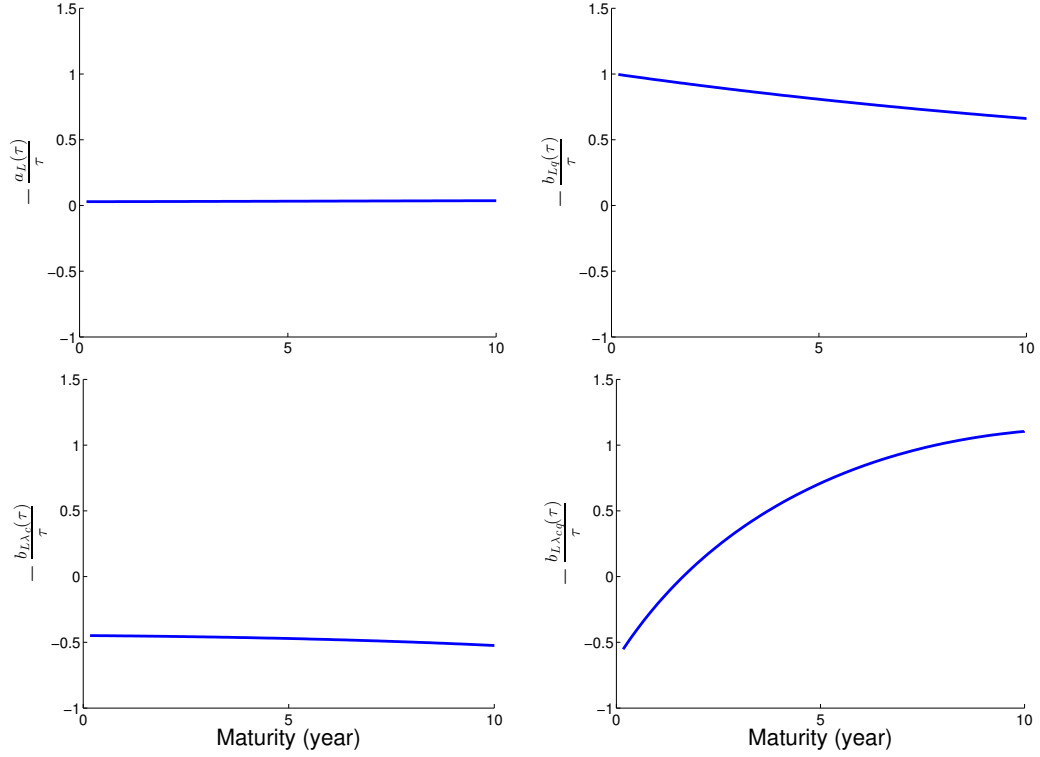
Notes: The nominal yield of a bond with maturity τ is

$$y_t^{\$,(\tau)} = -\frac{1}{\tau}\left(a_L(\tau) + b_{Lq}(\tau)q_t + b_{L\lambda}(\tau)^\top \lambda_t\right).$$

The top-left panel plots the constant term, the top-right panel plots the coefficient multiplying q_t (expected inflation), the bottom-left panel plots the coefficient multiplying λ_c (non-inflation disaster probability), and the bottom right panel plots the coefficient multiplying λ_{cq} (inflation disaster probability). All are plotted as functions of years to maturity (τ).

Figure 1.4: Yield curve as functions of the state variables
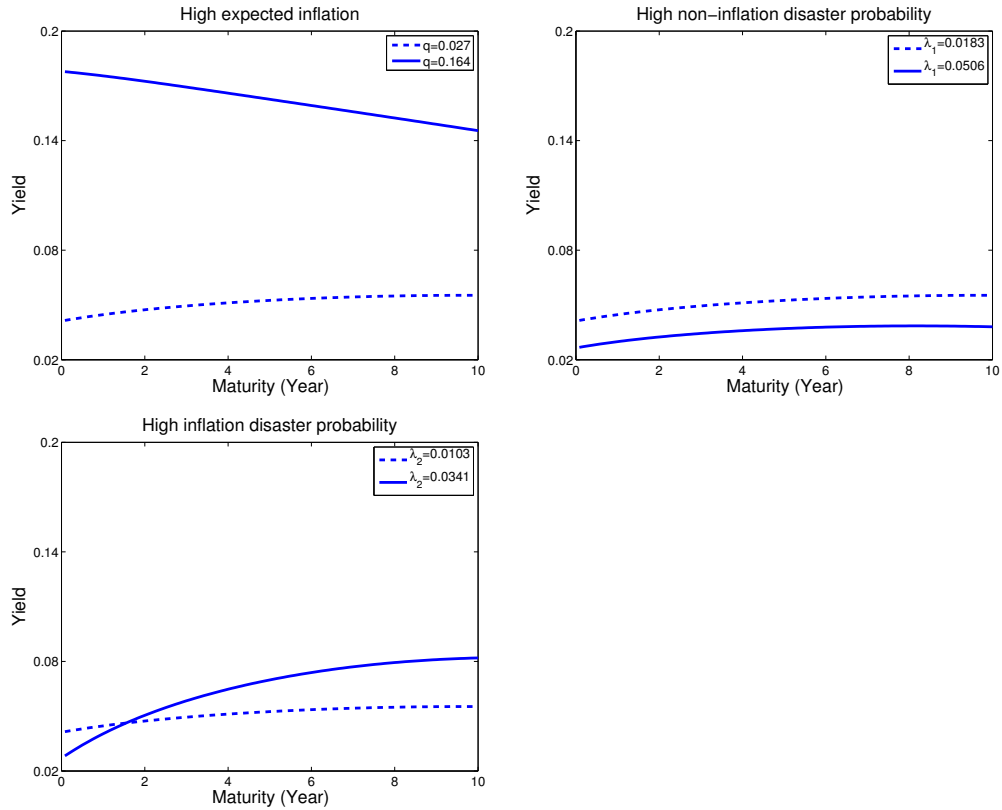
Notes: The figure plots the responses of the nominal yield curve to a shock of standard deviation on each of the three state variables. The dashed line represents the yield curve when all variables are fixed at their means. The solid line in the top-left panel represents high expected inflation; the solid line in the top-right panel represents high non-inflation disaster probability; and the solid line in the bottom-left panel represents high inflation disaster probability.

Figure 1.5: Risk premiums as a function of non-inflation disaster probability
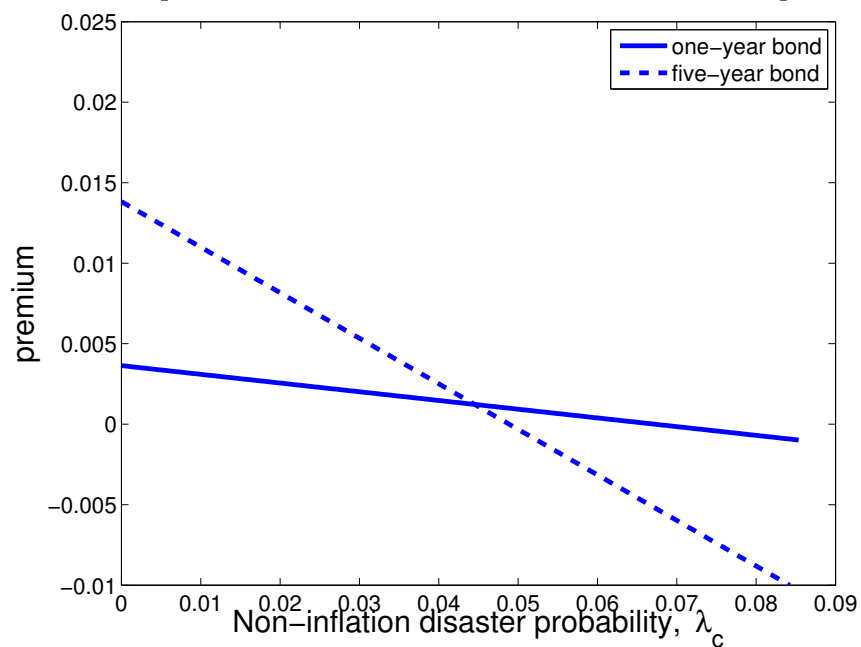
Notes: This figure shows the instantaneous expected nominal return on a one-year nominal zero coupon bond above the nominal risk-free rate (solid line) and the analogous premium for the five-year nominal zero coupon bond (dashed line). Premiums are shown as a function of the non-inflation disaster probability, λ_1, while λ_2 is fixed at its mean of 1.03%. Premiums are in annual terms.

Figure 1.6: Risk premiums as a function of inflation disaster probability
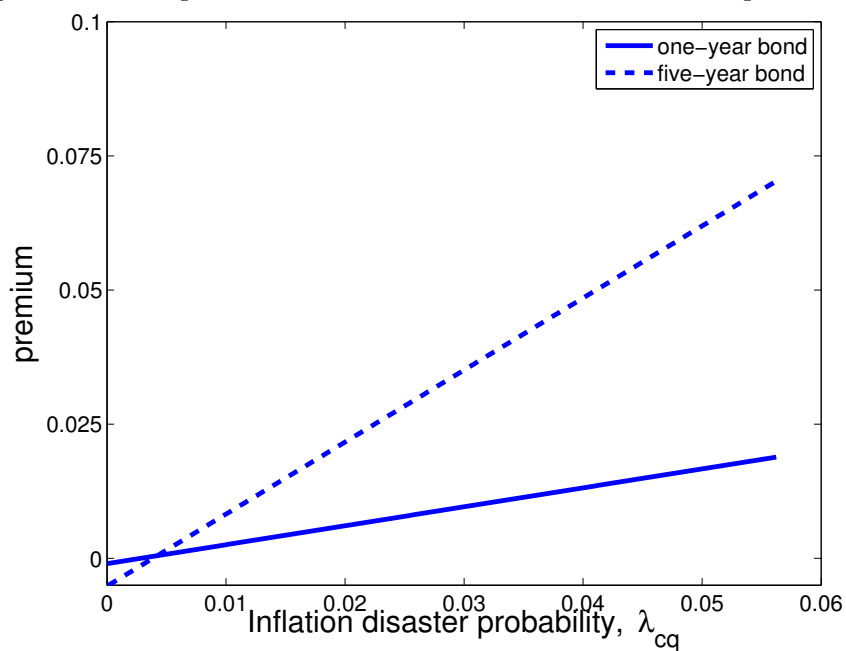
Notes: This figure shows the instantaneous expected nominal return on a one-year nominal zero coupon bond above the nominal risk-free rate (solid line) and the analogous premium for the five-year nominal zero coupon bond (dashed line). Premiums are shown as a function of the disaster probability, λ_2, while λ_1 is fixed at its mean of 1.83%. Premiums are in annual terms.

Figure 1.7: Average bond yield
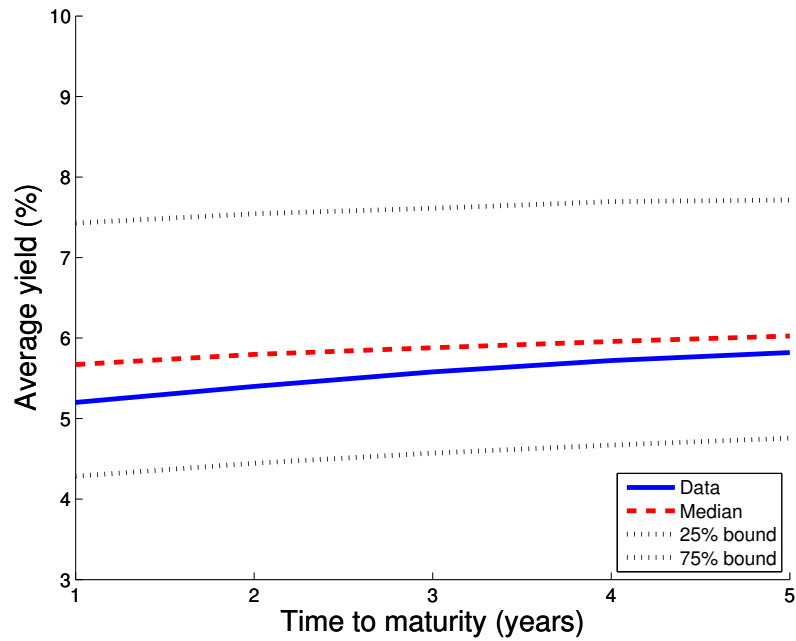
Notes: This figure plots the data and model-implied average nominal bond yield as a function of years to maturity. The solid line plots the average nominal bond yields in the data. The dashed line plots the median average bond yields in the small sample containing no consumption disasters, and the dotted lines plot the 25% and 75% bounds. Data moments are calculated using monthly data from 1952 to 2011. Data are constructed using the Fama-Bliss dataset from CRSP. All yields are in annual terms.

Figure 1.8: Volatility of bond yield
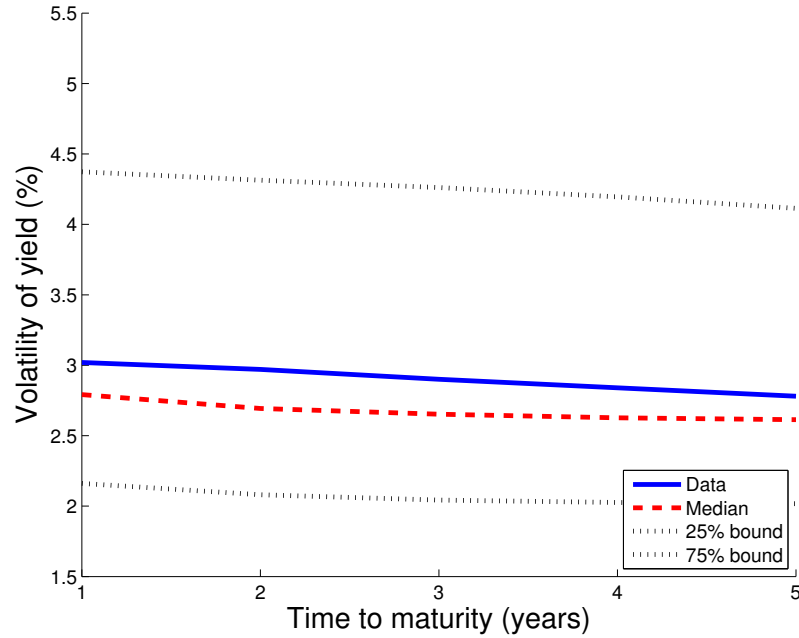
Notes: This figure plots the data and model-implied volatility of nominal bond yield as a function of years to maturity. The solid line plots the volatility of nominal bond yields in the data. The dashed line plots the median volatility of bond yields in the small-samples containing no consumption disasters, and the dotted lines plot the 25% and 75% bounds. Data moments are calculated using monthly data from 1952 to 2011. Data are constructed using the Fama-Bliss dataset from CRSP. All yields are in annual terms.

Figure 1.9: Campbell-Shiller long rate regression
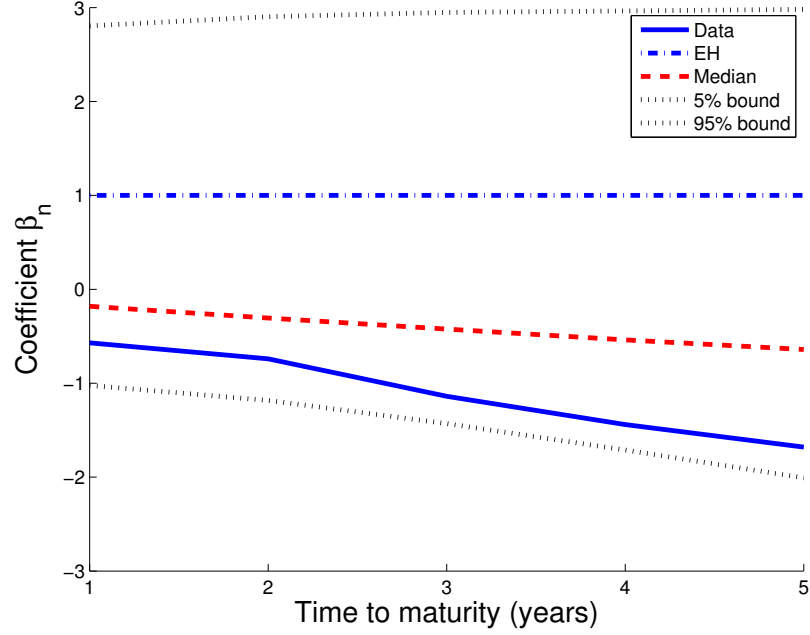
Notes: This figure reports the coefficients of the Campbell-Shiller regression.

$$y_{t+h}^{\$,(n-h)} - y_t^{\$,(n)} = \text{constant} + \beta_n \frac{1}{n-h} \left(y_t^{\$,(n)} - y_t^{\$,(h)} \right) + \text{error},$$

where $h = 0.25$. The solid line plots the coefficients in the data. The dash-dotted line plots the coefficients under the expectation hypothesis. The dashed line plots the median value of the coefficients in the small-samples containing no consumption disasters, and the dotted lines plot the 5% and 95% bounds. Data moments are calculated using monthly data from 1952 to 2011. Data are constructed using Fama-Bliss dataset from the CRSP.

Figure 1.10: Forward rate regression - First stage estimates
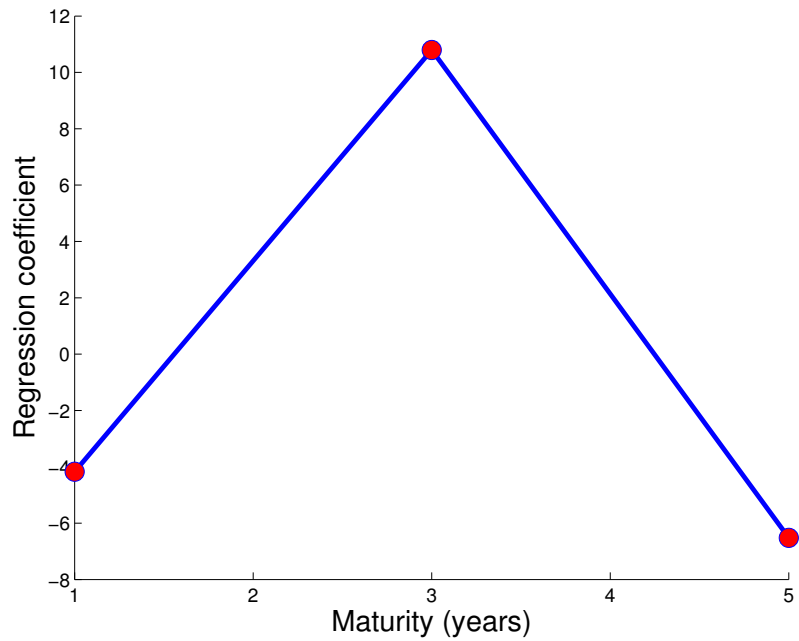
Notes: This figure plots the coefficient from regressing average excess bond returns on forward rates in the model. Average annual returns on two-, three-, four-, and five-year nominal bonds, in excess of the return on the one-year bond, are regressed on the one-, three-, and five-year forward rates. The figure shows the resulting coefficients as a function of the forward-rate maturity. About 35% of the small-sample having no consumption disaster have coefficients that form a tent shape, and this figure plots the average of the coefficients in these samples.

Figure 1.11: Principal component analysis
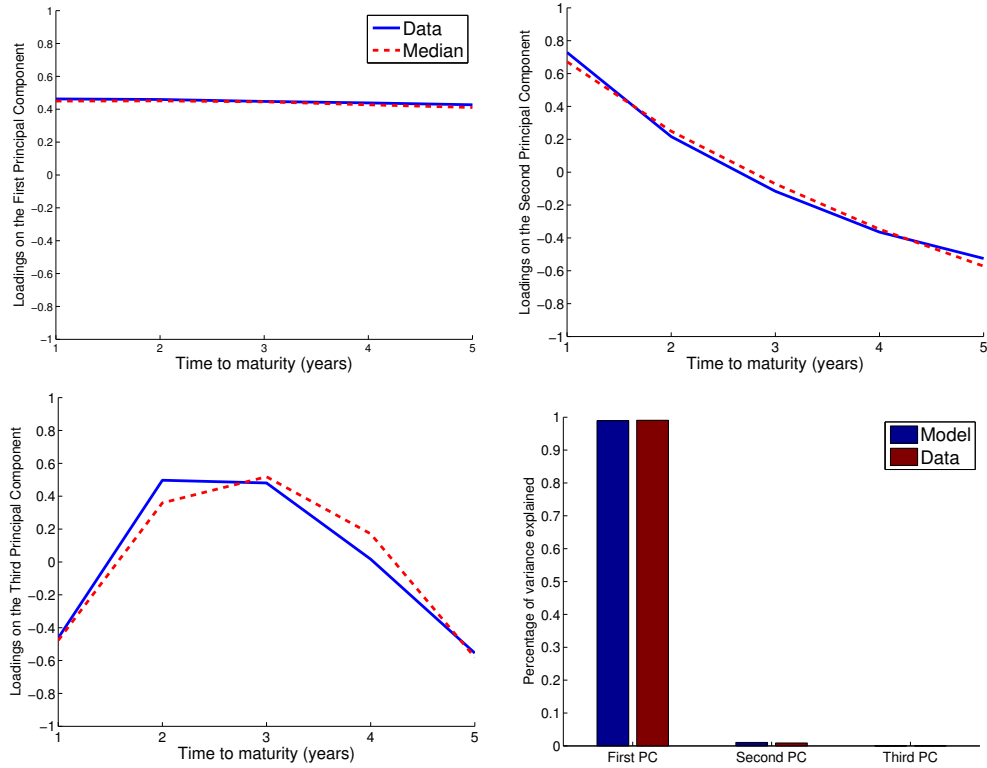
Notes: This figure plots the results from the principal component analysis. I report the median values from the subset of small-sample simulations that do not contain any disasters. The top-left panel plots the loadings on the first principal component, the top-right panel plots the loadings on the second principal component, and the bottom-left panel plots the loadings on the third principal component. The bottom-right panel shows the percentage of variance explained by each of the principal components. Data are available at monthly frequency from June 1952 to December 2011.

Table 1.1: Summary statistics of consumption disasters

Panel A: All countries	
Number of consumption disasters	89
Number of consumption disasters with high inflation	30
Percentage of consumption disasters with high inflation (%)	33.71
Panel B: OECD countries	
Number of consumption disasters	53
Number of consumption disasters with high inflation	17
Percentage of consumption disasters with high inflation (%)	32.08

Table 1.2: Parameters

Panel A: Basic parameters	
Average growth in consumption (normal times) $\bar{\mu}$ (%)	2.02
Average growth in dividend (normal times) μ_D (%)	3.48
Volatility of consumption growth (normal times) σ (%)	2.00
Leverage ϕ	3.0
Rate of time preference β	0.010
Relative risk aversion γ	3.0
Panel B: Inflation parameters	
Average inflation $\bar{q}$ (%)	2.70
Volatility of expected inflation σ_q (%)	1.30
Volatility of realized inflation σ_p (%)	0.80
Mean reversion in expected inflation κ_q	0.09
Panel C: Non-inflation disaster parameters	
Average probability of non-inflation disaster $\bar{\lambda}_c$ (%)	1.83
Mean reversion in non-inflation disaster probability κ_{λ_c}	0.11
Volatility parameter for non-inflation disaster σ_{λ_c}	0.112
Minimum non-inflation disaster (%)	10
Power law parameter for non-inflation disaster	10
Panel D: Inflation disaster parameters	
Average probability of inflation disaster $\bar{\lambda}_{cq}$ (%)	1.03
Mean reversion in inflation disaster probability $\kappa_{\lambda_{cq}}$	0.11
Volatility parameter for inflation disaster $\sigma_{\lambda_{cq}}$	0.103
Minimum inflation disaster (%)	10
Power law parameter for inflation disaster	8

Table 1.3: Log consumption and dividend growth moments

Panel A: Consumption growth

	Data	No-Disaster Simulations			All Simulations			Population
		0.05	0.50	0.95	0.05	0.50	0.95	
mean	1.91	1.57	2.00	2.42	−0.19	1.45	2.22	1.29
standard deviation	1.41	1.68	1.99	2.29	1.85	3.86	9.02	5.05
skewness	−0.48	−0.51	−0.01	0.49	−6.06	−3.20	0.22	−6.61
kurtosis	3.49	2.22	2.82	3.97	2.50	16.18	43.29	69.83

Panel B: Dividend growth

	Data	No-Disaster Simulations			All Simulations			Population
		0.05	0.50	0.95	0.05	0.50	0.95	
mean	1.78	2.01	3.29	4.56	−3.28	1.64	3.97	1.16
standard deviation	6.57	5.05	5.97	6.86	5.56	11.57	27.06	15.14
skewness	−0.01	−0.51	−0.01	0.49	−6.06	−3.20	0.22	−6.61
kurtosis	5.26	2.22	2.82	3.97	2.50	16.18	43.29	69.83

Notes: Data moments are calculated using annual data from 1947 to 2010. Population moments are calculated by simulating data from the model at a monthly frequency for 60,000 years and then aggregating monthly growth rates to an annual frequency. I also simulate 10,000 60-year samples and report the 5th-, 50th- and 95th-percentile for each statistic both from the full set of simulations and for the subset of samples for which no consumption disasters occur.

Table 1.4: Inflation moments

| | Data | No-Disaster Simulations | | | All Simulations | | | Population |
		0.05	0.50	0.95	0.05	0.50	0.95	
Mean	3.74	0.27	3.65	12.59	0.91	6.12	28.51	11.16
Standard deviation	3.03	1.77	2.89	13.55	1.92	5.54	31.34	20.63
AC(1)	0.66	0.61	0.84	0.93	0.65	0.87	0.95	0.95

Notes: Data moments are calculated using annual data from 1947 to 2010. Population moments are calculated by simulating data from the model at a monthly frequency for 60,000 years and then aggregating monthly growth rates to an annual frequency. I also simulate 10,000 60-year samples and report the 5th-, 50th- and 95th-percentile for each statistic both from the full set of simulations and for the subset of samples for which no consumption disasters occur. All numbers are in annual level terms.

Table 1.5: Nominal Yield Moments

Panel A: Average nominal bond yield

Maturity	Data	No-Disaster Simulations			All Simulations			Population
		0.05	0.50	0.95	0.05	0.50	0.95	
1-year	5.20	2.39	5.67	13.09	2.49	7.39	23.26	10.86
2-year	5.40	2.66	5.80	13.06	2.71	7.51	23.27	10.92
3-year	5.58	2.88	5.88	12.93	2.89	7.60	23.15	10.92
4-year	5.72	3.03	5.96	12.85	3.03	7.65	22.97	10.89
5-year	5.82	3.18	6.03	12.71	3.14	7.67	22.68	10.83

Panel B: Volatility of nominal bond yield

Maturity	Data	No-Disaster Simulations			All Simulations			Population
		0.05	0.50	0.95	0.05	0.50	0.95	
1-year	3.02	1.67	2.79	10.96	1.90	5.09	20.67	12.97
2-year	2.97	1.62	2.69	10.51	1.83	4.96	20.18	12.72
3-year	2.90	1.58	2.65	10.22	1.80	4.88	19.59	12.46
4-year	2.84	1.56	2.63	9.87	1.77	4.80	19.13	12.20
5-year	2.78	1.54	2.61	9.58	1.76	4.73	18.67	11.93

Notes: Panel A reports the average nominal bond yield and Panel B reports the volatility of the nominal bond yield. Data moments are calculated using monthly data from 1952 to 2011. Population moments are calculated by simulating data from the model at a monthly frequency for 60,000 years. I also simulate 10,000 60-year samples and report the 5th-, 50th- and 95th-percentile for each statistic both from the full set of simulations and for the subset of samples for which no consumption disasters occur. All yields are in annual terms.

Table 1.6: Correlation between principal components and state variables

	PC1	PC2	PC3
expected inflation	0.92	0.09	0.11
non-inflation disaster risks	−0.05	0.07	0.82
inflation disaster risks	0.06	−0.90	0.07

Notes: This table reports the correlation between each principal component and each state variable in the model. I report the median value drawn from the subset of small-sample simulations having no consumption disasters.

Table 1.7: Campbell-Shiller long rate regression

		No-Disaster Simulations			All Simulations			
Maturity	Data	0.05	0.50	0.95	0.05	0.50	0.95	Population
1-year	-0.57	-1.02	-0.18	2.80	-0.93	0.31	3.65	0.44
2-year	-0.74	-1.18	-0.31	2.90	-1.08	0.30	3.76	0.57
3-year	-1.14	-1.43	-0.42	2.95	-1.31	0.27	3.87	0.67
4-year	-1.44	-1.71	-0.54	2.96	-1.56	0.25	3.93	0.74
5-year	-1.68	-2.01	-0.64	2.98	-1.80	0.23	3.96	0.79

Notes: This table reports the coefficients of the Campbell-Shiller regression.

$$y_{t+h}^{\$,(n-h)} - y_t^{\$,(n)} = \text{constant} + \beta_n \frac{1}{n-h} \left(y_t^{\$,(n)} - y_t^{\$,(h)} \right) + \text{error},$$

where $h = 0.25$ and each row represents a bond with a different maturity (n). Data moments are calculated using quarterly data from June 1952 to December 2011.

Table 1.8: Cochrane-Piazzesi forward rate regression

Panel A: Coefficient

Maturity	Data	No-Disaster Simulations			All Simulations			Population
		0.05	0.50	0.95	0.05	0.50	0.95	
2-year	0.44	0.32	0.40	0.48	0.33	0.41	0.49	0.54
3-year	0.83	0.73	0.80	0.87	0.74	0.81	0.88	0.90
4-year	1.26	1.19	1.20	1.21	1.19	1.20	1.21	1.17
5-year	1.47	1.46	1.59	1.73	1.44	1.57	1.71	1.39

Panel B: R^2-statistics

Maturity	Data	No-Disaster Simulations			All Simulations			Population
		0.05	0.50	0.95	0.05	0.50	0.95	
2-year	0.16	0.02	0.15	0.48	0.01	0.11	0.41	0.01
3-year	0.17	0.03	0.16	0.47	0.02	0.11	0.41	0.01
4-year	0.20	0.03	0.16	0.44	0.02	0.12	0.39	0.01
5-year	0.18	0.03	0.17	0.41	0.02	0.12	0.37	0.01

Notes: This table reports the results from the second stage of the Cochrane-Piazzesi single factor regression. It reports the coefficient on the linear combination of forward rates on nominal bonds and the R^2-statistics from regressing excess bond return on the single forward rate factor. I consider bonds with maturities of two, three, four and five years. Data are monthly from 1952 to 2011.

Table 1.9: Market moments

	Data	No-Disaster Simulations			All Simulations			Population
		0.05	0.50	0.95	0.05	0.50	0.95	
$E[R^{(0.25)}]$	1.25	1.03	2.03	2.54	-0.56	1.57	2.40	1.35
$\sigma(R^{(0.25)})$	2.75	0.90	1.26	2.13	0.98	1.64	3.29	2.15
$E[R^m - R^{(0.25)}]$	7.25	3.12	5.06	7.87	2.09	4.78	8.57	5.04
$\sigma(R^m)$	17.79	9.68	13.75	19.91	11.21	17.80	27.44	18.91
Sharpe ratio	0.41	0.25	0.37	0.50	0.11	0.28	0.45	0.27
$\exp(E[p-d])$	32.51	30.71	36.13	39.06	24.48	33.77	38.35	32.66
$\sigma(p-d)$	0.43	0.07	0.15	0.30	0.09	0.21	0.44	0.29
$AR1(p-d)$	0.92	0.46	0.73	0.90	0.53	0.79	0.92	0.88

Notes: Data moments are calculated using annual data from 1947 to 2010. Population moments are calculated by simulating monthly data from the model for 60,000 years and then aggregating to an annual frequency. We also simulate 10,000 60-year samples and report the 5th-, 50th-, and 95th-percentile for each statistic from the full set of simulations and for the subset of samples for which no disasters occur. $R^{(0.25)}$ denotes the three-month Treasury Bill return where $R^{(0.25)} = R_t^{\$,(0.25)} \frac{P_{t+1}}{P_t}$. R^m denotes the return on the aggregate market, and $p-d$ denotes the log price-dividend ratio.

Table 1.10: Long-horizon regressions of returns on the price-dividend ratio (One-year holding period)

	Data	t-stat	No-Disaster Simulations			All Simulations			Population
			0.05	0.50	0.95	0.05	0.50	0.95	
			Panel A: Aggregate Market						
Coef.	−0.12	[−1.89]	−0.63	−0.34	−0.17	−0.53	−0.23	0.03	−0.13
R^2	0.07		0.07	0.17	0.29	0.00	0.08	0.23	0.04
			Panel B: Bond Market						
	Data	t-stat	No-Disaster Simulations			All Simulations			Population
			0.05	0.50	0.95	0.05	0.50	0.95	
Coef.	0.02	[1.19]	−0.15	−0.00	0.13	−0.13	0.02	0.18	0.02
R^2	0.02		0.00	0.03	0.23	0.00	0.02	0.19	0.00

Notes: This table reports the results from regressing one-year aggregate market excess returns and average nominal bond excess return on the price-dividend ratios. Data are annual from 1953 to 2010. For the data coefficients, I report t-statistics constructed using Newey-West standard errors. Population moments are calculated by simulating monthly data from the model for 60,000 years and then aggregating to an annual frequency. I also simulate 10,000 60-year samples and report the 5th-, 50th- and 95th-percentile for each statistic from the full set of simulations and for the subset of samples for which no disasters occur.

Table 1.11: Long-horizon regressions of returns on the linear combination of forward rates (One-year holding period)

	Data	t-stat	No-Disaster Simulations			All Simulations			Population
			0.05	0.50	0.95	0.05	0.50	0.95	
Panel A: Aggregate Market									
Coef.	0.72	[0.48]	−3.14	0.39	2.84	−3.06	0.54	3.24	−0.28
R^2	0.00		0.00	0.03	0.16	0.00	0.02	0.14	0.00
Panel B: Bond Market									
Coef.	1.03	[3.19]	0.86	1.17	1.55	0.81	1.20	1.68	1.59
R^2	0.20		0.02	0.17	0.45	0.01	0.12	0.40	0.01

Notes: This table reports the results from regressing one-year aggregate market excess returns and average nominal bond excess return on the linear combination of forward rates. Data are annual from 1953 to 2010. For the data coefficients, I report t-statistics constructed using Newey-West standard errors. Population moments are calculated by simulating monthly data from the model for 60,000 years and then aggregating to an annual frequency. I also simulate 10,000 60-year samples and report the 5th-, 50th- and 95th-percentile for each statistic from the full set of simulations and for the subset of samples for which no disasters occur.

Table 1.12: Long-horizon regressions of returns on the price-dividend ratio (Five-year holding period)

| | | | Panel A: Aggregate Market | | | | | | |
| | | | No-Disaster Simulations | | | All Simulations | | | |
	Data	t-stat	0.05	0.50	0.95	0.05	0.50	0.95	Population
Coef.	-0.28	$[-2.87]$	-1.55	-1.05	-0.57	-1.51	-0.84	0.13	-0.52
R^2	0.13		0.13	0.44	0.69	0.01	0.25	0.61	0.12
			Panel B: Bond Market						
			No-Disaster Simulations			All Simulations			
	Data	t-stat	0.05	0.50	0.95	0.05	0.50	0.95	Population
Coef.	0.07	$[1.83]$	-0.54	0.00	0.43	-0.51	0.08	0.62	0.07
R^2	0.09		0.00	0.07	0.47	0.00	0.06	0.46	0.01

Notes: This table reports the results from regressing five-year aggregate market excess returns and average nominal bond excess return on the price-dividend ratios. Data are annual from 1953 to 2010. For the data coefficients, I report t-statistics constructed using Newey-West standard errors. Population moments are calculated by simulating monthly data from the model for 60,000 years and then aggregating to an annual frequency. I also simulate 10,000 60-year samples and report the 5th-, 50th- and 95th-percentile for each statistic from the full set of simulations and for the subset of samples for which no disasters occur.

Table 1.13: Long-horizon regressions of returns on the linear combination of forward rates (Five-year holding period)

	Data	t-stat	Panel A: Aggregate Market						
			No-Disaster Simulations			All Simulations			
	Data	t-stat	0.05	0.50	0.95	0.05	0.50	0.95	Population
Coef.	2.02	[0.68]	−11.03	0.97	8.61	−11.61	1.41	10.89	−1.21
R^2	0.01		0.00	0.08	0.45	0.00	0.07	0.40	0.00
			Panel B: Bond Market						
			No-Disaster Simulations			All Simulations			
	Data	t-stat	0.05	0.50	0.95	0.05	0.50	0.95	Population
Coef.	1.83	[2.53]	1.04	3.37	5.83	0.84	3.66	7.18	6.45
R^2	0.11		0.02	0.26	0.65	0.01	0.24	0.64	0.05

Notes: This table reports the results from regressing five-year aggregate market excess returns and average nominal bond excess return on the linear combination of forward rates. Data are annual from 1953 to 2010. For the data coefficients, I report t-statistics constructed using Newey-West standard errors. Population moments are calculated by simulating monthly data from the model for 60,000 years and then aggregating to an annual frequency. I also simulate 10,000 60-year samples and report the 5th-, 50th- and 95th-percentile for each statistic from the full set of simulations and for the subset of samples for which no disasters occur.

CHAPTER 2 : Rare Booms and Disasters in a Multi-sector Endowment Economy

(with Jessica A. Wachter)

2.1. Introduction

This paper introduces a representative agent asset pricing model in which the endowment and the aggregate dividend are subject to larges rare negative shocks (disasters) and large rare positive shocks (booms). We consider a two-sector model for the economy: the growth sector is the claim to the stream of dividends arising from the rare booms, while the value sector is the claim to the remaining dividend stream. The two sectors add up to the aggregate market. We show that this parsimonious model can explain important features of stock market data. As shown in earlier work, a time-varying probability of rare disasters can account for the high equity premium, high stock market volatility and return predictability exhibited by the aggregate market.[1] Beyond addressing these earlier points, our work also explains the cross-section of stock returns.

The possibility of rare booms has been little studied in comparison to rare disasters. This may be because the implications of rare booms for the equity premium, a focus of earlier work, are relatively minor. Because of decreasing marginal utility, the representative agent requires little compensation for bearing the risk of rare booms, even if they are large.[2] However, when assets have varying exposure to the booms, the impact on the cross-section can be substantial. The model implies that investors are willing to hold the growth portfolio

---

[1]For the equity premium result, see Rietz (1988), Longstaff and Piazzesi (2004), Veronesi (2004) and Barro (2006). For the volatility and predictability results, see Gabaix (2012), Gourio (2012) and Wachter (2012).

[2]An exception is the literature on technological innovations. Pastor and Veronesi (2009) show how the transition from idiosyncratic to systematic risk can explain time series patterns of returns in innovative firms around technological revolutions. In the present paper, we assume for simplicity that the risk of the technology is systematic from the start. Jovanovic and Rousseau (2003) show how technological revolutions can have long-lived effects, in that the firms that capitalize on such revolutions continue to have high market capitalization in a manner consistent with our model. These papers do not study the value premium. In recent work, Bekaert and Engstrom (2010) propose a model in which the economy is also subject to shocks in which bad events predominate and shocks in which good events predominate. Their model differs from ours in that they focus on explaining aggregate market and consumption moments with an agent with habit-like preferences.

despite its low return because of the small possibility of a high payout. The growth portfolio has a high covariance with the market because it is subject to a time-varying risk of booms as well as a time-varying risk of disaster; once a boom occurs the resulting dividend stream has the same disaster exposure as the rest of the economy. In fact, the model accurately predicts that the growth portfolio has a market beta greater than one while the value portfolio has a market beta less than one. This combination of high betas with low expected returns allows the model to explain the striking failure of the Capital Asset Pricing Model (CAPM) observed in the data (Fama and French (1992)).

Our model introduces several innovations beyond those described above. First, we model disasters and booms as influencing the drift rate of fundamentals, rather than fundamentals directly. This allows our model to capture the fact that disasters and booms unfold slowly, as emphasized by Constantinides (2008). The assumption of recursive utility implies that there is still a substantial equity premium.[3] Second, we introduce a novel way to model value and growth assets that allows the dividends on value to grow more slowly than those of the aggregate market, but still implies value and growth add up to the market, and price ratios are stationary.

A number of other papers also offer risk-based explanations for the relatively high expected returns on value stocks (the value premium).[4] It is likely that the value premium has multiple causes, and it is not the purpose of this article to rule out other explanations. One difficulty with these risk-based explanations is that a value premium arises because returns on the value portfolio are more risky than the growth portfolio. This, however, is not the case in the data. In our model, growth is in fact more risky. We break the link between risk

---

[3]Bansal et al. (2010) also model large shocks to the growth rate in a setting with a constant probability of disaster. Like the present paper, Nakamura et al. (2011) address the Constantinides (2008) critique; the focus of their empirical paper is to accurately capture the disaster distribution in complex setting where only numerical solutions are available. In contrast, the focus of this paper is to account for the aggregate market and cross-sectional moments using a relatively simple model with analytical solutions. Another strand of the literature incorporates non-normal shocks into the drift and volatility of the endowment process to model multifrequency or business-cycle fluctuations: see Calvet and Fisher (2007, 2008), Lettau et al. (2008) and Bhamra et al. (2010).

[4]For example, Ai and Kiku (2013), Berk et al. (1999), Carlson et al. (2004), Gârleanu et al. (2012), Gomes et al. (2003), Hansen et al. (2008), Novy-Marx (2010), Santos and Veronesi (2010) and Zhang (2005).

and return in two ways: first, while population returns on growth may be higher, in any given sample, it is not unlikely that a value premium will be observed in the data. Second, the risk in growth arises from rare booms, which occur in times of low marginal utility. Hence investors do not require compensation for bearing this risk.[5]

Besides addressing the sign and magnitude of the value premium, our model can also account for the time-series behavior of the value premium and its relation to the equity premium. As is well-known, the price-dividend ratio can predict excess returns on the aggregate market, implying that the equity premium is varying over time (Campbell and Shiller (1988)). The value spread can predict the return on the value-minus-growth portfolio, implying that it, too, has a time-varying risk premium (Cohen et al. (2003)). However, these risk premiums appear to have little to do with one-another; the price-dividend ratio has almost no predictive power for the value spread. In our model, a two-factor structure for risk premia arise naturally, and it is thus capable of explaining this result.

The remainder of the paper is organized as follows. Section 2.2 describes and solves the model. Section 2.3 discusses the quantitative fit of the model to the data. Section 2.4 concludes.

2.2. Model

2.2.1. Endowment and preferences

We assume an endowment economy with an infinitely-lived representative agent. Aggregate consumption (the endowment) follows a diffusion process with time-varying drift:

$$\frac{dC_t}{C_t} = \mu_{Ct}\, dt + \sigma dB_{Ct},\tag{2.1}$$

---

[5]Other studies succeed in breaking the link between risk and return using mechanisms other than what we consider here. These include Campbell and Vuolteenaho (2004) and Campbell et al. (2010), who model growth and value in an ICAPM setting, and Lettau and Wachter (2007), who assume an exogenous stochastic discount factor. These studies, however, do not assume a representative agent pricing assets in equilibrium in which cash flows must add up to the market.

where B_{Ct} is a standard Brownian motion. The drift of the consumption process is given by

$$\mu_{Ct} = \bar{\mu}_C + \mu_{1t} + \mu_{2t}, \tag{2.2}$$

where

$$d\mu_{jt} = -\kappa_{\mu_j}\mu_{jt}dt + Z_{jt}dN_{jt}, \tag{2.3}$$

for $j = 1, 2$. This model allows expected consumption growth to be subject to two types of (large) shocks. The rare events N_{jt} each follow a Poisson process (that is, for a given t, N_{jt} has a Poisson distribution). In what follows, we will consider the first type ($j = 1$) to be disasters, so that $Z_{1t} \leq 0$ and the second type ($j = 2$) to be booms, so that $Z_{2t} \geq 0$. When a disaster occurs, the process μ_{1t} jumps downward. It then mean-reverts back (absent any other bad shocks). Likewise, when a boom occurs, the process μ_{2t} jumps upward. It too reverts back. This model allows for smooth consumption (as in the data), that nonetheless goes through periods of extreme growth rates in one direction or another. Writing down two separate processes influencing expected consumption growth (as opposed to one process with two types of shocks) simplifies pricing of different sectors and allows disasters to be shorter-lived than booms, as the data suggest.

In what follows, the magnitude of the jumps will be random with a time-invariant distribution. That is, Z_{jt} has distribution ν_j. We will use the notation E_{ν_j} to denote expectations taken over the distribution ν_j. The intensity of the Poisson shock N_j is governed by λ_{jt}, which is stochastic, and follows the process

$$d\lambda_{jt} = \kappa_{\lambda_j}(\bar{\lambda}_j - \lambda_{jt})\,dt + \sigma_{\lambda_j}\sqrt{\lambda_{jt}}\,dB_{\lambda_j t}. \tag{2.4}$$

where $B_{\lambda_j t}, j = 1, 2$ are independent Brownian motions, that are each independent of B_{Ct}. Furthermore, we assume that the Poisson shocks N_{jt} are independent of each other, and of the Brownian motions. Define $\lambda_t = [\lambda_{1t}, \ \lambda_{2t}]^\top$, $\mu_t = [\mu_{1t}, \ \mu_{2t}]^\top$, $B_{\lambda t} = [B_{\lambda_1 t}, \ B_{\lambda_2 t}]^\top$ and

$B_t = [B_{Ct}, \ B_{\lambda t}^\top]^\top.$[6]

We assume the continuous-time analogue of the utility function defined by Epstein and Zin (1989) and Weil (1990), that generalizes power utility to allow for preferences over the timing of the resolution of uncertainty. The continuous-time version is formulated by Duffie and Epstein (1992); we use the case that sets the parameter associated with the elasticity of intertemporal substitution (EIS) equal to one. Define the utility function V_t for the representative agent using the following recursion:

$$V_t = E_t \int_t^\infty f(C_s, V_s) \, ds, \tag{2.5}$$

where

$$f(C_t, V_t) = \beta(1 - \gamma)V_t \left(\log C_t - \frac{1}{1 - \gamma} \log((1 - \gamma)V_t) \right). \tag{2.6}$$

We follow common practice in interpreting γ as risk aversion and β as the rate of time preference. We assume throughout that $\gamma > 0$ and $\beta > 0$.

2.2.2. The value function

Let W_t denote the wealth of the representative agent and $J(W_t, \mu_t, \lambda_t)$ the value function. In equilibrium, it must be the case that $J(W_t, \mu_t, \lambda_t) = V_t$. The following describes the value function and its properties. The proof of Theorem 2.1 is in Appendix A.2.2.

Theorem 2.1. *Assume parameter values satisfy Assumption A.2. Then the value function J takes the following form:*

$$J(W_t, \mu_t, \lambda_t) = \frac{W_t^{1-\gamma}}{1 - \gamma} I(\mu_t, \lambda_t), \tag{2.7}$$

where

$$I(\mu_t, \lambda_t) = \exp \left\{ a + b_\mu^\top \mu_t + b_\lambda^\top \lambda_t \right\}, \tag{2.8}$$

for vectors $b_\mu = [b_{\mu_1}, b_{\mu_2}]^\top$ and $b_\lambda = [b_{\lambda_1}, b_{\lambda_2}]^\top$. The coefficients a, b_{μ_j} and b_{λ_j} for $j = 1, 2$

---

[6]We assume throughout that $\kappa_\mu j, \kappa_{\lambda_j}, \bar{\lambda}_j$ and σ_{λ_j}, for $j = 1, 2$, are strictly positive.

take the following form:

$$a = \frac{1-\gamma}{\beta}\left(\bar{\mu}_C - \frac{1}{2}\gamma\sigma^2\right) + (1-\gamma)\log\beta + \frac{1}{\beta}b_\lambda^\top(\kappa_\lambda * \bar{\lambda}) \tag{2.9}$$

$$b_{\mu_j} = \frac{1-\gamma}{\kappa_{\mu_j} + \beta}, \tag{2.10}$$

$$b_{\lambda_j} = \frac{1}{\sigma_{\lambda_j}^2}\left(\beta + \kappa_{\lambda_j} - \sqrt{(\beta + \kappa_{\lambda_j})^2 - 2E_{\nu_j}\left[e^{b_{\mu_j}Z_{jt}} - 1\right]\sigma_{\lambda_j}^2}\right). \tag{2.11}$$

Here and in what follows, we use the notation $*$ to denote element-by-element multiplication of vectors of equal dimension.

As the next corollary shows, an investor is made better off (as measured by the value function), by an increase in the components of expected consumption growth or by an increase in the probability of a boom. The investor is made worse off by an increase in the probability of disaster.

Corollary 2.1. *The value function is increasing in μ_{jt} for $j = 1, 2$, decreasing in λ_{1t}, and increasing in λ_{2t}.*

Proof To fix ideas, consider $\gamma > 1$. It suffices to show $b_{\lambda_1} > 0$, $b_{\lambda_2} < 0$, and $b_{\mu_j} < 0$ for $j = 1, 2$. It follows immediately from (2.10) that $b_{\mu_j} < 0$. Because $Z_1 < 0$ and $b_{\mu_1} < 0$, $E_{\nu_1}\left[e^{b_{\mu_1}Z_{1t}} - 1\right] > 0$. Therefore,

$$\sqrt{(\beta + \kappa_{\lambda_1})^2 - 2E_{\nu_1}\left[e^{b_{\mu_1}Z_{1t}} - 1\right]\sigma_{\lambda_1}^2} < \beta + \kappa_{\lambda_1}.$$

It follows that $b_{\lambda_1} > 0$. Because $Z_2 > 0$ and $b_{\mu_2} < 0$, $E_{\nu_2}\left[e^{b_{\mu_2}Z_{2t}} - 1\right] < 0$. Therefore,

$$\sqrt{(\beta + \kappa_{\lambda_2})^2 - 2E_{\nu_2}\left[e^{b_{\mu_2}Z_{2t}} - 1\right]\sigma_{\lambda_2}^2} > \beta + \kappa_{\lambda_2}$$

and $b_{\lambda_2} < 0$. □

The riskfree rate takes a particularly simple form:

Corollary 2.2. *Let r_t denote the instantaneous risk-free rate in this economy, then r_t is*

given by

$$r_t = \beta + \mu_{Ct} - \gamma\sigma^2. \tag{2.12}$$

2.2.3. The aggregate market

Let D_t denote the dividend on the aggregate market. Assume that dividends follow the process

$$\frac{dD_t}{D_t} = \mu_{Dt}\, dt + \phi\sigma\, dB_{Ct}, \tag{2.13}$$

where

$$\mu_{Dt} \;=\; \bar{\mu}_D + \phi\mu_{1t} + \phi\mu_{2t}.$$

This structure allows dividends to respond by a greater amount than consumption to booms and disasters (this is consistent with the U.S. experience, as shown in Longstaff and Piazzesi (2004)). For parsimony, we assume that the parameter, namely, ϕ, governs the dividend response to normal shocks, booms and disasters. This ϕ is analogous to leverage in the model of Abel (1999), and we will refer to it as leverage in what follows.

Prices

We price equity claims using no-arbitrage and the state-price density. Duffie and Skiadas (1994) show that the state-price density π_t equals

$$\pi_t = \exp\left\{\int_0^t \frac{\partial}{\partial V} f\left(C_s, V_s\right) ds\right\} \frac{\partial}{\partial C} f\left(C_t, V_t\right). \tag{2.14}$$

Let $H\left(D_t, \mu_t, \lambda_t, \tau\right)$ denote the time t price of a single future dividend payment at time $t + \tau$. Then

$$H(D_t, \mu_t, \lambda_t, s - t) = E_t\left[\frac{\pi_s}{\pi_t} D_s\right].$$

The following corollary gives the solution for H up to ordinary differential equations. This

corollary is a special case of Theorem A.3, given in Appendix A.2.2.

Corollary 2.3. *The solution for the function H is as follows*

$$H(D_t, \mu_t, \lambda_t, \tau) = D_t \exp \left\{ a_\phi(\tau) + b_{\phi\mu}(\tau)^\top \mu_t + b_{\phi\lambda}(\tau)^\top \lambda_t \right\}, \tag{2.15}$$

where $b_{\phi\mu}(\tau) = [b_{\phi\mu_1}(\tau), \ b_{\phi\mu_2}(\tau)]^\top$ and $b_{\phi\lambda}(\tau) = [b_{\phi\lambda_1}(\tau), \ b_{\phi\lambda_2}(\tau)]^\top$. Furthermore, for $j = 1, 2$,

$$b_{\phi\mu_j}(\tau) = \frac{\phi - 1}{\kappa_{\mu_j}} \left(1 - e^{-\kappa_{\mu_j}\tau} \right), \tag{2.16}$$

while $b_{\phi\lambda_j}(\tau)$ (for $j = 1, 2$) and $a_\phi(\tau)$ satisfy the following:

$$\frac{db_{\phi\lambda_j}}{d\tau} = \frac{1}{2}\sigma_{\lambda_j}^2 b_{\phi\lambda_j}(\tau)^2 + \left(b_{\lambda_j}\sigma_{\lambda_j}^2 - \kappa_{\lambda_j} \right) b_{\phi\lambda_j}(\tau) + E_{\nu_j} \left[e^{b_{\mu_j} Z_{jt}} \left(e^{b_{\phi\mu_j}(\tau)Z_{jt}} - 1 \right) \right] \tag{2.17}$$

$$\frac{da_\phi}{d\tau} = \bar{\mu}_D - \bar{\mu}_C - \beta + \gamma\sigma^2 (1 - \phi) + b_{\phi\lambda}(\tau)^\top \left(\kappa_\lambda * \bar{\lambda} \right) \tag{2.18}$$

with boundary conditions $b_{\phi\lambda_j}(0) = a_\phi(0) = 0$.

Let $F(D_t, \mu_t, \lambda_t)$ denote the value of the market portfolio (namely, the price of the claim to the entire future dividend stream). Then

$$F(D_t, \mu_t, \lambda_t) = \int_0^\infty H(D_t, \mu_t, \lambda_t, \tau) \, d\tau.$$

Corollary 2.3 implies that the price-dividend ratio, which we will denote by a function G, can be written as

$$G(\mu_t, \lambda_t) = \int_0^\infty \exp \left(a_\phi(\tau) + b_{\phi\mu}(\tau)^\top \mu_t + b_{\phi\lambda}(\tau)^\top \lambda_t \right) d\tau. \tag{2.19}$$

The expressions in Corollary 2.3 show how prices respond to innovations in expected consumption growth and in changing disaster probabilities. Because $\phi > 1$, (2.16) shows that innovations to expected consumption growth increase the price-dividend ratio. The presence of the $\phi - 1$ term shows that this is a trade-off between the effect of expected consumption

64

growth on the riskfree rate and on dividend cash flows. In our recursive utility model, the cash flow effect dominates and asset prices fall during disasters and rise during booms. Moreover, the more persistent is the process for the mean (the lower is κ_{μ_j}), the greater is the effect of a change in μ_{jt} on prices.[7] Finally, an increase in the probability of a disaster lowers the price-dividend ratio, while an increase in the probability of a boom raises it. These effects are summarized in the following corollary.

Corollary 2.4. *The price-dividend ratio $G(\mu_t, \lambda_t)$ is increasing in the components of expected consumption growth μ_{jt} (for $j = 1, 2$), decreasing in the probability of a disaster λ_{1t} and increasing in the probability of a boom λ_{2t}.*

The fact that $G(\mu_t, \lambda_t)$ is increasing in μ_{jt} follows immediately from the form of (2.16). The results for λ_{1t} and λ_{2t} are less obvious. We give a full proof in Appendix A.2.2 and discuss the intuition here. Consider the ODE (2.17). The functions $b_{\phi\lambda_j}(\tau)$ would be identically zero without the last term $E_{\nu_j}\left[e^{b_{\mu_j} Z_{jt}} \left(e^{b_{\phi\mu_j}(\tau)Z_{jt}} - 1 \right)\right]$. It is this term that determines the sign of $b_{\phi\lambda_j}(\tau)$, and thus how prices respond to changes in probabilities.

To fix ideas, consider disasters ($j = 1$). The last term in (2.17) can itself be written as a sum of two terms:

$$
E_{\nu_1}\left[e^{b_{\mu_1} Z_{1t}} \left(e^{b_{\phi\mu_1}(\tau)Z_{1t}} - 1 \right)\right] =
$$
$$
-\underbrace{E_{\nu_1}\left[\left(e^{b_{\mu_1} Z_{1t}} - 1 \right)\left(1 - e^{b_{\phi\mu_1}(\tau)Z_{1t}} \right)\right]}_{\text{Risk premium effect}} \quad + \quad \underbrace{E_{\nu_1}\left[e^{b_{\phi\mu_1}(\tau)Z_{1t}} - 1 \right]}_{\text{Cash flow and riskfree rate effect}} \quad (2.20)
$$

The first of the terms in (2.20) is one component of the equity premium, indeed it is what we will refer to as the static disaster premium, terminology that we discuss in more detail in the next section.[8] When the risk of a disaster increases, the static equity premium increases. Because an increase in the discount rate lowers the price-dividend ratio, this term appears in (2.20) with a negative sign. The second term in (2.20) is the expected price response in the

---

[7]The derivative of (2.16) with respect to κ_{μ_j} equals $(\kappa_{\mu_j}\tau + 1)e^{-\kappa_{\mu_j}\tau} - 1$ which is negative, because $e^{\kappa_{\mu_j}\tau} > \kappa_{\mu_j}\tau + 1$.

[8]More precisely, this is the static disaster premium for zero-coupon equity with maturity τ.

event of a disaster.[9] It represents the combined effect of the disaster on cash flows and on the riskfree rate. The net effect is negative, as described above. Thus the response of equity values to changes in the probability of a disaster is determined by a risk premium effect, and a (joint) cash flow and riskfree rate effect. Both effects turn out to be negative; our calibration implies that they are roughly of equal magnitude (the full risk premium however is of much greater magnitude since it also includes compensation for time-varying λ_{1t}). A similar structure holds for booms. However, in the case of booms, the joint riskfree-rate and cash flow effect is positive, and it dominates the risk premium effect.[10]

The equity premium

Here, we give an expression for the instantaneous equity premium and discuss its properties. This will be useful in understanding the quantitative results in Section 2.3.

First, we define the *jump operator*, which denotes how a process responds to an occurrence of a rare event. Namely, let X_t be any pure diffusion process (X_t can be a vector), and let μ_{jt}, $j = 1, 2$ be defined as above. Consider a scalar, real-valued function $h(\mu_{1t}, \mu_{2t}, X_t)$. Define the jump operator $\mathcal{J}$ as follows:

$$\mathcal{J}_1(h(\mu_{1t}, \mu_{2t}, X_t)) = h(\mu_1 + Z_1, \mu_2, X_t)$$

$$\mathcal{J}_2(h(\mu_{1t}, \mu_{2t}, X_t)) = h(\mu_1, \mu_2 + Z_2, X_t).$$

Further, define

$$\bar{\mathcal{J}}_j(h(\mu_{1t}, \mu_{2t}, X_t)) = E_{\nu_j} \mathcal{J}_j(h(\mu_{1t}, \mu_{2t}, X))$$

for $j = 1, 2$, and

$$\bar{\mathcal{J}}(h(\mu_{1t}, \mu_{2t}, X_t)) = \left[\bar{\mathcal{J}}_1(h(\mu_{1t}, \mu_{2t}, X_t)), \ \bar{\mathcal{J}}_2(h(\mu_{1t}, \mu_{2t}, X_t)) \right]^{\top}.$$

---

[9]Again, more precisely, it is the price response of zero-coupon equity with maturity τ.

[10]The relative magnitude of these terms can be seen by comparing the risk premiums with the observed expected returns in samples when no jumps occur (namely Figures 2.5 and 2.7 with Figures 2.9 and 2.10). The term on the left hand side of (2.20) corresponds to the observed static premium in no-jump samples while the first term on the right hand side corresponds to the static premium in population.

Using Ito's Lemma and the definition above, we can write the process for the aggregate stock price $F_t = F(D_t, \mu_t, \lambda_t)$ as follows:

$$\frac{dF_t}{F_{t^-}} = \mu_{F,t}\, dt + \sigma_{F,t}\, dB_t + \sum_j \frac{\mathcal{J}_j(F_t)}{F_{t^-}}\, dN_{jt}.$$

The instantaneous expected return is the expected change in price, plus the dividend yield:

$$r_t^m = \mu_{F,t} + \frac{D_t}{F_t} + \frac{1}{F_t}\lambda_t^\top \bar{\mathcal{J}}(F_t). \tag{2.21}$$

Corollary 2.5. *The equity premium relative to the risk-free rate r is*

$$r_t^m - r_t = \phi\gamma\sigma^2 \underbrace{-\sum_j \lambda_{jt} E_{\nu_j}\left[\left(e^{b_{\mu_j} Z_{jt}} - 1\right)\frac{\mathcal{J}_j(G_t)}{G_t}\right]}_{\text{static rare event premium}} \underbrace{-\sum_j \lambda_{jt}\frac{1}{G_t}\frac{\partial G}{\partial \lambda_j}b_{\lambda_j}\sigma_{\lambda j}^2}_{\lambda\text{-premium}}. \tag{2.22}$$

As Corollary 2.5 shows, the equity premium is the sum of three terms. The first is the standard term arising from the consumption Capital Asset Pricing Model (CCAPM) of Breenden (1979). The second term is the premium directly attributable to rare events. It arises from the co-movement in prices and in marginal utility when one of these events occurs. We will call this term the *static rare event premium* (we include the negative sign in the definition of the premium). This term can itself be divided into the *static disaster premium* and the *static boom premium*:

$$\text{static disaster premium:} \quad -\lambda_{1t} E_{\nu_1}\left[\left(e^{b_{\mu_1} Z_{1t}} - 1\right)\frac{\mathcal{J}_1(G_t)}{G_t}\right]$$
$$\text{static boom premium:} \quad -\lambda_{2t} E_{\nu_2}\left[\left(e^{b_{\mu_2} Z_{2t}} - 1\right)\frac{\mathcal{J}_2(G_t)}{G_t}\right]$$

If a rare event occurs, instantaneous current dividends do not change, but future dividends do. This is why the formulas above contain the price dividend ratio G_t (it would also be correct to substitute G_t with F_t). Note that this is the premium that would obtain if the probability of the rare event λ_{jt} were constant. It is for this reason that we refer to these

terms as the static rare event premium.[11]

Finally, the third term in (2.22) represents the compensation the investor requires for bearing the risk of changes in the rare event probabilities (again, the definition should be viewed as including the negative sign). Accordingly, we call this the λ-*premium*. This term can also be divided into the compensation for time-varying disaster probability (the λ_1-premium) and compensation for time-varying boom probability (the λ_2-premium). Note that under power utility, only the CCAPM term would appear in the risk premium. This is because, in the power utility model, only the instantaneous co-movement with consumption matters for risk premia, not changes to the consumption distribution.

We next address the question of how these various terms contribute to the equity premium. The following corollary describes the signs of these terms:

Corollary 2.6.   *1. The static disaster and boom premiums are positive.*

2. *The λ_1-premium (the premium for time-varying disaster probability) is positive. The λ_2-premium (the premium for time-varying boom probability) is also positive.*

Proof To show the first statement, recall that $b_{\mu_j} < 0$ for $j = 1, 2$ (Corollary 2.1). First consider disasters ($j = 1$). Note $Z_1 < 0$, so $e^{b_{\mu_1} Z_{1t}} - 1 > 0$. Furthermore, because G is increasing in μ_1 (Corollary 2.4), $\mathcal{J}_1(G_t) < 0$. It follows that the static disaster premium is positive. Now consider booms ($j = 2$). Because $Z_2 > 0$, $e^{b_{\mu_2} Z_{2t}} - 1 < 0$. Because G is increasing in μ_2, $\mathcal{J}_2(G_t) > 0$. Therefore the static boom premium is also positive.

To show the second statement, first consider disasters ($j = 1$). Recall that $b_{\lambda_1} > 0$ (Corollary 2.1). Further, $\partial G / \partial \lambda_1 < 0$ (Corollary 2.4). For booms ($j = 2$), each of these quantities takes the opposite sign. The result follows. $\square$

The intuitive content of Corollary 2.6 is that both booms and disasters increase the risk of equities for the representative agent. They do so both because of the direct (static) effect

---

[11]However, the term "static premium" is somewhat of a misnomer here, since even the direct effect of rare events on the price-dividend ratio is a dynamic one.

stemming from happens to equities in these events, and because of an indirect (dynamic) effect, due to what happens to equities (as a result of rational forecasts of what *would* happen in these events) during normal times.

It is also useful to consider the return the econometrician would observe in an sample without rare events. We will distinguish these expected returns using the subscript nj ("no jump"). This expected return is simply given by the drift rate in the price, plus the dividend yield

$$r_{\text{nj},t}^m = \mu_{F,t} + \frac{D_t}{F_t}.$$

Based on this definition, the fact that $\frac{\bar{\mathcal{J}}(F_t)}{F_t} = \frac{\bar{\mathcal{J}}(G_t)}{G_t}$ and on Corollary 2.5, these expected returns can be calculated as follows:

Corollary 2.7. *The observed expected excess return in a sample without jumps is*

$$r_{\text{nj},t}^m - r_t = \phi\gamma\sigma^2 - \sum_j \lambda_{jt} E_{\nu_j} \left[e^{b_{\mu_j} Z_{jt}} \frac{\mathcal{J}_j(G_t)}{G_t} \right] - \sum_j \lambda_{jt} \frac{1}{G_t} \frac{\partial G}{\partial \lambda_j} b_{\lambda_j} \sigma_{\lambda_j}^2 \qquad (2.23)$$

This expression differs from (2.22) in that the contribution directly due to rare events is equal to $-\sum_j \lambda_{jt} E_{\nu_j} \left[e^{b_{\mu_j} Z_{jt}} \frac{\mathcal{J}_j(G_t)}{G_t} \right]$ as opposed to $-\sum_j \lambda_{jt} E_{\nu_j} \left[\left(e^{b_{\mu_j} Z_{jt}} - 1 \right) \frac{\mathcal{J}_j(G_t)}{G_t} \right]$. We will refer to the $j = 1$ term as the observed static disaster premium in a sample without jumps and the $j = 2$ term as the observed static boom premium in a sample without jumps.

Corollary 2.8. *The observed static disaster premium in a sample without jumps is positive. The observed static boom premium in a sample without jumps is negative.*

Proof The result follows from the fact that G is increasing in μ_1 and μ_2, and hence $\mathcal{J}_1(G) < 0$ and $\mathcal{J}_2(G) > 0$. $\qquad\qquad\square$

Note that the observed disaster premium is positive, just like the true disaster premium. However, the observed boom premium is negative, the oppose sign to the true boom premium.[12]

---

[12]We refer to these as the observed premiums to distinguish them from the true risk premiums (note that, unlike true risk premiums, they do not in fact represent a return for risk). In practice, it will be nearly

2.2.4. Growth and value sectors

The value sector is defined as the claim to cash flows that are not subject to the positive jumps, but are otherwise identical to those of the market. We will use the superscript v to denote processes related to the value sector and the subscript g to denote processes related to the growth sector. The dividend process for the value sector is as follows:

$$\frac{dD_{t,s}^v}{D_{t,s}^v} = \mu_{Ds}^v ds + \phi\sigma dB_{Cs}, \tag{2.24}$$

where $\mu_{Dt}^v = \bar{\mu}_D + \phi\mu_{1t}$, and with the boundary condition $D_{t,t}^v = D_t$. The price of the value sector claim can be determined in the same way as the price of the claim to the aggregate market (see Corollary 2.9 below).

The growth sector is defined as the residual. Let $D_{t,s}^g = D_s - D_{t,s}^v$. Define $F_{t,s}^g$ to be the price of the growth claim. Then, by the absence of arbitrage,

$$F_{t,s}^g = F_s - F_{t,s}^v.$$

As long as there are no positive jumps, the dividend on the value claim and the aggregate market are identical. However, when a positive jump takes place, the market dividend begins to diverge permanently from the value dividend. The dividend on the value sector will henceforth grow at a lower rate than the aggregate dividend, with the dividend on the growth claim comprising the difference.

In this setting, thinking of the value and the growth claim as long-lived assets would imply a value claim that makes up a vanishingly small portion of the aggregate market as time passes. The asset pricing implications of defining the value claim in this way would not be very interesting. Therefore, we do not think of the value claim as being a long-lived asset (indeed, because markets are complete, the actual assets that are specified do not affect the

---

impossible to distinguish the separate terms in (2.23). The terminology "observed static disaster premium" and "observed static boom premium" is used for convenience, not to suggest that these terms can in fact be observed separately from other parts of the expected excess return.

equilibrium). If one wishes to think of long-lived assets, the following interpretation may be helpful (though note that given that the value and growth claim are priced by no-arbitrage, this interpretation is not necessary): Every time there is a positive jump, the growth sector is disbanded. Some of the capital is used to start a new growth sector, and some goes into the rest of the economy. The value of the claims to the new growth and value sectors are adjusted so that the owners of the previous growth sector still receive the value of the claim to the (previous) growth dividends. In effect, the owners of the growth sector are diluting the owners of the value sector in the event of a positive jump.

Prices

Let $H^v\left(D^v_{t,s}, \mu_s, \lambda_s, \tau\right)$ denote the time t price of a single future value sector dividend payment at time $s + \tau$. Recall that π_t is the state-price density, defined in (2.14). As in the case of the aggregate market,

$$H^v(D^v_{t,s}, \mu_s, \lambda_s, u - s) = E_s\left[\frac{\pi_u}{\pi_s} D^v_{t,u}\right].$$

Furthermore,

$$F^v\left(D^v_{t,s}, \mu_s, \lambda_s\right) = \int_0^\infty H^v\left(D^v_{t,s}, \mu_s, \lambda_s, \tau\right) d\tau. \tag{2.25}$$

The following corollary is a special case of Theorem A.3, given in Appendix A.2.2.

Corollary 2.9. *The solution for the function H^v is as follows:*

$$H^v\left(D^v_{t,s}, \mu_s, \lambda_s, \tau\right) = D^v_{t,s} \exp\left\{a^v_\phi(\tau) + b^v_{\phi\mu}(\tau)^\top \mu_s + b^v_{\phi\lambda}(\tau)^\top \lambda_s\right\},$$

where $b^v_{\phi\mu}(\tau) = [b^v_{\phi\mu_1}(\tau),\ b^v_{\phi\mu_2}(\tau)]^\top$ and $b^v_{\phi\lambda}(\tau) = [b^v_{\phi\lambda_1}(\tau),\ b^v_{\phi\lambda_2}(\tau)]^\top$. Furthermore,

$$b^v_{\phi\mu_1}(\tau) = \frac{\phi - 1}{\kappa_{\mu_1}}\left(1 - e^{-\kappa_{\mu_1}\tau}\right) \tag{2.26}$$

$$b^v_{\phi\mu_2}(\tau) = -\frac{1}{\kappa_{\mu_2}}\left(1 - e^{-\kappa_\mu 2\tau}\right), \tag{2.27}$$

while $b^v_{\phi\lambda_j}(\tau)$ (for $j = 1, 2$) and $a_\phi(\tau)$ satisfy

$$\frac{db^v_{\phi\lambda_j}}{d\tau} = \frac{1}{2}\sigma^2_{\lambda_j} b^v_{\phi\lambda_j}(\tau)^2 + \left(b_{\lambda_j}\sigma^2_{\lambda_j} - \kappa_{\lambda_j}\right) b^v_{\phi\lambda_j}(\tau) + E_{\nu_j}\left[e^{b_{\mu_j}Z_{jt}}\left(e^{b^v_{\phi\mu_j}(\tau)Z_{jt}} - 1\right)\right], \tag{2.28}$$

$$\frac{da^v_\phi}{d\tau} = \bar{\mu}_D - \bar{\mu}_C - \beta + \gamma\sigma^2(1 - \phi) + b^v_{\phi\lambda_j}(\tau)^\top(\kappa_\lambda * \bar{\lambda}) \tag{2.29}$$

with boundary conditions $b_{\phi\lambda_j}(0) = a_\phi(0) = 0$.

It follows from (2.25) and Corollary 2.9 that the price-dividend ratio on the value sector is

$$G^v(\mu_t, \lambda_t) = \int_0^\infty \exp\left(a^v_\phi(\tau) + b^v_{\phi\mu}(\tau)^\top\mu_t + b^v_{\phi\lambda}(\tau)^\top\lambda_t\right) d\tau. \tag{2.30}$$

The dynamics of this price-dividend ratio are given by the following:

Corollary 2.10. *The price-dividend ratio for the value claim $G^v(\mu_t, \lambda_t)$ is increasing in μ_{1t}, decreasing in μ_{2t}, and decreasing in the probability of a rare event λ_{jt}, for $j = 1, 2$.*

Though the dividends on the value sector are not exposed to positive jumps, the value sector still depends on μ_{2t} and therefore on λ_{2t} because of the effect of μ_{2t} on the riskfree rate.

Risk premia

Risk premia on the value claim can be derived similarly to those on the aggregate market. As we will see, however, they behave quite differently.[13]

Corollary 2.11. *The value sector premium relative to the risk-free rate r is*

$$r^v_t - r_t = \phi\gamma\sigma^2 - \sum_j \lambda_{jt}E_{\nu_j}\left[\left(e^{b_{\mu_j}Z_{jt}} - 1\right)\frac{\mathcal{J}_j(G^v_t)}{G^v_t}\right] - \sum_j \lambda_{jt}\frac{1}{G^v_t}\frac{\partial G^v}{\partial\lambda_j}b_{\lambda_j}\sigma^2_{\lambda j} \tag{2.31}$$

The three terms in (2.31) have an analogous interpretation to those for the market premium, and can also be signed.

---

[13]The proofs of these results are directly analogous to those for the market, and therefore we do not repeat them.

Corollary 2.12.   *1. The static disaster premium for the value sector is positive.*

2. *The static boom premium for the value sector is negative.*

3. *The λ_1-premium on the value sector is positive.*

4. *The λ_2-premium on the value sector is negative.*

Finally, the following corollary characterizes the observed expected return in a sample without jumps

Corollary 2.13. *The observed expected excess return on the value sector in a sample without jumps is*

$$r_{\mathrm{nj},t}^{v} - r_t = \phi\gamma\sigma^2 - \sum_j \lambda_{jt} E_{\nu_1}\left[e^{b_{\mu_j} Z_{jt}} \frac{\mathcal{J}(G_t^v)}{G_t^v} \right] - \sum_j \lambda_{jt} \frac{1}{G_t^v} \frac{\partial G^v}{\partial \lambda_j} b_{\lambda_j} \sigma_{\lambda j}^2 \qquad (2.32)$$

Both the terms corresponding to disaster and boom risk in this expression are positive. As in the case of the aggregate market, the sign of the disaster component is the same as in the risk premium, while the sign of the boom component is reversed.

Corollary 2.14. *In a sample without jumps, the observed disaster and boom premiums for the value sector are positive.*

The corollaries in this section state that the premiums related to disaster risk (the static disaster premium and the λ_1-premium) are positive for the value sector, just as they are for the aggregate market. The premiums related to boom risk (the static boom premium and the λ_2-premium) are negative for the value sector, though they are positive for the aggregate market. In population, the expected returns on the value sector will therefore be lower than those on the aggregate market. In a sample without jumps, however, this effect may be (and, for reasonable parameter values, will be) reversed. The reason is that the static boom premium switches signs: in a sample without booms, it is negative for the aggregate market, but positive for the value sector. This will produce an observed value premium.

73

2.3. Quantitative results

2.3.1. Calibration

Data

To calibrate the rare events, we use international consumption data described in detail in Barro and Ursua (2008), and updated by Barro and Ursua to include data on 43 countries. These data contain annual observations on real, per capita consumption; start dates vary from early in the 19th century to the middle of the 20th century.

Our aggregate market data come from CRSP. We define the market return to be the gross return on the value-weighted CRSP index. Dividend growth is computed from the dividends on this index. The price-dividend ratio is price divided by the previous 12 months of dividends to remove the effect of seasonality in dividend payments (in computing this dividend stream, we assume that dividends on the market are not reinvested). We compute market returns and dividend growth in real terms by adjusting for inflation using changes in the consumer price index (also available from CRSP). For the government bill rate, we use real returns on the 3-month Treasury Bill. We also use real, per capital expenditures on non-durables and services for the U.S., available from the Bureau of Economic Analysis. These data are annual, begin in 1947, and end in 2010. Focusing on post-war data allows for a clean comparison between U.S. data and hypothetical samples in which no rare events take place.

Data on value and growth portfolio are from Ken French's website. CRSP stocks are sorted annually into deciles based on their book-to-market ratios. Our growth claim is an extreme example of a growth stock; it is purely a claim to positive extreme events and nothing else. In the data, it is more likely that growth stocks are a combination of this claim and the value claim. To avoid modeling complicated share dynamics, we identify the growth claim with the decile that has the lowest book-to-market ratio, while the value claim consists of a

portfolio (with weights defined by market equity) of the remaining nine deciles. A standard definition of the value spread is the log book-to-market ratio of the value portfolio minus the log book-to-market ratio of the growth portfolio (Cohen et al. (2003)). In our endowment economy, book value can be thought of as the dividend. However, the dividend on the growth claim is identically equal to zero (though of course this claim has future non-zero dividends), and for this reason, there is no direct analogue of the value spread. We therefore compute the value spread in the model as the log dividend-price ratio on the value portfolio minus the log dividend-price ratio on the aggregate market. For comparability, we compute the same quantity in the data. Where our non-standard definition might be an issue is our predictability results; we have checked that these results are robust to the more standard data definition.

Parameter values

We report parameter values in Table 2.1. Average consumption growth and the volatility of consumption growth equal their post-war averages over a set of developed countries as in Barro (2006). These are both about 2%. We calibrate dividend growth to be slightly higher: 3.55%. Given the construction of CRSP dividends, there no reason to assume that dividends and consumption should grow at the same rate. Indeed, CRSP dividends do not include repurchases; presumably these imply that dividends are likely to be higher some time in the future, and that the sample mean is not a good indicator of the true mean. For this reason, we choose the mean of the dividend growth distribution that is implied by the level of the price-dividend ratio in the data.

Leverage, ϕ, is chosen to be 3.5. This implies that the volatility of log dividends is 3.5 times that of log consumption. In our data, the ratio is 4.66. However, this value would most likely imply too great a response of dividends to consumption disasters; we therefore choose a smaller and more conservative value. We choose a low rate of time preference to obtain a realistic government bill rate.[14] Relative risk aversion is equal to 3.

---

[14]Further lowering this value leads there to be no solution to the investor's optimization problem.

The average probability of a disaster is chosen to be 2.86%, which is the value calibrated by Barro and Ursua (2008) for OECD countries.[15] The persistence in the price-dividend ratio is nearly entirely determined by the persistence in the disaster probability. We therefore choose a low rate of mean reversion: $\kappa_\lambda 1 = 0.11$. With this choice, the median small-sample value of the persistence of the price-dividend ratio is 0.78; the value in the data is 0.92. This suggests the possibility of lowering $\kappa_\lambda 1$ still further (which would increase the effect of disaster risk on the equity premium and volatility); however, insisting that the model fit the very large degree of persistence in the data greatly widens the parameter range at which the value function fails to exist. The volatility σ_{λ_1} is chosen to be 9.4%, which leads to a realistic volatility for the aggregate market.

The disaster distribution, and the mean reversion in the disaster component of the expected consumption growth ($\kappa_\mu 1$) are chosen to fit the distribution of consumption declines, reported in Table 2.2 and the left panels Figures 2.1 and 2.2. These results suggest that the consumption growth reverts to its normal level relatively quickly, suggesting a high value for $\kappa_\mu 1$ (we choose 1.0). To calculate the size of the jumps, we assume a power law distribution (see Gabaix (2009) for a discussion of the properties of power law distributions). Following Barro and Ursua (2008), we consider 10% as the smallest magnitude of the disaster. Our calibration procedure suggests a power law parameter of 7 (the lower this parameter, the heavier the tail of the power law). Barro and Jin (2011) find similar results using maximum likelihood.[16] Table 2.2 also reports the distribution of declines in a model in which all the decline takes place immediately. This model fits the data less well, substantially over-predicting the number of large declines at the one-year horizon.

We follow a similar strategy for booms (data for large positive consumption events are reported in Table 2.3 and the right panels of Figures 2.1 and 2.2). Parameter values for the

---

[15]We calibrate the size of the disasters to the full set of samples and the average probability to the OECD subsample. In both cases, we are choosing the more conservative measure, because the OECD sub-sample has rarer, but more severe disasters.

[16]Barro and Jin (2011) estimate that the power law parameter is 6.86%. They also argue that the distribution is better characterized by a double power law, with a lower exponent for larger disasters. In this sense our choice of a single power with a coefficient of 7 is conservative.

λ_{2t} process are chosen to fit the mean, volatility, and mean reversion in the value spread as reported in Table 2.7. Booms in the data do not seem to be as heavy-tailed as disasters, but they die out somewhat more slowly. We choose a minimum value of 5%, a mean reversion coefficient of 0.60, and a power law parameter of 20. Our results are not sensitive to the precise choices of these parameter values.

2.3.2. Prices and expected returns as functions of the state variables

Prices

Figures 2.3 and 2.4 show terms in the expressions for the price-dividend ratio on the market (2.19) and the corresponding quantity for the value claim (2.30). These expressions are an integral of exponential-linear terms. Each of these terms can be interpreted as the ratio of the price of a zero-coupon equity claim to the current dividend. The integral is over τ, which can be interpreted as the maturity of these claims. Figure 2.3 shows the functions $b_{\phi\mu_j}(\tau)$ and $b^v_{\phi\mu_j}(\tau)$ as a function of τ, and Figure 2.4 does the same for the functions $b_{\phi\lambda_j}(\tau)$ and $b^v_{\phi\lambda_j}(\tau)$. The persistence of the state variables, combined with the effect of the duration of the claims implies that the magnitude of these functions is increasing in τ, as the figures show.

We first discuss the effect of variation in the mean of consumption on the price-dividend ratios. It is useful to discuss this first, because the effect of μ on the price is ultimately what determines the effect of λ. Note that both $b_{\phi\mu_1}(\tau)$ and $b_{\phi\mu_2}(\tau)$ are positive, reflecting the fact that the market is exposed to both positive and negative jumps in dividend growth. Greater average dividend growth, whether it arises from the absence of a disaster or the presence of a boom, increases the price-dividend ratio. Both terms converge to their limits in a relatively short time, reflecting the fact that neither booms nor disasters are highly persistent in the model. The fact that $b_{\phi\mu_2}(\tau)$ takes longer to converge reflects the greater persistence of booms than disasters, as does the fact that $b_{\phi\mu_2}(\tau)$ is larger in magnitude that $b_{\phi\mu_1(\tau)}$ (because in fact the distribution of immediate responses is larger for disasters

than for booms).

The response of the value claim to disasters, reflected in $b_{\phi\mu_1}^v(\tau)$ is nearly the same as that of the market as a whole. However, the response to booms is quite different. The reason, of course, is that the cash flows on the value claim are not exposed to booms. Indeed, the price of the value claim is decreasing in μ_{2t} because of the effect of μ_{2t} in interest rates. As explained above, the price response to μ_{jt} is determined by the tradeoff between the cash flow effect and the interest rate effect. Because $\phi > 1$, the cash-flow effect dominates for both types of shocks for the aggregate market. For the value claim, the cash-flow effect dominates for μ_{1t}. However, there is no cash flow effect for μ_{2t} on the value claim (this can be seen by comparing Equation 2.26 with 2.27; the first of these terms has a ϕ while the second does not). Thus the riskfree rate effect implies than an increase in expected consumption growth arising from booms decreases the price of the value claim.

Figure 2.4 shows the functions $b_{\phi\lambda_j}(\tau)$ (which multiply λ_{jt} in the expression for the market price-dividend ratio) and $b_{\phi\lambda_j(\tau)}^v$ (which multiply λ_{jt} in the expression for the price-dividend ratio on the value claim). $b_{\phi\lambda_1}(\tau)$ and $b_{\phi\lambda_1}^v(\tau)$ are negative, implying that an increase in the probability of a disaster lowers prices. These coefficients are similar, though slightly greater in magnitude for the market portfolio because of the greater duration of this claim.

Note first that $b_{\phi\lambda_2}(\tau)$ is positive, implying that an increase in the probability of a boom increases the value of the market. The magnitude of this effect is about half the size of that of disasters. The reason is that there is asymmetry in the value function regarding booms and disasters. Consider the last term in (2.17): $E_{\nu_j}\left[e^{b_{\mu_j}Z_{jt}}\left(e^{b_{\phi\mu_j}(\tau)Z_{jt}} - 1\right)\right]$. The magnitude of this function is determined in large part by this relatively simple expression. For booms, the term $b_{\mu_j}Z_{jt}$ is negative, implying that the immediate effect of a positive jump on prices, given by $e^{b_{\phi\mu_j}(\tau)Z_{jt}} - 1$, is scaled down.[17] For disasters, however, $b_{\mu_j}Z_{jt}$ is positive, implying that the effect of a negative jump is scaled up. Finally, an increase in

---

[17]The expression $e^{b_{\phi\mu_j}(\tau)Z_{jt}} - 1$ gives the percent change in the price of zero-coupon equity with maturity τ.

the probability of a boom decreases the price of the value claim, because of the riskfree rate effect described above.

Risk premia

Figures 2.5–2.10 decompose risk premia into the compensation for the various sources of risk in the economy. These results are useful for understanding the simulation results that follow. Figure 2.5 shows the equity premium (left panel) and the risk premium on the value claim (right panel) as a function of the probability of a disaster. The risk premium (defined in Section 2.2.3) represents the expected instantaneous return on the asset less the riskfree rate.

In both panels, the solid line shows the full risk premium. This can be decomposed into the static rare event premium and the λ-premium (the compensation for time-varying risk of rare events). The static rare event premium is shown by the dashed line. The static rare event premium can itself be decomposed into the static boom premium and the static disaster premium. The static disaster premium is shown by the dashed-dotted line. Finally, there is the premium for risk in consumption in normal times, as would obtain in the CCAPM (dashed line).

As Figure 2.5 shows, the CCAPM premium is negligible, not surprisingly, given the low value of risk aversion. Both the static rare event premium and the full premium are increasing in the probability of a disaster. While the static rare event premium is substantial, the full premium is more than twice as large, indicating that the risk of time-varying rare event risk is important. For the market portfolio, the static disaster premium lies below the rare event premium, indicating that the static boom premium is positive; however for the value claim it is negative. In both cases, it is small in comparison with the other components (at least when the probability of a boom is fixed at its mean). The static boom premium arises from the co-movement of marginal utility and prices during rare events. Holding all else equal, marginal utility changes less in response to a boom than to a disaster.

Figure 2.6 shows the total λ-premium and the disaster component, the λ_1-premium. As this figure shows, nearly the entire λ-premium is accounted for by disaster risk. The λ_2-premium is negligible. Why this difference? Recall that the λ_1-premium is given by

$$\underbrace{-b_{\lambda_1}}_{\text{price of risk}} \times \underbrace{\frac{1}{G}\frac{\partial G}{\partial \lambda_1}\sigma_{\lambda 1}^2 \lambda_{1t}}_{\text{risk loading}}.$$

An analogous expression holds for the λ_2-premium. From evaluating the terms in this expression, we see that two forces contributing to make the compensation for time-varying disasters much greater than for booms. First, the price of risk for time-varying disasters is much larger in magnitude; b_{λ_1} is 11.7, while b_{λ_2} is -3.9. Second, changes in the probability of disaster have a much greater effect on the price-dividend ratio than do changes in the probability of a boom (that is, $\partial G/\partial \lambda_1$ is about twice the magnitude of $\partial G/\partial \lambda_2$).

Figures 2.7 and 2.8 repeat Figures 2.5 and 2.6, except that risk premia are shown as functions of the probability of a boom. The main conclusion from these figures is the same; except when the probability of a boom is very high, the booms have little contribution to risk premiums.

It is tempting to conclude from this analysis that the presence of booms will have little impact on the cross-section of asset returns. However, while booms have a relatively small impact on true risk premia, their impact on observed risk premia can be large. Whether the sample contains jumps or not makes little difference for disasters, as comparing Figure 2.5 with Figure 2.9 shows. However, for booms, the difference is substantial. Because the λ_2-premium is the same (for a given value of the state variables) regardless of whether booms take place or not, the entire difference must arise from the static boom premium. As shown in Sections 2.2.3 and 2.2.4, the static boom premium switches sign, depending on whether booms are observed are not: In population, the boom premium is positive. However, this value is more than entirely due to the realized return should a boom take place. In normal times, the investors receive a lower-than-average return. Figure 2.10 shows that, for the

market, the premium for booms lowers the equity premium by 1% (per annum) when the probability is at its average value, and possibly much more as the probability of a boom increases. Because the value claim is only exposed to boom risk through the effect on discount rates, the effect is much smaller and in the opposite direction.

2.3.3. Simulation results

In what follows, we consider the population and small-sample properties of the model. Both require a stationary distribution for the rare event probabilities. We show this stationary distribution in Figure 2.11. The solid line shows the probability density function for the disaster probability λ_1, while the dashed line shows the probability density function for the boom probability λ_2. The mean of the disaster probability is greater, as can be seen from the fact that the solid line lies above the dotted line for most of the relevant range. However, the boom probability is more skewed; the chance of unusually high values of the probability is greater for booms than for disasters. This can be seen from the fact that, for the tail of the distribution, the dashed line lies above the solid line.[18]

To evaluate the quantitative succes of the model, we simulate monthly data for 600,000 years, and also simulate 10,000 60-year samples. For each sample, we initialize the λ_{jt} processes using a draw from the stationary distribution. In the tables, we report population values for each statistic, percentile values from the small-sample simulations, and percentile value for the subset of small-sample simulations that do not contain jumps. It is this subset of simulations that is the most interesting comparison for postwar data.

The aggregate market

Table 2.4 reports moments of log growth rates of consumption and dividends. There is little skewness or kurtosis in postwar annual consumption data.[19] Postwar dividend growth

---

[18]As Cox et al. (1985) discuss, the stationary distribution for λ_{jt} is Gamma with shape parameter $2\kappa_j \bar{\lambda}_j / \sigma_{\lambda_j}^2$ and scale parameter $\sigma_{\lambda_j}^2/(2\kappa_j)$. This characterization simplifies drawing from the stationary distribution.

[19]In the definition of kurtosis that we use, three is the value for the normal distribution.

exhibits somewhat more skewness and kurtosis. The simulated paths of consumption and dividends for the no-jump samples are, by definition, normal, and the results reflect this. However, the full set of simulations does show significant non-normality; the median kurtosis is seven for consumption and dividend growth. Kurtosis exhibits a substantial small-sample bias. The last column of the table reports the population value of this measure, which is 37.

Table 2.5 reports simulation results for the aggregate market. The model is capable of explaining most of the equity premium: the median value among the simulations with no disaster risk is 4.8%; in the data it is 7.2%. Moreover, the data value is below the 95th percentile of the values drawn from the model indicating the data value is not high enough to reject the model at the 10% level.

Several other recent papers note that the equity premium can be explained by allowing for consumption disasters. However, this paper departs from most of the literature in that the disasters are to expected rather than realized consumption growth. Our results thus speak to a debate concerning whether properly accounting for the smoothness of consumption growth, and the multiperiod nature of disasters, greatly reduces their effect. Barro (2006) calibrates the disaster sizes using a peak-to-trough measure of disasters. In the data, these disasters typically unfold over several years. Barro's model, and that used by a number of subsequent papers treats the disasters as occurring instantaneously. Constantinides (2008) and Julliard and Ghosh (2012) show that if instead the annual declines in consumption are used, the disasters explain only a small portion of the equity premium. In effect, converting the disasters to annual from multiperiod increases their frequency, but greatly reduces their size. Further increasing the frequency to monthly and beyond further reduces the effect. This debate recalls earlier concerns raised in response to the rare disaster model of Rietz (1988) (see Mehra and Prescott (1988)).

To understand this debate, it is necessary to distinguish between two different ways of confronting the problem of the different frequency of consumption and returns. One response

is to model both the consumption data and the returns as occurring at the same frequency. Indeed, Barro (2006) notes that changing the frequency at which returns are measured has very little effect on the model's ability to explain the equity premium. That is, if one's goal is to explain long-horizon returns using long-horizon consumption growth, the disaster risk model is successful.[20] There are some drawbacks, however. Most of the literature focuses on the equity premium that is observable at short horizons. More importantly, explaining long-horizon returns in this way implicitly assumes a decision interval for agents that spans several years. This is not realistic.

A second response is to explicitly model the consumption declines as taking place over several periods, while allowing a realistically short decision interval. If one assumes that consumption growth is iid, but that there are more, smaller, disasters, then certainly it is difficult to explain the equity premium as noted above. If one considers these consumption declines as happening together, a power utility model with leverage below risk aversion would actually have greater difficulty in explaining the equity premium than in the iid case, because prices rise when further consumption declines become more likely. Equity thereby becomes a disaster hedge.[21]

How can one reconcile the fact that the model can explain multi-year returns (assuming a buy-and-hold investor) but not single-year returns (assuming an investor who can trade at realistic intervals)? Moreover, it seems odd, intuitively, that agents would not somehow take into account that disaster-years occur together. In fact, this result is a knife-edge property of power utility. Moving beyond power utility, even slightly (as in this paper; risk aversion and the EIS are not very different) implies that the agent takes more than just the instantaneous innovation to consumption growth into account when pricing assets. Indeed as Hansen (2012) notes, the recursive utility investor takes the long run into account when

---

[20]Constantinides (2008) discusses this precise issue. However, in his equation that addresses the long-horizon return and consumption growth problem, he does not take into account the fact that reducing the frequency raises the probability of disasters; for example, going from one to three years increases the probability of a disaster by a factor of three.

[21]This point is made in various contexts by Gourio (2008), Nakamura et al. (2011) and Wachter (2012).

pricing assets, similarly to the power utility investor with a long decision interval. Thus by making consumption smooth and allowing disasters to unfold slowly, we offer a plausible description of consumption dynamics that confronts the problem raised by Constantinides (2008) and others, but we can still explain a substantial fraction of the equity premium.

Before moving on to the cross-section, we note two limitations to the model's fit to the data. First, the government bond yield in the model is higher than in the data (2.9% vs. 1.25%). This fit could be improved by allowing a fraction of the disaster to hit consumption immediately (or a larger fraction than in the present calibration to hit within the first three months). In fact, results reported in Table 2.2 suggest that this might better fit the behavior of disasters in the data, and, provided that the fraction of the disaster that hits instantaneously would be relatively small, would not raise concerns regarding the discussion of consumption smoothness above. This effect would be straightforward to implement in the model, but would substantially complicate the notation and exposition without changing any of the underlying economics. We should also note that Treasury bill returns may in part reflect liquidity at the very short end of the yield curve (Longstaff (2000)); the model does a better job of explaining the return on the one-year bond.[22] Second, while the model can account for a substantial fraction of the volatility of the price-dividend ratio (the volatility puzzle, reviewed in Campbell (2003)), it cannot explain all of it, at least if we take the view that the postwar series in a sample without rare events. This is a drawback that the model shares with other models attempting to explain aggregate prices using time-varying moments (see the discussion in Bansal et al. (2012) and Beeler and Campbell (2012)) but parsimoniously-modeled preferences. It arises from strong general equilibrium effects: time-varying moments imply cash flow, riskfree rate, and risk premium effects, and one of these generally acts as an offset to the other two, limiting the effect time-varying moments have on prices. One possible response is that some behavior of the prices (i.e. the "bubble" in the late 1990s) may be beyond the reach of this type of model. Certainly this is a fruitful

---

[22]The model predicts a near-zero volatility for returns on this bill in samples without disasters. This is not a limitation, since the volatility in returns in the data is due to inflation, which is not captured in the model.

area for further research.

Unconditional moments of value and growth portfolios

Table 2.6 reports cross-sectional moments. Recall that the data moments are constructed using the growth portfolio as the top decile formed by sorting on book-to-market and the value portfolio as the remaining nine deciles. The resulting difference between the value and the growth portfolio is 1.34%. In samples without jumps, the model easily accounts for this difference; the median value is in fact 2.16%. The higher expected return does not come about because of an increase in volatility: the standard deviation of returns on the value portfolio in the model is in fact far lower than the standard deviation of growth returns. Moreover, the model correctly captures the relative Sharpe ratios of value and growth, as well as the Sharpe ratio on the value-minus-growth strategy. In population, the value premium is negative because growth stocks are in fact more risky than value in the model. However, this population number is not necessarily relevant for calibration in a rare events model; among the full set of simulated paths, the 95 percent critical value of the value premium is 3.35%, far above what is measured in the data. If the value premium does not represent a return for risk, what in the model makes it arise? As explained in Sections 2.2.4 and 2.3.2, it is because investors are willing to accept a lower return on growth in most periods, in return for an occasional very high payout.[23]

While we have chosen to match the data for the top growth portfolio and the remaining nine value deciles, our results can explain much of the more traditional value premium when value is the top decile and growth is the bottom decile. The 95% critical value in our no-jump simulations is 4%, close to the 6% observed in the data. A realistic extension of the model might involve value stocks having greater declines in disasters than growth stocks. This would of course increase the value premium. In most models, it would also, counterfactually, lead value stocks to have higher betas and higher volatilities than growth

---

[23] A value premium can also be observed in many, but not all developed economies, as reported by Fama and French (1992). Over their relatively short sample period, as in the U.S., these countries do not appear to have experienced large booms.

stocks. In the present model this need not be the case, as the main mechanism of the model counteracts this effect.

The focus of this manuscript is not so much on the raw expected returns but on the alphas and betas for value and growth stocks. As Table 2.6 shows, the model exhibits negative alphas for the growth portfolio and positive alphas for the value portfolio. Moreover, the beta on the value portfolio is below one, and the beta on the growth portfolio is above one, just as in the data. Indeed, the result is more extreme than in the data, reflecting the highly convex nature of growth returns in our model. Interestingly, the pattern for alphas and betas does not just characterize the median sample in the no-jump simulations, it also characterizes the median sample in the full set of simulations, as well as in population. Thus, unlike previous models of the value premium, our model is able to explain the patterns in betas on growth and value in the data. It does so in a way that is consistent with the patterns in expected returns.

The discussion of prices and risk premia in Sections 2.3.2 and 2.3.2 is useful in understanding why the betas on growth stocks are above one, and why the alphas are negative. First note that growth stocks are quite volatile because they account for the entire market's loading on the risk of booms. In the model, growth represents a highly levered claim on the innovations of the economy. Risk premia, on the other hand, arise almost entirely from disasters. They arise both from the co-movement of marginal utility and asset prices during disasters themselves, and from the covariance of asset prices with the risk of disasters during normal times. There is a large endogenous asymmetry between the effects of disasters and booms, stemming from the fact that the investor's marginal utility is relatively insensitive to positive events. Thus growth stocks have high volatility, but not the kind of volatility that leads to risk premia.

Return predictability

In a recent survey, Cochrane (2011) notes that time-varying risk premia are a common feature across asset classes. However, variables that predict excess returns in one asset class often fail in another, suggesting that more than one economic mechanism lies behind this common predictability.[24] For example, as the tables below show, the price-dividend ratio is a significant predictor of aggregate market returns, but fails to predict the value-minus-growth return. On the other hand, the value spread predicts the value-minus-growth return, but it is less successful than the price-dividend ratio at predicting the aggregate market return.

Table 2.8 shows the results of regressing the aggregate market portfolio return on the price-dividend ratio in the actual and simulated data. Not surprisingly given earlier work (Wachter (2012)), the model can reproduce the data finding that the price-dividend ratio predicts excess returns. This result arises from the fact that a high value of the disaster probability is followed, on average, high returns, because a higher than average premium compensates investors for taking on greater risk. As described above, a high disaster probability also pushes down the price-dividend ratio. A time-varying boom probability lowers the effect of predictability, since in a sample without jumps, times of higher-than-average boom probabilities signify lower-than-average returns. However, this effect is not large enough to overturn the effect of disasters. Note that in the full set of simulations, predictability is still present, but it is smaller. This is because more of the variance of stock returns arises from the (more volatile) realized dividends during these periods. In population, the magnitude of predictability is smaller, reflecting the well-known small-sample bias in predictive regressions.

In the data, the market return can also be predicted by the value spread, though with substantially smaller t-statistics and R^2 values (Table 2.9). The model also captures the

---

[24]Lettau and Wachter (2011) show that if a single factor drives risk premia, then population values of predictive coefficients should be proportional across asset classes.

sign and the relative magnitude of this predictability; in a sample without jumps, the median R^2 is 3% at the 1-year horizon, compared with a data value of 5%. The coefficient implies that high realizations of the value spread are associated with low future market returns. Like the price-dividend ratio, the value spread is a function of the probability of disaster, so the intuition above goes through in this case. The reason is that the market is somewhat more sensitive to changes in disaster risk than the value spread (though the cash flow effects are similar) because of its greater duration. Thus the price of the value claim declines by less than the price of the market when the risk of a disaster rises. Of course, the value spread is also determined by the boom probability, which has minimal effects on the market expected return. This is why the R^2 values are much lower in this case.

Table 2.10 shows that, in contrast to the market portfolio, the value-minus-growth return cannot be predicted by the price-dividend ratio. The data coefficient is positive and insignificant. This fact represents a challenge for models that seek to simultaneously explain market returns and returns in the cross-section since the forces that explain time-variation in the equity premium also lead to time-variation in the value premium (e.g. Lettau and Wachter (2011), Santos and Veronesi (2010)); this reasoning would lead the coefficient to be negative. The present model does, however, predict a positive coefficient. A high value of the price-dividend ratio on the market indicates a relatively high probability of a boom. In samples without rare events, the return on growth will be lower than the return on value when the boom probability is high. In the population, the coefficient is negative (and quite small); times of high λ_2 precede periods of high returns on growth when jumps occur with their proper frequency.[25]

One might think that the reason that the value-minus-growth return cannot be predicted

---

[25] The median coefficient across all simulations is also positive, on account of small-sample bias. This bias arises from the negative correlation between shocks to the price-dividend ratio and shocks to the value-minus-growth return. Shocks to the disaster probability decrease the price-dividend ratio; both value and growth returns fall, but growth falls by more because of its higher duration. Shocks to the boom probability increase the price-dividend ratio; value returns fall but growth returns rise. This bias is conceptually the same as for regressions of the market portfolio on the price-dividend ratio (see Stambaugh (1999)), but, because the correlation is negative rather than positive, it is in the opposite direction.

by the price-dividend ratio is that it is not very predictable. This is, however, not the case. Table 2.11 shows that, as in the data, the value spread predicts the value-minus-growth return with a positive sign in samples without jumps. The median R^2 value at a 1-year horizon is 9%, compared with a data value of 10%. At a 5-year horizon, the value in the model is 34%, it is 21% in the data. The intuition is the same as for the regressions on the price-dividend ratio. When the probability of a boom is high (but the boom does not occur), the realized return on value is high relative to growth. The R^2 values are much higher than for the price-dividend ratio because the value spread is primarily driven by the probability of a boom, while the price-dividend ratio is only driven by this probability to a small extent.[26]

To summarize, the joint predictive properties of the price-dividend ratio and the value spread would be quite difficult to explain with a model in which single factor drives risk premia; they therefore constitute independent evidence of a multiple-factor structure of the kind presented here.

2.4. Conclusion

This paper has addressed the question of how growth stocks can have both low returns and high risk, as measured by variance and covariance with the market portfolio. It does so within a framework that is also consistent with what we know about the aggregate market portfolio; namely the high equity premium, high stock market volatility, and time-variation in the equity premium. The problem can be broken into two parts: why is the expected return on growth lower, and why is the abnormal return relative to the CAPM negative? This latter question is important, because one does not want to increase expected return through a counterfactual mechanism.

This paper answers the first of these questions as follows: Growth stocks have, in population,

---

[26]In population, the effect works in the opposite direction because high values of the boom probability predict low returns on value relative to growth. The resulting R^2 coefficients are very small. For the set of all simulations, the median coefficient is again positive because of small-sample bias, as explained in footnote 25.

a slightly higher expected return. In finite samples, however, this return may be measured as lower. The answer to the second question is different, because the abnormal return relative to the CAPM appears both in population and samples characterized by a value premium. The abnormal return result arises because risk premia are determined by two sources of risk, each of which is priced very differently by the representative agent. Covariance during disasters, and covariance with the changing disaster probability is assigned a high price by the representative agent because marginal utility is low in these states. However, growth stock returns are highly influenced by booms, and by the time-varying probability of booms. Because marginal utility is low in boom states, the representative agent does not require compensation for holding this risk. This two-factor structure is also successful in accounting for the joint predictive properties of the market portfolio and of the value-minus-growth return.

A number of extensions of the present framework are possible. In this paper, we have specified the growth and the value claim in a stark manner. Extending our results to a setting with richer firm dynamics would allow one to answer a broader set of questions. Further, we have chosen a relatively simple specification for the latent variables driving the economy. An open question is how the specification of these variables affects the observable quantities. We leave these interesting topics to future research.

Figure 2.1: Tails of the one-year consumption growth rate distribution

Panel A: Model

Panel B: Data
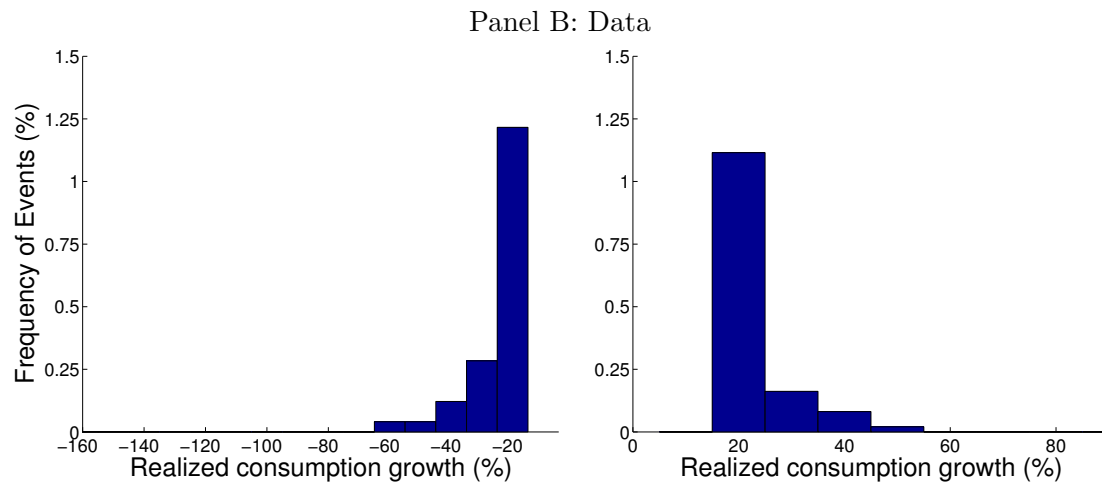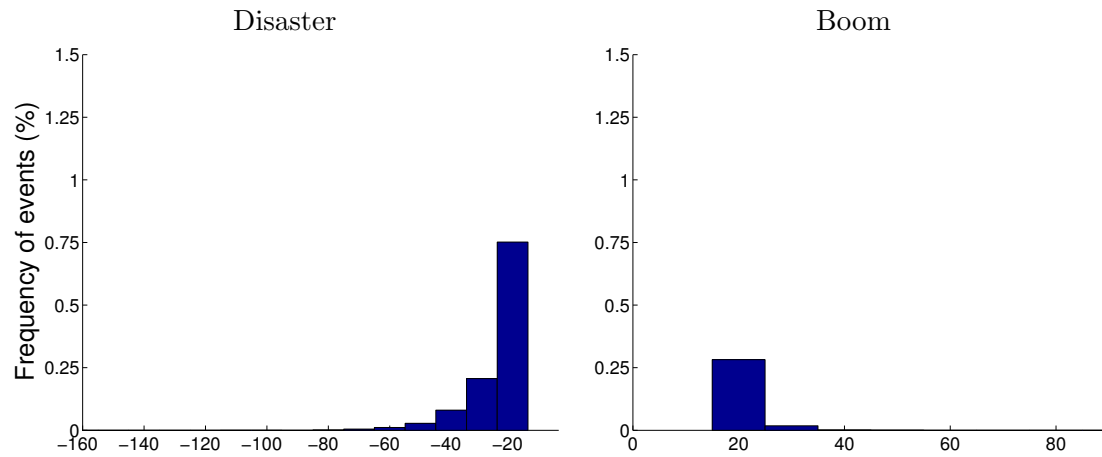
Note: This figure shows histograms of one-year consumption growth rates. The right panel considers growth rates above 15%. The left panel considers growth rates below -15%. The frequency is calculated by the number of observations within a range, divided by the total number of observations in the sample. Panel A shows results from simulated data from the model. Panel B shows results from the data. Data are from Barro and Ursua (2008). For the consumption booms, we exclude observations between 1944 and 1953.

Figure 2.2: Tails of the five-year consumption growth rate distribution

Panel A: Model

Panel B: Data
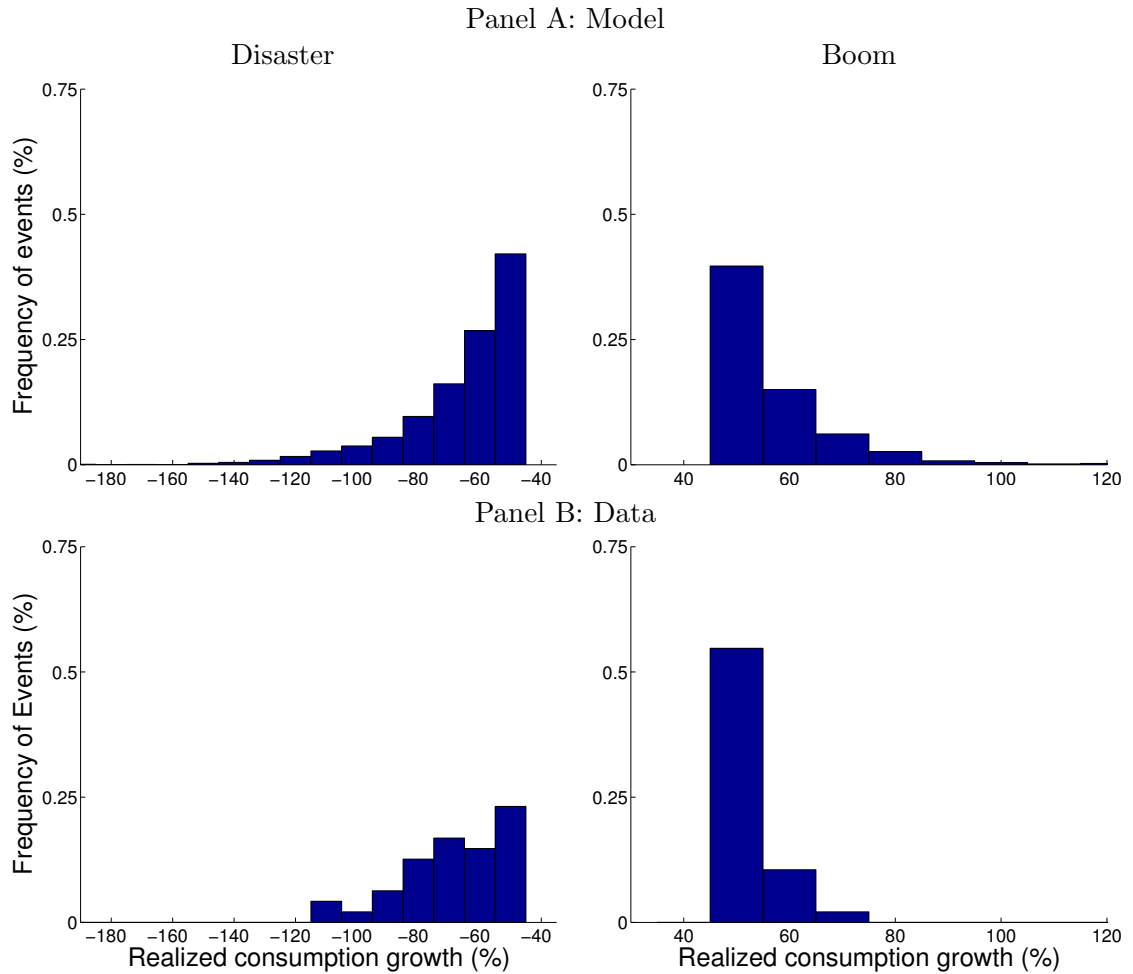
Notes: This figure shows histograms of five-year consumption growth rates. The right panel considers growth rates above 45%. The left panel considers growth rates below -45%. Panel A shows results from simulated data from the model. Panel B shows results from the data. Data are from Barro and Ursua (2008). For the consumption booms, we exclude five-year periods beginning between 1940 and 1948.

Figure 2.3: Solution for the price-dividend ratio: Coefficients on terms in the expected growth rate
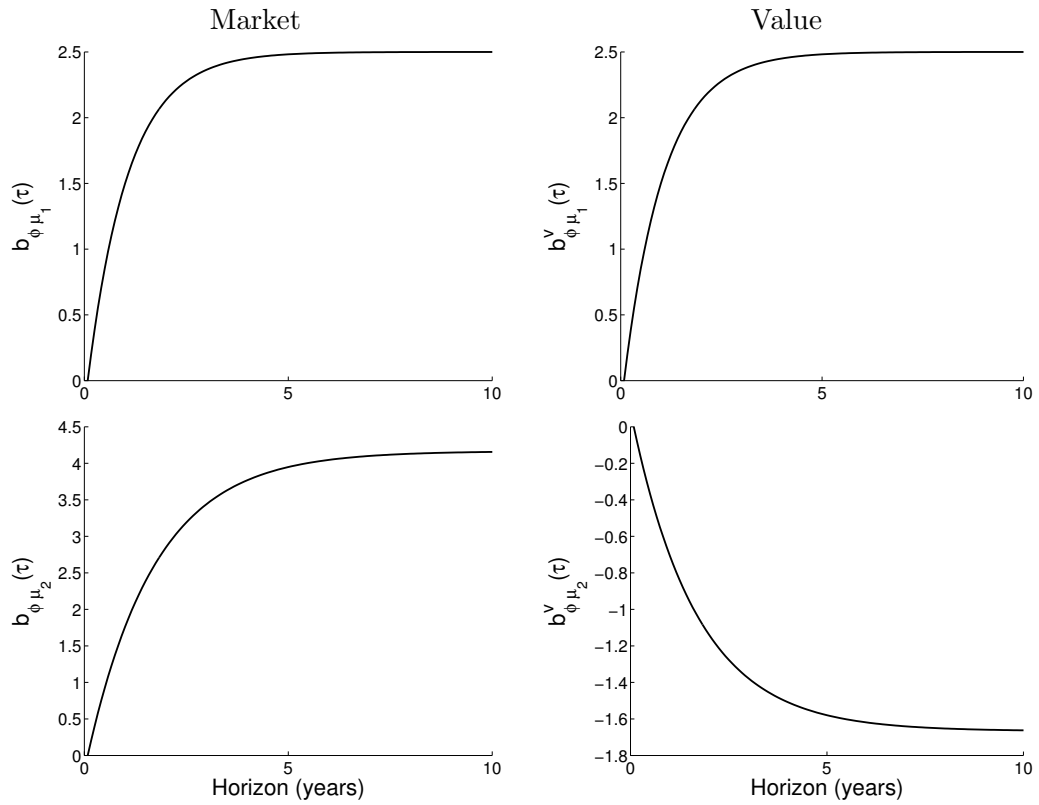
Notes: The left panel shows the coefficients multiplying μ_{1t} and μ_{2t} in the price-dividend ratio for the market. The right panel shows the analogous coefficients for the value claim. The scales on the right and left for $b_{\phi\mu_2}$ differ.

Figure 2.4: Solution for the price-dividend ratio: Coefficients on the jump probabilities
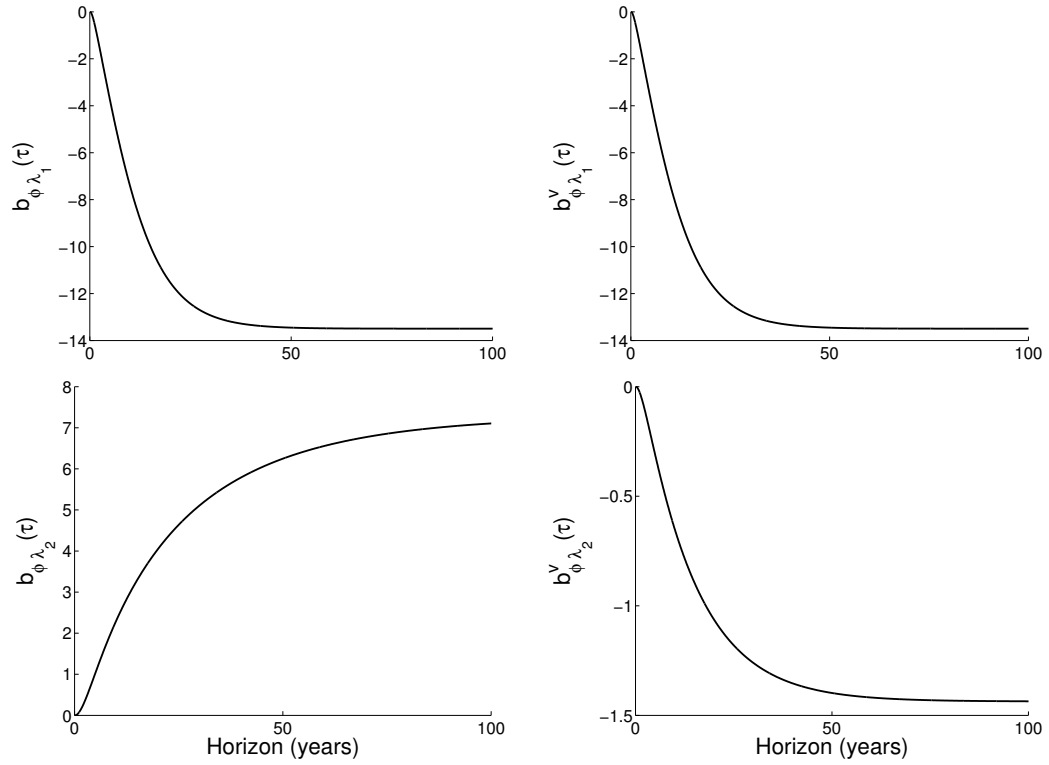
Notes: The left panel shows the coefficients multiplying λ_{1t} (the probability of a disaster) and λ_{2t} (the probability of a growth miracle) in the price-dividend ratio for the market. The right panel shows the analogous coefficients for the value claim. The scales on the right and the left for $b_{\phi\lambda_2}(\tau)$ differ.

Figure 2.5: Risk premiums as functions of the probability of disaster

Notes: The figure shows components of the equity premium (left figure) and of the risk premium on the value claim (right figure). The solid line represents the full premium, the dotted line the CCAPM premium, the dashed-dotted line the static disaster premium and the dashed line the static rare event premium (namely, the static disaster premium plus the static boom premium). Premiums are shown as a function of the disaster probability, λ_1, while the boom probability, λ_2, is fixed at its mean of 2.5%. The vertical line represents the mean of the disaster probability. Premiums are defined relative to the riskfree rate and are in annual terms.

Figure 2.6: λ-premiums (compensation for changing rare event probabilities) as functions of the probability of disaster.
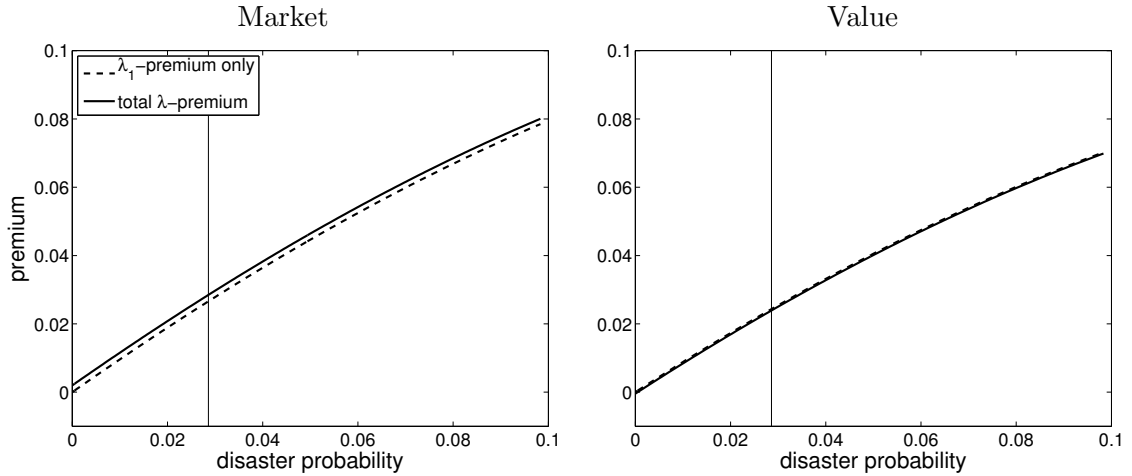
Notes: The solid line shows the component of the equity premium (left figure) and of the risk premium on the value claim (right figure) that compensates for the risk of changing rare event probabilities. This term, referred to as the λ-premium, can be divided into the compensation for disaster probabilities (λ_1-premium; shown by the dashed line) and the compensation for boom probabilities (λ_2-premium). Premiums are shown as a function of the disaster probability, λ_1, while the boom probability, λ_2, is fixed at its mean of 2.5%. The vertical line represents the mean of the disaster probability. Premiums are defined relative to the riskfree rate and are in annual terms.

Figure 2.7: Risk premiums as functions of the probability of a boom
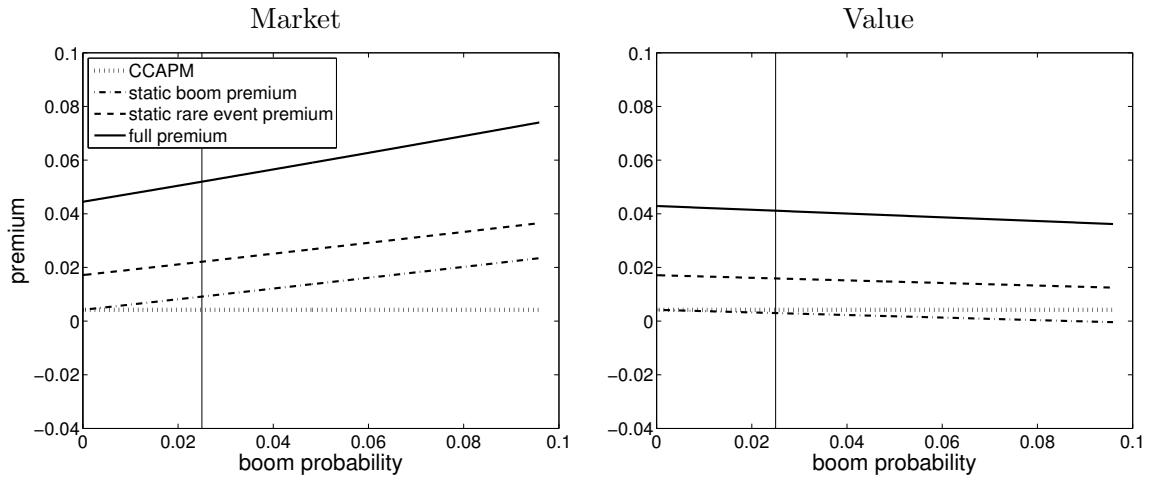
Notes: The figure shows components of the equity premium (left figure) and of the risk premium on the claim (right figure). The solid line represents the full premium, the dotted line the CCAPM premium, the dashed-dotted line the static boom premium and the dashed line the static rare event premium (namely, the static disaster premium plus the static boom premium). Premiums are shown as a function of the boom probability, λ_2, while the disaster probability, λ_1, is fixed at its mean of 2.86%. The vertical line represents the mean of the boom probability. Premiums are defined relative to the riskfree rate and are in annual terms.

Figure 2.8: λ-premiums (compensation for changing rare event probabilities) as a function of the probability of a boom.
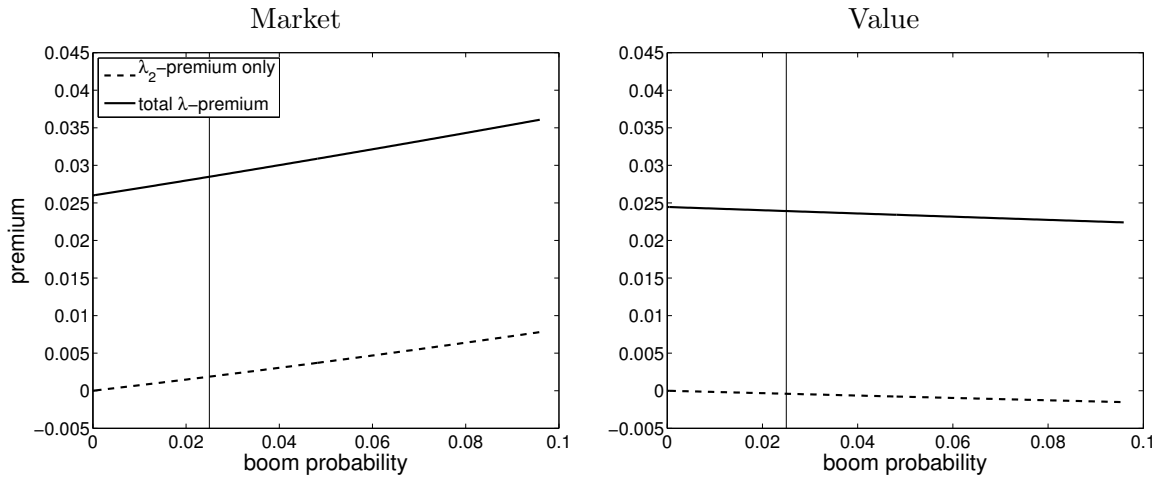
Notes: The solid line shows the component of the equity premium (left figure) and of the risk premium on the value claim (right figure) that compensates for the risk of changing rare event probabilities. This term, referred to as the λ-premium, can be divided into the compensation for disaster probabilities (λ_1-premium) and the compensation for boom probabilities (λ_2-premium; shown by the dashed line). Premiums are shown as a function of the boom probability, λ_2, while the disaster probability, λ_1, is fixed at its mean of 2.86%. The vertical line represents the mean of the boom probability. Premiums are defined relative to the riskfree rate and are in annual terms.

Figure 2.9: Observed expected excess returns in a sample without jumps as a function of disaster probability

Notes: This figure shows expected realized returns in excess of the riskfree rate in a sample without jumps. The left panel shows expected excess returns on the market, while the right panel shows expected excess returns on the value claim. The solid line represents the full premium, the dotted line the CCAPM premium, the dashed-dotted line the static disaster premium (observed in a sample without jumps) and the dashed line the static rare events premium (also observed in a sample without jumps; this is the sum of the static disaster premium and the static boom premium). Premiums are shown as a function of the disaster probability, λ_1, while the boom probability, λ_2, is fixed at its mean of 2.5%. The vertical line represents the mean of the disaster probability. Premiums are defined relative to the riskfree rate and are in annual terms.

Figure 2.10: Observed expected excess returns in a sample without jumps as a function of boom probability

Notes: This figure shows expected realized returns in excess of the riskfree rate in a sample without jumps. The left panel shows expected excess returns on the market, while the right panel shows expected excess returns on the value claim. The solid line represents the full premium, the dotted line the CCAPM premium, the dashed-dotted line the static disaster premium (observed in a sample without jumps) and the dashed line the static rare events premium (also observed in a sample without jumps; this is the sum of the static disaster premium and the static boom premium). premiums are shown as a function of the boom probability, λ_2, while the disaster probability, λ_1, is fixed at its mean of 2.86%. The vertical line represents the mean of the boom probability. Premiums are defined relative to the riskfree rate and are in annual terms.

Figure 2.11: Stationary distributions of rare event probabilities
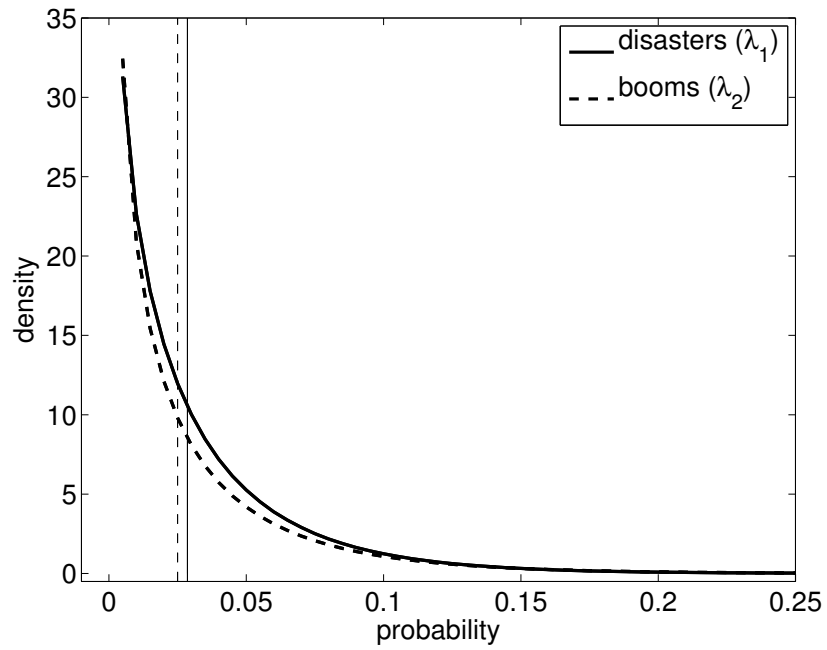
Notes: The figure shows the probability density function of the disaster probability λ_1 and the boom probability λ_2. The probabilities are in annual terms. The vertical solid line shows the location of the mean of the disaster probability while the vertical dashed line shows the location of the mean of the boom probability.

Table 2.1: Parameter values

Panel A: Basic parameters	
Average growth in consumption (normal times) $\bar{\mu}_C$ (%)	2.02
Average growth in dividend (normal times) $\bar{\mu}_D$ (%)	3.55
Volatility of consumption growth (normal times) σ (%)	2.00
Leverage ϕ	3.5
Rate of time preference β	0.012
Relative risk aversion γ	3.0
Panel B: Disaster parameters	
Average probability of disaster $\bar{\lambda}_1$ (%)	2.86
Mean reversion in disaster probability κ_{λ_1}	0.11
Volatility parameter for disasters σ_{λ_1}	0.094
Mean reversion in expected consumption growth κ_{μ_1}	1.00
Minimum consumption disaster (%)	10
Power law parameter for consumption disaster	7
Panel C: Boom parameters	
Average probability of boom $\bar{\lambda}_2$ (%)	2.50
Mean reversion in boom probability κ_{λ_2}	0.05
Volatility parameter for booms σ_{λ_2}	0.070
Mean reversion in expected consumption growth κ_{μ_2}	0.60
Minimum consumption boom (%)	5
Power law parameter for consumption booms	20

Notes: Parameter values for the main calibration, expressed in annual terms.

Table 2.2: Extreme negative consumption events in the model and in the data

Panel A: 1-year rates of decline					
Growth rate	$5-15$	$15-25$	$25-35$	$35-45$	>45
Data	6.89	1.22	0.28	0.12	0.08
Model 1	2.71	0.68	0.20	0.08	0.05
Model 2	0.92	0.80	0.44	0.27	0.33
Panel B: 5-year rates of decline					
Growth rate	$35-45$	$45-55$	$55-65$	$65-75$	>75
Data	0.44	0.23	0.15	0.17	0.25
Model 1	0.65	0.41	0.27	0.17	0.25
Model 2	0.83	0.52	0.33	0.20	0.34

Notes: This table reports frequencies of rates of decline in consumption in the Barro and Ursua (2008) data and in data simulated from the model, for periods of lengths 1 and 5 years. Model 1 refers to the model presented in the text, with jumps in expected consumption growth. Model 2 refers to a model with jumps of the same size in realized consumption, but that is otherwise identical. We compute $(C_t - C_{t+h})/C_t$, where C is consumption and h is the relevant horizon. In both the model and in the data, growth rates are computed using overlapping annual observations. Frequencies are calculated by taking the number of observations within the given range divided by the total number of observations. Frequencies are expressed in percentage terms; for example, 1.22 refers to 1.22% of the observations.

Table 2.3: Extreme positive consumption events in the model and in the data

Panel A: 1-year growth rates					
Growth rate	$5 - 15$	$15 - 25$	$25 - 35$	$35 - 45$	> 45
Data	18.14	1.11	0.16	0.08	0.02
Model	10.44	0.28	0.02	0.00	0.00
Panel B: 5-year growth rates					
Growth rate	$35 - 45$	$45 - 55$	$55 - 65$	$65 - 75$	> 75
Data	2.13	0.55	0.11	0.02	0.00
Model	1.10	0.39	0.15	0.06	0.04

Notes: This table reports frequencies of growth rates in consumption in the Barro and Ursua (2008) data and in data simulated from the model, for periods of lengths 1 and 5 years. Namely, we compute $(C_{t+h} - C_t)/C_t$, where C is consumption and h is the relevant horizon. In both the model and in the data, growth rates are computed using overlapping annual observations. Frequencies are calculated by taking the number of observations within the given range divided by the total number of observations. Frequencies are expressed in percentage terms; for example, 1.11 refers to 1.11% of the observations. For the data, we exclude years following World War II as described in Figures 2.1 and 2.2.

Table 2.4: Log consumption and dividend growth moments

Panel A: Consumption growth

	Data	No-Jump Simulations			All Simulations			Population
		0.05	0.50	0.95	0.05	0.50	0.95	
mean	1.95	1.59	2.00	2.44	−0.10	1.78	3.15	1.70
standard deviation	1.45	1.69	1.99	2.28	1.94	3.32	6.96	4.19
skewness	−0.37	−0.49	0.01	0.48	−3.81	−0.91	1.42	−3.30
kurtosis	3.22	2.17	2.81	3.96	2.62	6.96	21.71	36.57

Panel B: Dividend growth

	Data	No-Jump Simulations			All Simulations			Population
		0.05	0.50	0.95	0.05	0.50	0.95	
mean	1.67	1.85	3.30	4.83	−4.04	2.54	7.32	2.25
standard deviation	6.46	5.93	6.97	7.99	6.79	11.63	24.37	14.67
skewness	0.10	−0.49	0.01	0.48	−3.81	−0.91	1.42	−3.30
kurtosis	4.66	2.17	2.81	3.96	2.62	6.96	21.71	36.57

Notes: Data moments are calculated using annual data from 1947 to 2010. Population moments are calculated from simulating data from the model at a monthly frequency for 600,000 years and then aggregating monthly growth rates to an annual frequency. We also simulate 10,000 60-year samples and report the 5th-, 50th- and 95th-percentile for each statistic both from the full set of simulations and for the subset of samples for which no jumps occur.

Table 2.5: Aggregate market moments

	Data	No-Jump Simulations			All Simulations			Population
		0.05	0.50	0.95	0.05	0.50	0.95	
$E[R^b]$	1.25	2.67	2.93	3.06	0.90	2.70	3.97	2.60
$\sigma(R^b)$	2.75	0.10	0.22	0.43	0.24	2.40	5.76	3.28
$E[R^m - R^b]$	7.25	2.57	4.83	7.36	2.11	5.48	11.37	5.95
$\sigma(R^m)$	17.8	11.1	15.1	21.7	13.5	21.3	37.6	25.6
Sharpe ratio	0.41	0.19	0.32	0.45	0.10	0.26	0.41	0.23
$\exp(E[p-d])$	32.5	26.6	32.1	35.9	21.3	31.0	40.9	30.5
$\sigma(p-d)$	0.43	0.09	0.18	0.32	0.13	0.28	0.54	0.39
AR1$(p-d)$	0.92	0.55	0.78	0.90	0.55	0.79	0.92	0.87

Notes: Data moments are calculated using annual data from 1947 to 2010. Population moments are calculated from simulating monthly data from the model for 600,000 years and then aggregating to an annual frequency. We also simulate 10,000 60-year samples and report the 5th-, 50th- and 95th-percentile for each statistic both from the full set of simulations and for the subset of samples for which no jumps occur. R^b denotes the government bond return, R^m denotes the return on the aggregate market and $p-d$ denotes the log price-dividend ratio.

Table 2.6: Cross-sectional moments

	Data	No-Jump Simulations			All Simulations			Population
		0.05	0.50	0.95	0.05	0.50	0.95	
$E[R^v - R^b]$	7.95	3.18	5.36	7.90	1.67	4.57	7.73	4.59
$E[R^g - R^b]$	6.62	0.34	3.32	7.07	0.43	7.54	25.94	9.67
$E[R^v - R^g]$	1.34	-0.36	2.16	3.90	-21.58	-2.70	3.35	-5.07
$\sigma(R^v)$	17.0	10.4	14.0	19.9	11.8	17.9	26.3	18.8
$\sigma(R^g)$	21.0	18.2	25.5	37.0	23.1	42.8	120.1	66.9
$\sigma(R^v - R^g)$	11.7	12.6	18.4	26.2	15.2	36.3	120.3	64.0
Sharpe ratio, value	0.48	0.25	0.38	0.53	0.09	0.27	0.44	0.24
Sharpe ratio, growth	0.32	0.02	0.13	0.23	0.02	0.17	0.30	0.14
Sharpe ratio, value-growth	0.11	-0.01	0.12	0.27	-0.22	-0.07	0.20	-0.08
alpha, value	1.26	0.77	1.25	2.38	0.15	1.30	6.05	1.57
alpha, growth	-1.26	-6.97	-4.91	-3.03	-13.68	-3.93	0.70	-2.97
alpha, value-growth	2.53	4.01	6.16	8.86	-0.35	5.37	18.66	4.54
beta, value	0.92	0.77	0.91	0.97	0.14	0.79	0.96	0.51
beta, growth	1.09	1.18	1.44	1.73	1.21	1.63	3.34	2.12
beta, value-growth	-0.16	-0.94	-0.54	-0.22	-0.35	5.37	18.66	-1.62

Note: Data moments are calculated using annual data from 1947 to 2010. Population moments are calculated from simulating monthly data from the model for 600,000 years and then aggregating to an annual frequency. We also simulate 10,000 60-year samples and report the 5th-, 50th- and 95th-percentile for each statistic both from the full set of simulations and for the subset of samples for which no jumps occur. R^v denotes the gross return on the value sector, R^g denotes the gross return on the growth sector, alpha denotes the loading of the constant term of the CAPM regression and beta denotes the loading on the market equity excess return of the CAPM regression. In the data, the growth portfolio is the lowest book-to-market decile. The remaining nine deciles comprise the value portfolio.

Table 2.7: Value spread moments

	Data	No-Jump Simulations			All Simulations			Population
		0.05	0.50	0.95	0.05	0.50	0.95	
$\exp(E[\log(\text{value spread})])$	1.23	1.16	1.20	1.32	1.16	1.26	1.71	1.32
$\sigma(\log(\text{value spread}))$	0.08	0.02	0.05	0.14	0.03	0.11	0.34	0.23
Value spread autocorrelation	0.79	0.57	0.80	0.93	0.55	0.78	0.92	0.89

Notes: Data moments are calculated using annual data from 1947 to 2010. Population moments are calculated from simulating monthly data from the model for 600,000 years and then aggregating to an annual frequency. We also simulate 10,000 60-year samples and report the 5th-, 50th- and 95th-percentile for each statistic both from the full set of simulations and for the subset of samples for which no jumps occur. The value spread is defined as the log of the book-to-market ratio for the value sector minus the book-to-market ratio for the aggregate market in the data, and as log price-dividend ratio for the aggregate market minus the log price-dividend ratio for the value sector in the model. In the data, the growth portfolio is the lowest book-to-market decile. The remaining nine deciles comprise the value portfolio.

Table 2.8: Long-horizon regressions of aggregate market returns on the price-dividend ratio

	Data	t-stat	No-Jump Simulations			All Simulations			Population
			0.05	0.50	0.95	0.05	0.50	0.95	

Panel A: 1-year horizon

	Data	t-stat	0.05	0.50	0.95	0.05	0.50	0.95	Population
Coef.	-0.12	$[-2.41]$	-0.56	-0.30	-0.16	-0.43	-0.16	0.03	-0.08
R^2	0.09		0.06	0.15	0.27	0.00	0.05	0.19	0.01

Panel B: 3-year horizon

	Data	t-stat	0.05	0.50	0.95	0.05	0.50	0.95	Population
Coef.	-0.29	$[-3.37]$	-1.17	-0.73	-0.40	-1.00	-0.43	0.10	-0.20
R^2	0.22		0.10	0.33	0.55	0.00	0.11	0.43	0.04

Panel C: 5-year horizon

	Data	t-stat	0.05	0.50	0.95	0.05	0.50	0.95	Population
Coef.	-0.41	$[-3.37]$	-1.54	-0.99	-0.52	-1.35	-0.63	0.16	-0.31
R^2	0.27		0.11	0.42	0.69	0.00	0.16	0.55	0.05

Notes: The table reports coefficients and R^2-statistics from predictive regressions of continuously compounded aggregate market returns in excess of the continuously compounded government bill rate. The predictor variable is the log of the price-dividend ratio on the market. Coef. refers to the coefficient on the predictor variable. Data are annual, from 1947 to 2010. For the data coefficients, we report t-statistics constructed using Newey-West standard errors. Population moments are calculated from simulating monthly data from the model for 600,000 years and then aggregating to an annual frequency. We also simulate 10,000 60-year samples and report the 5th-, 50th- and 95th-percentile for each statistic both from the full set of simulations and for the subset of samples for which no jumps occur.

Table 2.9: Long-horizon regressions of aggregate market returns on the value spread

	Data	t-stat	No-Jump Simulations			All Simulations			Population
Panel A: 1-year horizon									
			0.05	0.50	0.95	0.05	0.50	0.95	
Coef.	-0.50	$[-1.86]$	-1.55	-0.36	0.04	-1.24	-0.12	0.22	-3×10^{-3}
R^2	0.05		0.00	0.03	0.12	0.00	0.01	0.09	7×10^{-6}
Panel B: 3-year horizon									
			0.05	0.50	0.95	0.05	0.50	0.95	
Coef.	-1.18	$[-2.28]$	-3.85	-0.93	0.21	-3.15	-0.33	0.62	-5×10^{-3}
R^2	0.12		0.00	0.07	0.29	0.00	0.03	0.22	8×10^{-6}
Panel C: 5-year horizon									
			0.05	0.50	0.95	0.05	0.50	0.95	
Coef.	-1.28	$[-3.13]$	-5.53	-1.31	0.44	-4.76	-0.50	1.03	-4×10^{-3}
R^2	0.09		0.00	0.09	0.39	0.00	0.04	0.31	4×10^{-6}

Notes: The table reports coefficients and R^2-statistics from predictive regressions of continuously compounded aggregate market returns in excess of the continuously compounded government bill rate. The predictor variable is the value spread, defined in the model as the log price-dividend ratio of the aggregate market minus log price-dividend ratio of the value sector and in the data as the log book-to-market of the value sector minus log book-to-market of the aggregate market. Coef. refers to the coefficient on the predictor variable. Data are annual, from 1947 to 2010. For the data coefficients, we report t-statistics constructed using Newey-West standard errors. Population moments are calculated from simulating monthly data from the model for 600,000 years and then aggregating to an annual frequency. We also simulate 10,000 60-year samples and report the 5th-, 50th- and 95th-percentile for each statistic both from the full set of simulations and for the subset of samples for which no jumps occur.

Table 2.10: Long-horizon regressions of value-minus-growth returns on the price-dividend ratio

	Panel A: 1-year horizon							
			No-Jump Simulations			All Simulations			
	Data	t-stat	0.05	0.50	0.95	0.05	0.50	0.95	Population
Coef.	0.01	[0.37]	−0.04	0.14	0.48	−0.12	0.06	0.38	-5×10^{-3}
R^2	0.00		0.00	0.02	0.11	0.00	0.01	0.09	4×10^{-5}

	Panel B: 3-year horizon							
			No-Jump Simulations			All Simulations			
	Data	t-stat	0.05	0.50	0.95	0.05	0.50	0.95	Population
Coef.	0.05	[0.51]	−0.11	0.36	1.10	−0.32	0.16	0.96	-1×10^{-2}
R^2	0.01		0.00	0.06	0.27	0.00	0.03	0.21	1×10^{-4}

	Panel C: 5-year horizon							
			No-Jump Simulations			All Simulations			
	Data	t-stat	0.05	0.50	0.95	0.05	0.50	0.95	Population
Coef.	0.09	[0.76]	−0.14	0.53	1.62	−0.50	0.26	1.40	-2×10^{-2}
R^2	0.02		0.00	0.09	0.38	0.00	0.04	0.30	2×10^{-4}

Notes: The table reports coefficients and R^2-statistics from predictive regressions of continuously compounded returns on the value portfolio in excess of continuously compounded returns on the growth portfolio. The predictor variable is the log of the price-dividend ratio on the market. Coef. refers to the coefficient on the predictor variable. Data are annual, from 1947 to 2010. For the data coefficients, we report t-statistics constructed using Newey-West standard errors. Population moments are calculated from simulating monthly data from the model for 600,000 years and then aggregating to an annual frequency. We also simulate 10,000 60-year samples and report the 5th-, 50th- and 95th-percentile for each statistic both from the full set of simulations and for the subset of samples for which no jumps occur.

Table 2.11: Long-horizon regressions of value-minus-growth returns on the value spread

			Panel A: 1-year horizon						
			No-Jump Simulations			All Simulations			
	Data	t-stat	0.05	0.50	0.95	0.05	0.50	0.95	Population
Coef.	0.46	[2.52]	0.19	0.86	2.41	−0.13	0.26	1.70	-1×10^{-2}
R^2	0.10		0.01	0.09	0.22	0.00	0.02	0.15	2×10^{-4}
			Panel B: 3-year horizon						
			No-Jump Simulations			All Simulations			
	Data	t-stat	0.05	0.50	0.95	0.05	0.50	0.95	Population
Coef.	1.13	[2.44]	0.56	2.23	5.18	−0.37	0.76	3.98	-4×10^{-2}
R^2	0.19		0.05	0.23	0.47	0.00	0.05	0.36	4×10^{-4}
			Panel C: 5-year horizon						
			No-Jump Simulations			All Simulations			
	Data	t-stat	0.05	0.50	0.95	0.05	0.50	0.95	Population
Coef.	1.48	[2.37]	1.02	3.39	6.47	−0.61	1.24	5.44	-6×10^{-2}
R^2	0.21		0.07	0.34	0.60	0.00	0.09	0.49	6×10^{-4}

Notes: The table reports coefficients and R^2-statistics from predictive regressions of continuously compounded returns on the value portfolio in excess of continuously compounded returns on the growth portfolio. The predictor variable is the value spread, defined in the model as the log price-dividend ratio of the aggregate market minus log price-dividend ratio of the value sector and in the data as the log book-to-market of the value sector minus log book-to-market of the aggregate market. Coef. refers to the coefficient on the predictor variable. Data are annual, from 1947 to 2010. For the data coefficients, we report t-statistics constructed using Newey-West standard errors. Population moments are calculated from simulating monthly data from the model for 600,000 years and then aggregating to an annual frequency. We also simulate 10,000 60-year samples and report the 5th-, 50th- and 95th-percentile for each statistic both from the full set of simulations and for the subset of samples for which no jumps occur.

CHAPTER 3 : Dynamic Asset Allocation with Learning

3.1. Introduction

Standard market models under the efficient market hypothesis are based on the premise that financial markets are informationally efficient and thus assumed constant expected returns. Empirical work beginning with Shiller (1984) and Summers (1986), however, shows that stock returns are predictable. In particular, several papers including Campbell and Shiller (1988), and Fama and French (1989) demonstrate that investors can use predictor variables such as the dividend-price ratio to predict excess returns in the market. The advent of empirical evidence of return predictability has significant consequences for practical issues in portfolio choice theory, particularly for the role of learning.

Return predictability suggests that investors can use past and present data to inform their portfolio allocations at any given time of trade. While many portfolio choice papers acknowledge that investors learn from past data, few explicitly model this mechanism of learning dynamically. Nevertheless, if investors are able to learn, they not only learn from the past but must also account for the fact that they will continue to learn in the future when making their portfolio decisions. Still, intuitively one might think that if the historical data an investor observes is sufficiently long, accounting for her learning in the subsequent periods should likely have a minimal effect on the optimal portfolio decision.

This paper shows that even after the investor observes a full sample of historical data, accounting for learning in the future still has a large effect on her optimal portfolio decision. In particular, I find that dynamic learning induces a large negative hedging demand that increases with the investment horizon. Specifically, an investor with a 5-year investment horizon should decrease the percentage of wealth she allocates to the stock index by over 20 percent even after observing 83 years of data.

This paper is most related to recent work by Brandt et al. (2005) and Skoulakis (2007).

These papers also find that learning induces a negative hedging demand, but disagree on the magnitude of its effect on the investor's portfolio choice. Brandt et al. (2005) find that the negative hedging demand completely eliminates the positive hedging demand from time-varying investment opportunities, whereas Skoulakis (2007) finds that the negative effect induced by learning is not strong enough to drive out the positive hedging demand. Both of these papers, however, use relatively short sample periods. Brandt et al. (2005) choose one short sample period from 1986 to 1995, while Skoulakis (2007) uses several twenty-year sample periods with various starting years. Realistically, investors have access to a much longer sample period of data. Historical data on stock market returns and dividend yields are available from the 1920s onward. Limiting the sample period discards potentially vital information for the investor's portfolio choice decision, and may magnify or disguise the true role of learning. Furthermore, returns are likely to behave differently in a particular ten- or twenty-year sample. Using a shorter sample period may simply pick up the dynamics of the selected ten or twenty years and misrepresent the true dynamics of the return process to the investor.

Additionally, these two papers implement different numerical methods. Skoulakis (2007) uses standard backward induction and employs a feedforward neural network to approximate the value function, while Brandt et al. (2005) develop a less traditional method. The authors first Taylor approximate the value function, then simulate the sample path and use regression to calculate the conditional expectations in the value function. One criticism is that this method may not be accurate when the set of state variables is of high dimension, such that the numerical method drives the results.

To address this concern, I use Skoulakis (2007)'s numerical method to examine the effect of learning using the same predictor variable and ten-year sample period employed by Brandt et al. (2005). I find that learning induces a large negative hedging demand, and that the investor's allocation to the stock index decreases with her investment horizon. Learning plays an important role regardless of the type of numerical method used.

In this paper, I examine the effect of learning on an investor's asset allocation decision after she observes a full sample of data from 1927 to 2009. A criticism of this approach is that historical data may contain structural breaks that may artificially magnify the role of learning. Consequently, I construct the investor's optimal portfolio choice using simulated data and fix the parameters in the underlying processes to eliminate any concerns of a structural break. I not only find that learning still has a large effect on the investor's portfolio choice, but also that she needs to base her initial estimation on a data sample of more than 500 years before the effect of learning begins to diminish. Given that investors only have 83 years of data to draw from, this result only further emphasizes the fact that the dynamic role of learning cannot be ignored in the investors' portfolio choice problem.

The organization of the rest of the paper is as follows. Section 2 describes the investor's problem and the Bayesian updating framework. Section 3 provides the main empirical results. Section 4 addresses criticisms of the paper with extended results. Section 5 discusses the implications for the dynamic portfolio choice literature.

3.2. Portfolio choice problem with predictable returns

This section describes the framework of the Bayesian investor's portfolio choice problem. The investor has an investment horizon of more than one period and is able to rebalance her portfolio periodically.[1] While returns are viewed as predictable, the investor is also has uncertain about the true extent of the predictability of returns. Following the seminal work by Barberis (2000) and others in this literature, the model uses a Bayesian approach to incorporate this parameter uncertainty.

3.2.1. Dynamic asset allocation framework

I consider a simple investment opportunity set of two assets: a risk-free asset with continuously compounded risk-free return r_f and the stock index with continuously compounded

---

[1]The original contributions to the dynamic portfolio choice problem are Merton (1971) and Samuelson (1969). Campbell and Viceira (2002), Brandt (2010), and Wachter (2010) provide survey of portfolio choice literature.

excess return r_t over period t. There are short-sale constraints on both the risk-free asset and the stock index.

The investor observes data up to initial decision time T_0, and will observe new data every subsequent period. She has an investment horizon of $\hat{T}$ periods, and can only rebalance every L periods. Let $K = \hat{T}/L$, where K denotes the number of times she rebalances her portfolio during the $\hat{T}$ periods. Specifically, the investor's time horizon is divided into K intervals, $[t_0, t_1], \cdots, [t_{K-1}, t_K]$ where $t_k = T_0 + kL$ for $k = 0, \cdots, K-1$, and $t_K = T_0 + \hat{T}$. The investor rebalances at time t_k, while ω_k, $k = 0, \cdots, K-1$ denotes the corresponding portfolio weight in the stock index. To make the notations simpler, I use W_k in place of W_{t_k} to denote the investor's wealth at time t_k.

The investor has CRRA utility and maximizes the expected utility of wealth at the terminal date:

$$\max_{\omega_0, \cdots, \omega_{K-1}} E_T \left[\frac{W_K^{1-\gamma}}{1-\gamma} \right],$$

where γ is the coefficient of relative risk aversion.

The wealth process from time t_k to t_{k+1} is represented by:

$$W_{k+1} = W_k R_{p,k+1}.$$

$R_{p,k+1}$ is the investor's portfolio return from t_k to t_{k+1}, which can be expressed as

$$R_{p,k+1} = (1 - \omega_k) \exp(r_f L) + \omega_k \exp\left(r_f L + r_{k+1}^e\right), \tag{3.1}$$

where

$$r_{k+1}^e = r_{t_k+1} + \cdots + r_{t_{k+1}}$$

is the cumulative excess return on the stock index over the L periods from t_k to t_{k+1}.

The set of state variables at time t is denoted by $S(t)$, which characterizes the posterior

distribution.[2] Define the derived utility of wealth as

$$J\left(W_k, S\left(t_k\right), t_k\right) = \max_{\omega_k, \cdots, \omega_{K-1}} E_{t_k}\left[\frac{W_K^{1-\gamma}}{1-\gamma}\right].$$

The Bellman equation of optimality is

$$J\left(W_k, S\left(t_k\right), t_k\right) = \max_{\omega_k} E_{t_k}\left[J\left(W_{k+1}, S\left(t_{k+1}\right), t_{k+1}\right)\right].$$

By the homotheticity of the utility function, this becomes

$$J\left(W_{t_k},\ S\left(t_k\right),\ t_k\right) = \frac{W_{t_k}^{1-\gamma}}{1-\gamma} V\left(S\left(t_k\right), t_k\right).$$

When risk aversion $\gamma > 1$, the Bellman equation can be written as

$$V\left(S\left(t_k\right),\ t_k\right) = \min_{\omega_k} E_{t_k}\left\{R_{p,k+1}^{1-\gamma} \times V\left(S\left(t_{k+1}\right),\ t_{k+1}\right)\right\}, \tag{3.2}$$

with terminal condition $V\left(S\left(t_K\right), t_K\right) = 1$.

Equation (3.2) is the investor's value function. I solve this problem using standard backward induction. I first specify the return process and the Bayesian framework in order to compute the expectation numerically by drawing from the posterior.

In the empirical work of this paper, I assume that the investor observes quarterly data, but can only rebalance annually.

3.2.2. Bayesian framework – predictable returns

Similar to Barberis (2000), Brandt et al. (2005), and Skoulakis (2007), the excess returns r_{t+1} are predictable by some predictor variable x_t which follows an AR(1) process. Frequent

---

[2]For example, in the case of normally distributed i.i.d. returns with known variance, the historical mean of returns is the only state variable. The state variable could potentially be of infinite dimension if the data generating process is more complicated, and solving the investor's problem would become infeasible. This paper considers a data generating process based on Gaussian disturbances, therefore $S(t)$ has finite dimension.

choices of x_t include dividend yield, term spread, and payout yield. Following Brandt et al. (2005), this paper uses dividend yield as the choice of predictor.[3] Specifically, the data generating process is:

$$r_{t+1} = \alpha + \beta x_t + u_{t+1} \tag{3.3}$$

$$x_{t+1} = \theta + \rho x_t + v_{t+1}, \tag{3.4}$$

where

$$\begin{pmatrix} u_{t+1} \\ v_{t+1} \end{pmatrix} \mid x_t, \cdots, x_0, r_t, \cdots, r_1 \sim N\left(0,\ \Sigma\right),$$

and

$$\Sigma = \begin{bmatrix} \sigma_u^2 & \sigma_{uv} \\ \sigma_{uv} & \sigma_v^2 \end{bmatrix}.$$

This data generating process can be rewritten as

$$Y_t = Z_t \Theta + E_t,$$

where

$$y_t = [r_t,\ x_t]^\top,\ z_t = [1,\ x_t]^\top,\ \Theta = \begin{bmatrix} \alpha & \theta \\ \beta & \rho \end{bmatrix},\ \varepsilon_t = [u_t,\ v_t]^\top$$

$$Y_t = [y_1, \cdots, y_t]^\top,\ Z_t = [z_0, \cdots, z_{t-1}]^\top,\ E_t = [\varepsilon_1, \cdots, \varepsilon_t]^\top.$$

Parameters (Θ, Σ) are the unknown to the investor. Zellner and Chetty (1965), Klein and Bawa (1976), Brown (1979), and Bawa et al. (1979) show that the investor should use

---

[3]A vast amount of literature has documented predictability in excess returns. Examples include Fama and Schwert (1977), Keim and Stambaugh (1986), Hodrick (1992) Lettau and Ludvigson (2001), Lewellen (2004), Ang and Bekaert (2007), and Boudoukh et al. (2007). This paper does not attempt to contribute to the debate of the existence of predicability (or the choice of predictor variable), rather this paper takes this process as given and investigates how parameter uncertainty and learning affect the investor's portfolio choice.

the subjective posterior return distribution to maximize her expected utility when facing parameter uncertainty. Kandel and Stambaugh (1996) implement this idea in a single-period problem where the return process is given by (3.3) and (3.4) and show that the estimation risk can significantly decrease the optimal allocation to stock. In this paper, I also use a Bayesian approach to incorporate parameter uncertainty into a multi-period problem. Because the investor's posterior beliefs reflect information in the historical data and her prior beliefs about the parameters, I first need to specify the prior and the likelihood function. I assume the investor has standard diffuse prior beliefs about the parameters (Θ, Σ) (Jeffreys (1961)):

$$p(\Theta, \Sigma) \propto |\Sigma|^{-\frac{3}{2}}. \tag{3.5}$$

Furthermore, treating the initial observation of the regressor x_0 as non-stochastic,[4] the likelihood function of data up to time t is given by

$$p(D_t \mid \Theta, \Sigma, x_0) = (2\pi |\Sigma|)^{-\frac{T}{2}} \exp\left\{-\frac{1}{2}\mathrm{tr}\left[(Y_t - Z_t\Theta)^\top (Y_t - Z_t\Theta)\Sigma^{-1}\right]\right\}. \tag{3.6}$$

Combining (3.5) and (3.6) yields the posterior beliefs about the parameters:

$$p(\Theta, \Sigma) \propto |\Sigma|^{-\frac{T+3}{2}} \exp\left\{-\frac{1}{2}\mathrm{tr}\left[(Y_t - Z_t\Theta)^\top (Y_t - Z_t\Theta)\Sigma^{-1}\right]\right\}$$

$$p(\Sigma^{-1}|D) \sim \text{Wishart}\left(T - 3, \; S^{-1}\right) \tag{3.7}$$

$$p(\text{vec}\,(\Theta)\,|\Sigma, \; D) \sim \text{N}\left(\text{vec}\left(\widehat{\Theta}\right), \; \Sigma \otimes \left(Z'Z\right)^{-1}\right), \tag{3.8}$$

where $\hat{\Theta} = \left(Z^\top Z\right)^{-1} Z^\top Y$ and $S = \left(Y - Z\widehat{\Theta}\right)^\top \left(Y - Z\widehat{\Theta}\right)$.

The matrices $Z^\top Z$, $Z^\top Y$, and $Y^\top Y$ characterize the posterior distribution. Skoulakis (2007) shows that these matrices, and thus the posterior distribution, can be identified using eight

---

[4]Following Stambaugh (1999), I also consider the case where x_0 is treated as stochastic (exact likelihood). See Appendix A.3.3 for details.

state variables, $S = (s_1, \cdots, s_8)$, where s_8 is the current dividend yield.[5]

With these state variables and their laws of motion, one can evaluate the expectation $E_{t_k}[\cdot]$ in the Bellman equation (3.2) with the following steps:

1. Given state variables $S(t_k)$, construct matrices $Z^\top Z$, $Z^\top Y$, and $Y^\top Y$.

2. With these matrices, draw a large number (I) of Σ and Θ using (3.7) and (3.8) to obtain $(\Sigma, \Theta)_i$, $i = 1, \cdots, I$.

3. Using each $(\Sigma, \Theta)_i$, simulate return and predictor data for the next L periods and obtain $(r_{t_j+1}, r_{t_j+2}, \cdots, r_{t_{j+1}})_i$ and $(x_{t_j+1}, x_{t_j+2}, \cdots, x_{t_{j+1}})_i$ for $i = 1, \cdots, I$.

4. With each simulated data, obtain an updated state $S^i(t_{k+1})$ and portfolio return R_p using (3.1).

5. Then $E_{t_k} \left\{ R_{p,k+1}^{1-\gamma} \times V\left(S(t_{k+1}), t_{k+1}\right) \right\} = \frac{1}{I} \sum_{i=1}^{I} R_{p,k+1}^{(i)1-\gamma} \times V\left(S^{(i)}(t_{k+1}), t_{k+1}\right)$.

I solve the investor's portfolio choice problem numerically using backward induction.[6] At each point in time, I discretize each state space into n grids for a total of n^8 grid points. At each grid point $S_j(t)$, $j = 1, \cdots, n^8$, I calculate the posterior distribution of the parameters and use them to simulate a large number of return and dividend yield series. Finally, I obtain an updated state with each simulated path. Starting from the last rebalancing time t_{K-1}, we know that $V\left(S^{(i)}(t_K), t_K\right) = 1$. Thus from Equation (3.2), we know that $V\left(S_j(t_{K-1}), t_{K-1}\right) = \min_{\omega_k} \frac{1}{I} \sum_{i=1}^{I} R_{p,K}^{(i)1-\gamma} \times 1$. Following Barberis (2000), I evaluate the right hand side at $\omega = [0, 1, \cdots, 99]$ to find the optimal portion of wealth that should be invested in the stock index ω_{K-1} and the value $V(S_j(t_{K-1}), t_{K-1})$.

---

[5] Appendix A.3.2 provides the complete derivation of the state variables and their laws of motion.

[6] Schroder and Skiadas (1999), Wachter (2002), Kim and Omberg (1996) and Liu (2007) obtain closed-form solutions under certain assumptions about the parameters. Campbell and Viceira (1999) derive an approximate analytical solution to a consumption-saving problem with infinitely-lived investors. Brennan et al. (1997), Kogan and Uppal (2001), Das and Sundaram (2002), and Brandt et al. (2005) incorporate other features into the dynamic portfolio choice problem and use different numerical and approximation methods to solve it. This paper solves the dynamic problem by discretizing the state space. Balduzzi and Lynch (1999), Brandt (1999), Barberis (2000), Dammon et al. (2001), and Skoulakis (2007) use a similar method.

I repeat this procedure at each grid point and solve backward. From period t_{K-2}, we need to approximate the value function numerically. If the $S^{(i)}(t_{K-1})$ obtained from simulation does not fall on a grid point, one need an approximation for $V\left(S^i\left(t_{K-1}\right), t_{K-1}\right)$. In this paper, I use the Feedforward Neural Network introduced by Skoulakis (2007) to approximate the value function numerically.[7] The advantage of this method is that it requires fewer grid points for each state. In particular, for the results in this paper, I discretize each state into two grid points and solve the problem multiple times to check for accuracy.

3.2.3. Data

I use quarterly data on stock index returns and dividend yield from January 1927 to December 2009. The stock index is the value-weighted index of stocks traded on the NYSE/AMEX/NASDAQ, from The Center for Research in Security Prices (CRSP). Three-month Treasury Bill returns are provided by Ibbotson and Associates and are available on Kenneth French's website. Log excess return is defined by $r_t = r_t^m - r_t^b$, where $r^m = \log(R^m)$ is the log market return and $r^b = \log(R^b)$ is the log Treasury Bill return. Dividend yield in month t is constructed by dividing the total dividends paid during months $t - 11$ through t by the value of the index at the end of month t. Dividend yield in month $t - 1$ is used to predict the return of the quarter spanning from month t to $t + 2$. Following Barberis (2000), I set the continuously compounded risk-free return at 0.0108 per quarter (0.0036 per month), which corresponds to an annual risk-free rate of 4.4%.

3.3. Empirical results

In order to study the effect of learning on portfolio choice, this section considers two different problems. The first one is the same problem studied in Barberis (2000), where the investor acknowledges the parameter uncertainty but does not learn when new data are realized. In this case, the variation in the investment opportunity set induces a positive hedging demand. In the second case, the investor optimally learns about the return-generating process, this

---

[7]See Appendix A.3.1 for more detail.

learning induces a large negative hedging demand.

3.3.1. Parameter uncertainty with no learning

Table 3.1 presents the parameter estimates using quarterly data from 1927 to 2009. I form the posterior distribution by generating $300,000$ draws using Equations (3.7) and (3.8). Table 3.2 presents the portfolio weights in the stock index. Each column in Table 3.2 represents a different current dividend yield (x_{T_0}) and each row represents a different investment horizon $\widehat{T}$. In this case, the investor acknowledges parameter uncertainty, but her beliefs about the parameters are unchanged between the initial decision date T_0 and the terminal date $T_0 + \hat{T}$.

The estimate for β is 0.94 with a standard deviation of 0.42, revealing that a higher current dividend yield marginally predicts a higher excess return in the next period. The investor can therefore time the market by allocating more wealth to the stock index when the current dividend yield is high. Indeed, in Table 3.2 we see that for any given investment horizon, the wealth allocated to the stock index increase with current dividend yield x_{T_0}.

Dividend yield therefore governs the investment opportunity, and changes in dividend yield lead to changes in future expected return. A risk-averse investor will want to hedge against the risk of an unexpected drop in dividend yield. From Table 3.1, we see that excess return innovations and dividend yield innovations are negatively correlated and statistically significant ($\rho_{uv} = -0.86$ with a standard deviation of 0.01). In other words, excess returns will be unexpectedly high (higher wealth) when the dividend yield is unexpectedly low (bad investment opportunity, lower wealth). Thus the investor can hedge against future expected return movement by allocating more wealth to the stock index. Furthermore, this hedging demand increases with horizon, and we can see in Table 3.2 that for a given current dividend yield x_{T_0}, the portfolio weight in the stock index increases with the investment horizon.

3.3.2. Learning

Table 3.3 contains the main results of this paper, where the investor's decision incorporates learning and parameter uncertainty. More specifically, the investment decision takes into account that her posterior beliefs about the parameters change over time as new data become available.

To solve the portfolio choice problem, I discretize each of the eight states by three grid points: $s_j^{down} < s_j^{mean} < s_j^{up}$, $j = 1, \cdots, 8$. In particular, for $j = 1, \cdots, 7$, s_j^{mean} corresponds to the value of the state variable j at the end of the sample period in 2009. The grid for the last state variable s_8 is set to be comparable to the previous case: $(s_8^{down}, s_8^{mean}, s_8^{up})$ $= (0.02, 0.04, 0.06)$. Then I solve the problem twice, each time using a combination of two different sets of grids. More specifically, the left panel labeled *Specification 1* uses the combination of (S^{down}, S^{mean}), and the right panel labeled with *Specification 2* uses (S^{mean}, S^{up}), where $S^j = (s_1^j, \cdots, s_8^j)$. Columns labeled with an L are the portfolio weights in the stock index when the investor takes learning into account, and the results are reported for three different values of current dividend yield (x_{T_0} or s_8). Each of the other seven state variables s_j is fixed at s_j^{mean}. In order to demonstrate the effect of learning, Table 3.3 also shows the portfolio weights in the stock index when the investor ignores learning, corresponding to the results in the columns labeled with an *NL*.

Table 3.3 shows that learning induces a large negative hedging demand, which significantly decreases investor's allocation to stocks. To see the effect of learning, I fix the current dividend yield and compare the results with and without learning. One can see that the portfolio weight in the stock index is almost flat along the investment horizon with learning, and it increases with the investment horizon without learning. For example, the percentage of wealth that an investor with a five-year investment horizon allocates to the stock index drops between 23 to 43 percent (depending on the current dividend yield) when she takes learning into account. These changes are significant. For example, if current dividend yield is 4%, an investor who takes learning into account should invest 15% more of her wealth in

the risk-free asset. Furthermore, from Table 3.3 one can see that even with learning, the investor still attempts to time the market by investing more in the stock index when the current dividend yield is high.

As the investment horizon becomes shorter, the effect of learning decreases and the difference between the two allocation strategies diminishes. When the investment horizon is only one year, there is no rebalance in the future the investor has no new information to learn. As one would expect, the two strategies results in an identical portfolio allocation. Figure 3.1 shows the effect of learning by plotting the decline in percentage of wealth allocated to the stock index, in percentage term.

The benefit of the above procedure is that it significantly reduces the number of grid points. Furthermore, both of the specifications contain the set S^{mean}, therefore we can have more confidence if both specifications yield the same results for the case in which all state variables are at their means. Indeed, one can see that when current dividend yield is 0.04, the results of the two specifications are very close.[8]

It is easy to see how learning can introduce a negative hedging demand in the simple case where excess returns are i.i.d. and investor only learns about the mean. Consider

$$r_{t+1} = \mu + \varepsilon_{t+1}, \text{ where } \varepsilon_{t+1} \sim N(0, \sigma).$$

This case is also examined by Brennan (1998) and Barberis (2000). The only unknown is the mean excess return μ, and its posterior $\tilde{\mu}_{t+1}$ governs the investment opportunity. Brennan (1998) shows that an investor with power utility and relative risk aversion $\gamma > 1$ wishes to hedge against learning that the investment opportunity is bad (low $\tilde{\mu}_{t+1}$). Since $\tilde{\mu}_{t+1}$ is updated based on realized return r_{t+1}, hedging against low $\tilde{\mu}_{t+1}$ means hedging against low r_{t+1}, and she can achieve this by simply allocating less wealth to the risky stock index.

---

[8]To further validate the accuracy of these results, I also solve the problem using three grid points for each state, and the results are similar.

Xia (2001) also considers predictable returns, however, in her model there is only uncertainty about β. She finds that learning about predictability β induces a positive hedging demand when the dividend yield is below its mean, a zero hedging demand when it is at its mean, and a negative hedging demand when it is above its mean. In the present paper, the investor learns about the full set of parameters simultaneously and I find the net hedging demand to be negative, consistent with Brandt et al. (2005) and Skoulakis (2007).

3.4. Extensions

Section 3 shows that even after observing a long sample of data, learning still has a large negative effect on the investor's asset allocation decision. The question naturally arises is that if 83 years of data is insufficient for the investor to learn the information she needs to make the optimal investment decision, how many years are needed? In Section 4.1, I simulate artificial data with various sample lengths and test how much data the investor needs to observe before the effect of learning diminishes.

Previous literature agrees that learning induces a net negative hedging demand and decreases the weight of wealth the investor allocates to the stock index, but there is disagreement on the magnitude of the effect. In Section 4.2 I try to reconcile the difference in previous literature. In particular, I find that using data from 1986 to 1995 (the sample used by Brandt et al. (2005)) leads to a larger negative effect of learning and long term investment are even less attractive to the investor, which is consistent with their results.

3.4.1. Extending the sample length for initial estimation

Theoretically, as an investor sees more data, she should become more confident in her estimates of the parameters of the data generating process. Thus, accounting for learning in subsequent periods should have a minimal effect on her portfolio choice decision. Section 3.3 shows that data covering 83 years is not sufficient to diminish the effect of learning on the investor's portfolio allocation. The negative hedging demand induced by learning eliminates most of the positive hedging demand that comes from the correlation between

shocks to excess returns and shocks to dividend yields.

In this section, I simulate data with different sample lengths to investigate how long the initial sample that the investor observes needs to be for the effect of learning to vanish. Taking the parameter estimates in Table 1 as the true parameters of the underlying return-generating process, I simulate data with different time spans between 100 and $1,000$ years. The investor then makes her initial estimation based on the simulated data. Table 3.4 and Figure 3.2 display the results found when 500 years of data are available for the initial estimation. In Table 3.4 we can see that the allocation decisions with and without learning are very close to each other. In other words, after observing 500 years of data, the investor becomes much more confident about the information she learns in the historical data and cares less about the information she will see in the subsequent periods. Unfortunately, reliable asset market data is only available from the 1920s onward. Therefore, it is essential for investors with long horizons to take into account the effect of learning when making their investment decisions. Otherwise they are likely to end up with a suboptimal portfolio that is over-allocated to the stock index.

This exercise also alleviates the concern that structural breaks in the data drive the results in the previous section. If there was a structural break during the 83-year sample period, the long sample of data would not help the investor to learn about the data generating process. If a structural break exists, the investor might become even more uncertain about the parameters because of the change in the underlying data generation process. Fortunately, this is not a concern with simulated data since the parameters in the underlying data generating process are fixed. The effect of learning is still very strong even in this case, which shows that the results do not simply come from a structural break in the data.

3.4.2. Comparison with Brandt, Goyal, Santa-Clara, and Stroud (2005)

Brandt et al. (2005) and Skoulakis (2007) also solve similar portfolio choice problems. Brandt et al. (2005) use an alternative numerical technique to approach this problem. In

their paper, they first use a fourth-order Taylor expansion for the value function, and then simulate the sample path and use regression to calculate the conditional expectations in the value function. Using data from 1986 to 1995, they find that the negative hedging demand induced by learning about the parameters in the return generating process prevails over the positive hedging demand induced by the negative correlation between excess return and dividend yield shocks. As a result, the investor allocates less wealth to the stock index when she has a longer investment horizon. Skoulakis (2007) uses different risky asset returns, predictors, and sample periods, and also finds a negative hedging demand induced by learning. His paper, however, finds that the negative horizon effect is not strong enough to eliminate the positive hedging demand arises from the negative correlation between shocks, so the portion of wealth the investor allocates to the stock index still increases with the investment horizon.

Skoulakis (2007) argues that the solution technique of Brandt et al. (2005) may not be accurate under such a high dimensional state space. Applying the numerical method of Skoulakis (2007) to the data sample and predictors from Brandt et al. (2005), however, leads to results comparable to Brandt et al. (2005). After repeating the procedures above with data from January 1986 to December 1995, I find that the negative effect of learning dominates the positive hedging demand induced by time-varying investment opportunities. Figure 3.3 shows that the investor's optimal allocation to the stock index decreases with the length of her investment horizon, as suggested by Brandt et al. (2005).[9] Thus, this inconsistency in results is more likely caused by the different small samples they choose to form the initial estimation. Using a shorter sample period may simply pick up the dynamics of the selected ten or twenty years and misrepresent the true dynamics of the return process to the investor.

---

[9]My procedures are not identical to theirs as they allow for quarterly data observation and portfolio rebalance whereas the present paper only allows the investor to adjust her portfolio annually. Furthermore, they set the risk aversion parameter $\gamma = 10$, which might further amplify the negative effect compared to the value of $\gamma = 5$ used in this paper.

3.5. Conclusion

How much information does the past 83 years of asset market data convey? From a Bayesian investor's perspective, not so much. This paper finds that when excess returns are predictable and the long-horizon dynamic investor learns about the full set of parameters, learning induces a large negative hedging demand and significantly decreases the portfolio weight in the stock index.

Previous literature also finds this negative effect, but disagrees on its magnitude. I find results comparable to Brandt et al. (2005) when using their sample periods. Furthermore, using simulated data, I show that in order for the effect of learning to vanish, the investor must observe data covering more than 500 years – far longer than the span of currently available historical data. This simulation exercise also alleviates concerns that a structural break in the sample drives the results, since the effect of learning remains strong when data are generated from a fixed underlying process.

These results illustrate the importance of learning for investors who dynamically rebalance their portfolios. If investors fail to acknowledge that more information will become available between the initial decision date and the terminal date, they will over-allocate wealth to the stock market and obtain a suboptimal portfolio allocation.

Table 3.1: Parameter Estimates

Θ	
-0.0220	0.0028
(0.0170)	(0.0008)
0.9398	0.9255
(0.4183)	(0.0208)

Σ	
0.0116	**-0.8637**
(0.0009)	(0.0140)
-0.8637	2.8×10^5
(0.0140)	(2.3×10^{-6})

Notes: This table presents the initial estimation of the parameters. Quarterly data are from January 1927 to December 2009. The bold numbers in Σ are the estimates of the correlation between u and v ($\rho_{uv} \equiv \frac{\sigma_{uv}}{\sqrt{\sigma_u^2}\sqrt{\sigma_v^2}}$). $\Theta = \begin{bmatrix} \alpha & \theta \\ \beta & \rho \end{bmatrix}$, and $\Sigma = \begin{bmatrix} \sigma_u^2 & \rho_{uv} \\ \rho_{uv} & \sigma_v^2 \end{bmatrix}$.

Table 3.2: Portfolio Allocation to Risky Asset

Horizon (years)	Current $\frac{d}{p}$		
	0.02	0.04	0.06
1	8	39	69
2	9	43	77
3	11	47	83
4	13	51	87
5	15	54	91
6	17	56	95
7	18	59	96
8	20	60	98
9	20	65	99
10	21	65	99

Notes: This table presents the percentage of wealth allocated to the stock index when the investor does not take into account the effect of future learning. Quarterly data are from January 1927 to December 2009. The relative risk aversion coefficient $\gamma = 5$. Each column denotes a different current dividend yield, x_{T_0}, and each row denotes a different investment horizon, $\hat{T}$.

Table 3.3: Portfolio Allocation to the Stock Index - Historical Data

|	Specification 1				Specification 2			
d/p	0.02		0.04		0.04		0.06	
Horizon	L	NL	L	NL	L	NL	L	NL
1	9	9	41	41	41	41	70	70
2	10	10	41	45	40	45	70	77
3	11	12	41	48	43	48	71	83
4	9	14	40	52	41	52	69	87
5	9	16	42	55	41	55	70	91

Notes: This table reports the portfolio allocation results using historical data. The numbers represent the percentage of wealth the investor allocates to the stock index. Label NL corresponds to parameter uncertainty without learning and label L corresponds to optimal learning. In specification 1, each state j is discretized into two grids, (s_j^{down}, s_j^{mean}) and in specification 2 each state j is discretized into two grids, (s_j^{mean}, s_j^{up}). In particular, $(s_8^{down}, s_8^{mean}, s_8^{up}) = (0.02, 0.04, 0.06)$. The results are reported for different current dividend yields. All other state variables are fixed at their means. Relative risk aversion coefficient $\gamma = 5$. Horizon is in years. The initial estimation is based on quarterly data are from January 1927 to December 2009.

Table 3.4: Portfolio Allocation to the Stock Index - Simulation

|	Specification 1				Specification 2			
d/p	0.02		0.04		0.04		0.06	
Horizon	L	NL	L	NL	L	NL	L	NL
1	7	7	44	44	44	44	80	80
2	9	9	46	49	47	49	83	91
3	11	11	50	55	52	55	87	97
4	13	13	53	59	58	59	92	99

Notes: This table reports the portfolio allocation results using simulated data. The numbers represents the percentage of wealth the investor allocates to the stock index. Label NL corresponds to parameter uncertainty without learning and label L corresponds to optimal learning. In specification 1, each state j is discretized into two grids, (s_j^{down}, s_j^{mean}) and in specification 2 each state j is discretized into two grids, (s_j^{mean}, s_j^{up}). In particular, $(s_8^{down}, s_8^{mean}, s_8^{up}) = (0.02, 0.04, 0.06)$. The results are reported for different current dividend yields. All other state variables are fixed at their means. Relative risk aversion coefficient $\gamma = 5$. Horizon is in years. The initial estimation is based on $2,000$ quarters (500 years) of simulated data using the parameter estimates in Table 3.1.

Figure 3.1: Effect of Learning - Full Sample
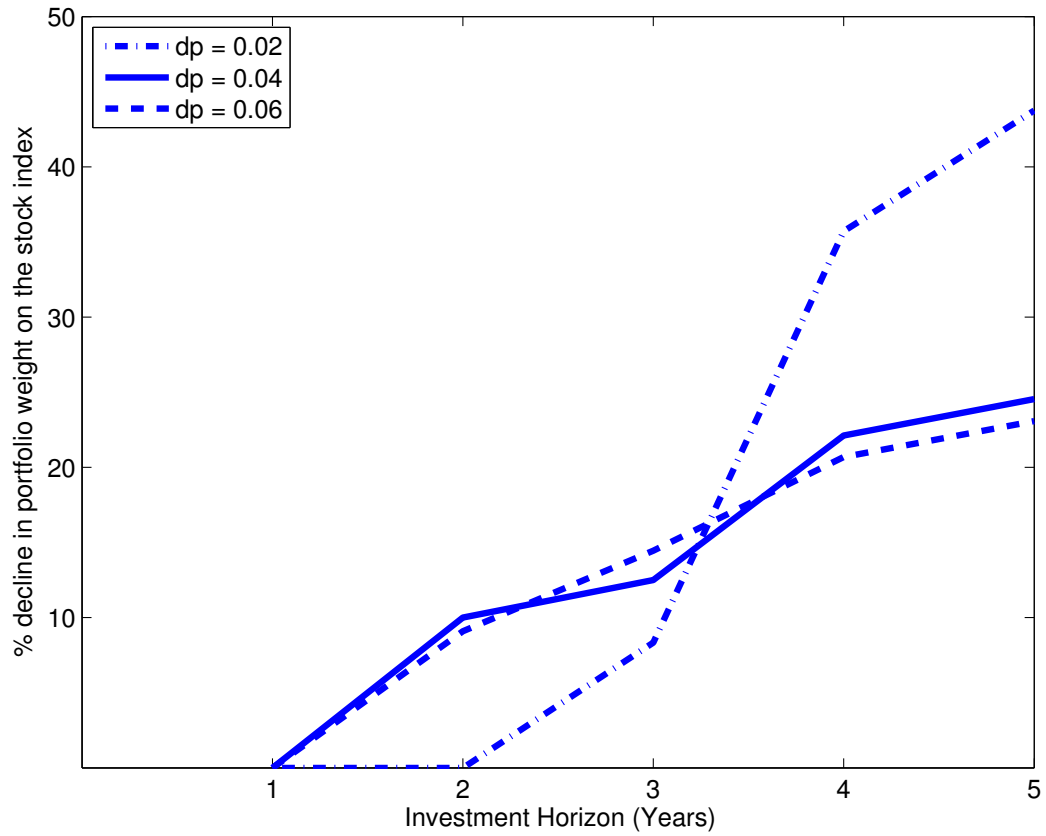
Notes: This figure plots the the percentage decline in portfolio weight on the stock index when learning is taken into account. Quarterly data are from January 1927 to December 2009. The dash-dotted line is for $d/p = 0.02$, the solid line is for $d/p = 0.04$, and the dashed line is for $d/p = 0.06$. For $d/p = 0.04$, I plotted the average of the results from the two specifications.

Figure 3.2: Effect of Learning - Simulated Data
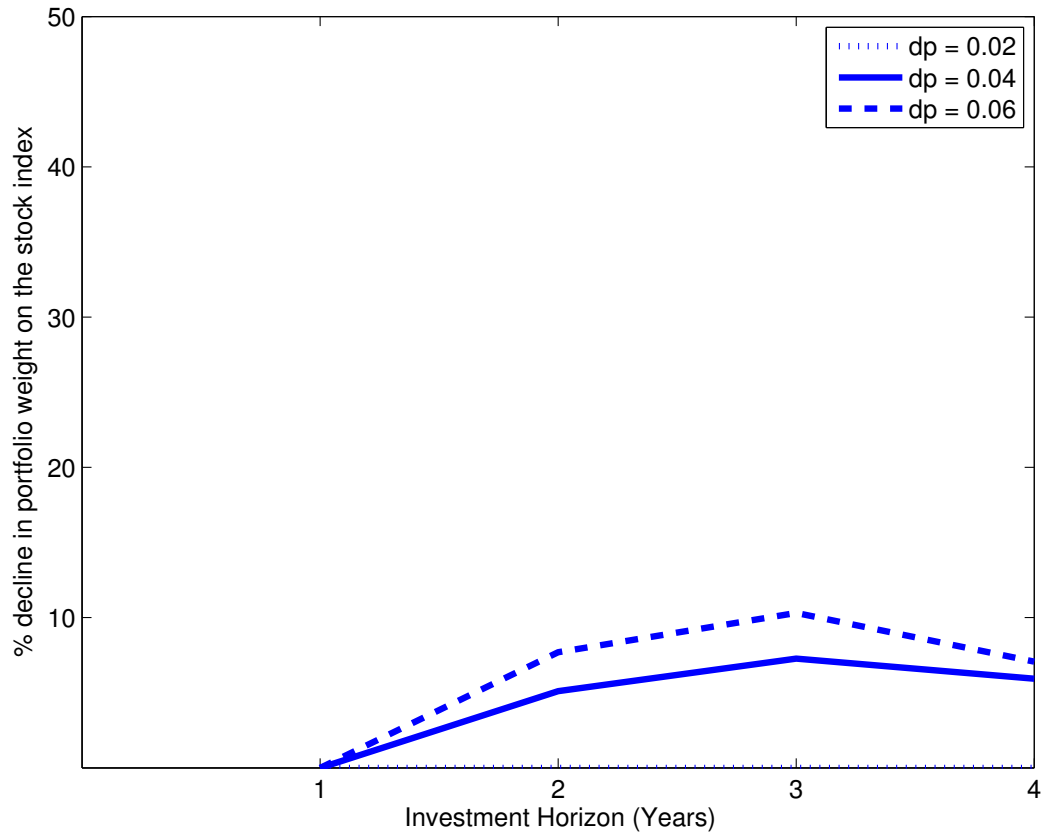
Notes: This figure plots the the percentage decline in portfolio weight on the stock index when learning is taken into account. 2,000 quarters of data are simulated using the parameter estimates in Table 3.1. The dash dotted line is for $d/p = 0.02$, the solid line is for $d/p = 0.04$, and the dashed line is for $d/p = 0.06$. For $d/p = 0.04$, I plotted the average of the results from the two specifications.

Figure 3.3: Comparison with Brandt et al. (2006)
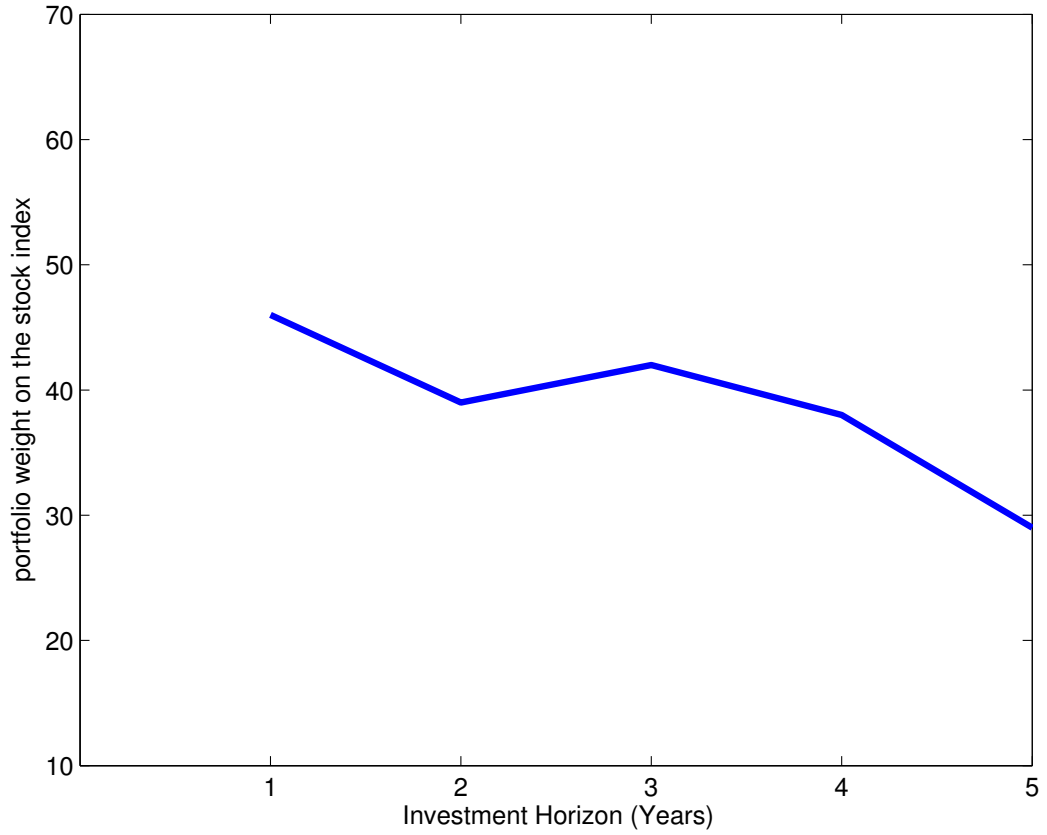
Notes: This figure plots the percentage of wealth allocated to the stock index against investment horizon. Initial estimations for the parameters are formed using quarterly data from January 1986 to December 1995. All state variables are at their means, current dividend yield is set at 0.06.

A.1. Appendix for Rare Disasters and the Term Structure of Interst Rates

A.1.1. Model derivation

Notation

Definition A.1. *Let X be a jump-diffusion process. Define the jump operator of X with respect to the jth type of jump as the following:*

$$\mathcal{J}_j(X) = X_{t_j} - X_{t_{j-}} \quad j \in \{c, cq, q\},$$

for t_{j-} such that a type-j jump occurs. Then define

$$\bar{\mathcal{J}}_j(X) = E_{\nu_j}\left[X_{t_j} - X_{t_{j}-}\right] \quad j \in \{c, cq, q\},$$

and

$$\bar{\mathcal{J}}(X) = [\bar{\mathcal{J}}_c(X), \ \bar{\mathcal{J}}_{cq}(X), \ \bar{\mathcal{J}}_q(X)]^\top.$$

The value function

Proof of Theorem 1.1 Let S denote the value of a claim to aggregate consumption, and conjecture that the price-dividend ratio for the consumption claim is constant:

$$\frac{S_t}{C_t} = l,$$

for some constant l. This relation implies that S_t satisfies

$$dS_t = \mu S_{t-}dt + \sigma S_{t-}dB_{ct} + (e^{Z_{ct}} - 1)S_t dN_{ct} + (e^{Z_{cq,t}} - 1)S_t dN_{cq,t}. \tag{A.1}$$

Consider an agent who allocates wealth between S and the risk-free asset. Let α_t be the fraction of wealth in the risky asset S_t, and let c_t be the agent's consumption. The wealth process is then given by

$$
dW_t = \left(W_t \alpha_t \left(\mu - r_t + l^{-1} \right) + W_t r_t - c_t \right) dt + W_t \alpha_t \sigma dB_{ct}
$$
$$
+ \alpha_t W_t \left((e^{Z_{ct}} - 1) S_t dN_{ct} + (e^{Z_{cq,t}} - 1) S_t dN_{cq,t} \right),
$$

where r_t denotes the instantaneous risk-free rate. Optimal consumption and portfolio choices must satisfy the following Hamilton-Jacobi-Bellman (HJB) equation:

$$
\sup_{\alpha_t, C_t} \left\{ J_W \left(W_t \alpha_t \left(\mu - r_t + l^{-1} \right) + W_t r_t - c_t \right) + \kappa_{\lambda_c} \left(\bar{\lambda}_c - \lambda_{ct} \right) + \kappa_{\lambda_{cq}} \left(\bar{\lambda}_{cq} - \lambda_{cq,t} \right) \right.
$$
$$
+ \frac{1}{2} J_{WW} W_t^2 \alpha_t^2 \sigma^2 + \frac{1}{2} \left(J_{\lambda_c \lambda_c} \sigma_{\lambda_c}^2 \lambda_{ct} + J_{\lambda_{cq} \lambda_{cq}} \sigma_{\lambda_{cq}}^2 \lambda_{cq,t} \right)
$$
$$
+ \lambda_{ct} E_{\nu_c} \left[J \left(W_t \left(1 + \alpha_t \left(e^{Z_{ct}} - 1 \right) \right), \lambda_t \right) - J \left(W_t, \lambda_t \right) \right]
$$
$$
\left. + \lambda_{cq,t} E_{\nu_{cq}} \left[J \left(W_t \left(1 + \alpha_t \left(e^{Z_{cq,t}} - 1 \right) \right), \lambda_t \right) - J \left(W_t, \lambda_t \right) \right] + f \left(c_t, V_t \right) \right\} = 0, \quad \text{(A.2)}
$$

where J_n denotes the first derivative of J with respect to variable n, for n equal to λ_i or W, and J_{nm} denotes the second derivative of J with respect to n and m.

In equilibrium, $\alpha_t = 1$ and $c_t = W_t l^{-1}$. Substituting these policy functions into (A.2) implies

$$
J_W W_t \mu + J_{\lambda_c} \kappa_{\lambda_c} \left(\bar{\lambda}_c - \lambda_{ct} \right) + J_{\lambda_{cq}} \kappa_{\lambda_{cq}} \left(\bar{\lambda}_{cq} - \lambda_{cq,t} \right) + \frac{1}{2} J_{WW} W_t^2 \sigma^2
$$
$$
+ \frac{1}{2} \left(J_{\lambda_c \lambda_c} \sigma_{\lambda_c}^2 \lambda_{ct} + J_{\lambda_{cq} \lambda_{cq}} \sigma_{\lambda_{cq}}^2 \lambda_{cq,t} \right) + \lambda_{ct} E_{\nu_c} \left[J \left(W_t e^{Z_{ct}}, \lambda_t \right) - J \left(W_t, \lambda_t \right) \right]
$$
$$
+ \lambda_{cq,t} E_{\nu_{cq}} \left[J \left(W_t e^{Z_{cq,t}}, \lambda_t \right) - J \left(W_t, \lambda_t \right) \right] + f \left(c_t, V_t \right) = 0. \quad \text{(A.3)}
$$

By the envelope condition $f_C = J_W$, we obtain $\beta = l^{-1}$. Given the consumption-wealth

ratio, it follows that

$$f(c_t, V_t) = f\left(W_t l^{-1}, J(W_t, \lambda_t)\right) = \beta W_t^{1-\gamma}\left(\log\beta - \frac{\log I(\lambda_t)}{1-\gamma}\right). \tag{A.4}$$

Substituting (A.4) and (1.6) into (A.3) and dividing both sides by $W_t^{1-\gamma} I(\lambda_t)$, we find

$$
\begin{aligned}
\mu + I^{-1}(1-\gamma)^{-1}&\left(I_{\lambda_c}\kappa_{\lambda_c}(\bar{\lambda}_c - \lambda_{ct}) + I_{\lambda_{cq}}\kappa_{\lambda_{cq}}(\bar{\lambda}_{cq} - \lambda_{cq})\right) - \frac{1}{2}\gamma\sigma^2 \\
&+ \frac{1}{2}I^{-1}\left(I_{\lambda_c\lambda_c}\sigma^2_{\lambda_c}\lambda_{ct} + I_{\lambda_{cq}\lambda_{cq}}\sigma^2_{\lambda_{cq}}\lambda_{cq,t}\right) \\
&+ (1-\gamma)^{-1}\left(\lambda_c E_{\nu_c}\left[e^{(1-\gamma)Z_c} - 1\right] + \lambda_{cq}E_{\nu_{cq}}\left[e^{(1-\gamma)Z_{cq}} - 1\right]\right) \\
&+ \beta\left(\log\beta - \frac{\log I(\lambda_t)}{1-\gamma}\right) = 0,
\end{aligned}
$$

where I_{λ_j} denotes the first derivative of I with respect to λ_j and $I_{\lambda_j\lambda_j}$ denotes the second derivative for $j \in \{c, cq\}$.

Collecting terms in λ_{jt} results in the following quadratic equation for b_j:

$$\frac{1}{2}\sigma^2_{\lambda_j}b_j^2 - (\kappa_{\lambda_j} + \beta)b_j + E_{\nu_j}\left[e^{(1-\gamma)Z_j} - 1\right],$$

for $j \in \{c, cq\}$, implying

$$b_j = \frac{\kappa_{\lambda_j} + \beta}{\sigma^2_{\lambda_j}} \pm \sqrt{\left(\frac{\kappa_{\lambda_j} + \beta}{\sigma^2_{\lambda_j}}\right)^2 - 2\frac{E_{\nu_j}\left[e^{(1-\gamma)Z_j} - 1\right]}{\sigma^2_{\lambda_j}}},$$

Collecting constant terms results in the following characterization of a in terms of b:

$$a = \frac{1-\gamma}{\beta}\left(\mu - \frac{1}{2}\gamma\sigma^2\right) + (1-\gamma)\log\beta + \frac{1}{\beta}b^\top\left(\kappa_\lambda * \bar{\lambda}\right).$$

Here and in what follows, I use $*$ to denote element-by-element multiplication of vectors of equal dimension. Given the form of $I(\lambda)$, $I_{\lambda_j} = b_j I$ and $I_{\lambda_j\lambda_j} = b_j^2 I$ for $j \in \{c, cq\}$. Because there are no interaction terms, the solution takes the same form as when there is only a

single type of jump. As in (Wachter, 2012, Appendix A.1) we take the negative root of the corresponding equation for b_j to find:

$$b_j = \frac{\kappa_{\lambda_j} + \beta}{\sigma_{\lambda_j}^2} - \sqrt{\left(\frac{\kappa_{\lambda_j} + \beta}{\sigma_{\lambda_j}^2}\right)^2 - 2\frac{E_{\nu_j}\left[e^{(1-\gamma)Z_j} - 1\right]}{\sigma_{\lambda_j}^2}}.$$

$\square$

Proof of Corollary 1.1 Since $\gamma > 1$, if $Z_j < 0$, then the second term in the square root of (1.9) is positive. Therefore the square root term is positive but less than $\frac{\kappa_j + \beta}{\sigma_j^2}$, and $b_j > 0$. Similarly, if $Z_j > 0$ then the second term in the square root of (1.9) is negative. Therefore the square root term is positive and greater than $\frac{\kappa_j + \beta}{\sigma_j^2}$, and $b_j < 0$. $\square$

Proof of Corollary 1.2 The risk-free rate is obtained by taking the derivative of the HJB (A.2) with respect to α_t, evaluating at $\alpha_t = 1$, and setting it equal to 0. The result immediately follows. $\square$

The state-price density

Duffie and Skiadas (1994) show that the state-price density π_t equals

$$\pi_t = \exp\left\{\int_0^t f_V\left(C_s, V_s\right) ds\right\} f_C\left(C_t, V_t\right),$$

where f_C and f_V denote derivatives of f with respect to the first and second argument respectively. Note that the exponential term is deterministic. From equation (1.4), I obtain

$$f_C\left(C_t, V_t\right) = \beta\left(1 - \gamma\right)\frac{V}{C}.$$

From the equilibrium condition $V_t = J\left(\beta^{-1}C_t, \lambda_t\right)$, together with the form of the value function (1.6), I get

$$f_C\left(C_t, V_t\right) = \beta^\gamma C_t^{-\gamma} I(\lambda_t). \tag{A.5}$$

Applying Ito's Lemma to (A.5) implies

$$\frac{d\pi_t}{\pi_{t-}} = \mu_{\pi t}dt + \sigma_{\pi t}dB_t + \left(e^{-\gamma Z_{ct}} - 1\right)dN_{ct} + \left(e^{-\gamma Z_{cq,t}} - 1\right)dN_{cq,t}, \qquad (A.6)$$

where

$$\sigma_{\pi t} = \left[-\gamma\sigma,\ 0,\ 0,\ b_c\sigma_c\sqrt{\lambda_{ct}},\ b_{cq}\sigma_{cq}\sqrt{\lambda_{cq,t}}\right]. \qquad (A.7)$$

It also follows from no-arbitrage that

$$
\begin{aligned}
\mu_{\pi t} &= -r_t - \left(\lambda_{ct}E_{\nu_c}\left[e^{-\gamma Z_{ct}} - 1\right] + \lambda_{cq,t}E_{\nu_{cq}}\left[e^{-\gamma Z_{cq,t}} - 1\right]\right) \\
&= -\beta - \mu + \gamma\sigma^2 - \left(\lambda_{ct}E_{\nu_c}\left[e^{(1-\gamma)Z_{ct}} - 1\right] + \lambda_{ct}E_{\nu_{cq}}\left[e^{(1-\gamma)Z_{cq,t}} - 1\right]\right). \qquad (A.8)
\end{aligned}
$$

From (A.6) we can see that in the event of a disaster, marginal utility (as represented by the state-price density) jumps upward. This implies that investors require compensation for bearing disaster risks. The first element of (A.7) implies that the standard diffusion risk in consumption is priced; more importantly, changes in λ_{jt} are also priced as reflected by the last two elements of (A.7).

The nominal state-price density $\pi^\$$ equals

$$\pi_t^\$ = \frac{\pi_t}{P_t}. \qquad (A.9)$$

The nominal state-price density follows

$$\frac{d\pi_t^\$}{\pi_{t-}^\$} = \mu_{\pi t}^\$ dt + \sigma_{\pi t}^\$ dB_t + \left(e^{-\gamma Z_{ct}} - 1\right)dN_{ct} + \left(e^{-\gamma Z_{cq,t}} - 1\right)dN_{cq,t}, \qquad (A.10)$$

where

$$\sigma_{\pi t}^\$ = \left[-\gamma\sigma,\ -\sigma_P,\ 0,\ b_c\sigma_{\lambda_c}\sqrt{\lambda_{ct}},\ b_{cq}\sigma_{\lambda_{cq}}\sqrt{\lambda_{cq,t}}\right], \qquad (A.11)$$

and

$$\mu_{\pi t}^{\$} = -\beta - \mu + \gamma\sigma^2 - q_t + \sigma_P^2 - \left(\lambda_{ct}E_{\nu_c}\left[e^{(1-\gamma)Z_{ct}} - 1\right] + \lambda_{ct}E_{\nu_{cq}}\left[e^{(1-\gamma)Z_{cq,t}} - 1\right]\right). \quad \text{(A.12)}$$

By comparing (A.11) to (A.7), we can see that the second element is no longer zero. This implies that the diffusion risk in inflation is also priced in the nominal state-price density. By comparing (A.12) to (A.8), we can see that the expected inflation and volatility of realized inflation also affect the drift of the nominal state-price density.

Proof of Corollary 1.3 It follows from no-arbitrage that

$$\mu_{\pi t}^{\$} = -r_t^{\$} - \left(\lambda_{ct}E_{\nu_c}\left[e^{-\gamma Z_{ct}} - 1\right] + \lambda_{cq,t}E_{\nu_{cq}}\left[e^{-\gamma Z_{cq,t}} - 1\right]\right),$$

where $\mu_{\pi t}^{\$}$ is given by (A.12). Therefore the nominal risk-free rate on a nominal bond, $r_t^{\$}$ is

$$r_t^{\$} = \beta + \mu - \gamma\sigma^2 + q_t - \sigma_P^2 + \lambda_{ct}E_{\nu_c}\left[e^{-\gamma Z_{ct}}\left(e^{Z_{ct}} - 1\right)\right] + \lambda_{qt}E_{\nu_{cq}}\left[e^{-\gamma Z_{cq,t}}\left(e^{Z_{cq,t}} - 1\right)\right].$$

$\square$

A.1.2. Pricing general zero-coupon equity

This section provides the price of a general form of a zero-coupon equity, both in real terms and in nominal terms. The dividend on the aggregate market and the face value on the bond market will be special cases.

Real assets

First I will consider the price of a real asset. Consider a stream of cash-flow that follows a jump-diffusion process:

$$\frac{dD_t}{D_{t^-}} = \mu_D\,dt + \sigma_D\,dB_t + (e^{\phi_{D,c}Z_{ct}} - 1)\,dN_{ct}^D + (e^{\phi_{D,cq}Z_{cq,t}} - 1)\,dN_{cq,t}^D. \quad \text{(A.13)}$$

This stream of cash-flow is subject to Poisson shocks dN_{jt}^D, $j \in \{c, cq\}$. The arrival time of these Poisson shocks are linked to the arrival time of consumption disasters.

Assumption A.1. *When a consumption disaster happens, this cash-flow stream experiences a jump with probability p_D; that is, for $j \in \{c, cq\}$.*

- *If $dN_{jt} = 0$, then $dN_{jt}^D = 0$.*

- *If $dN_{jt} = 1$, then*

$$dN_{jt}^D = \begin{cases} 1 & \textit{with probability } p_D \\ \\ 0 & \textit{otherwise.} \end{cases}$$

With this assumption, $\phi_{D,j}$ denotes the jump multiplier for a type-j jump, for $j \in \{c, cq\}$.

Lemma A.1. *Let $H(D_t, \lambda_t, \tau)$ denote the time t price of a single future cash-flow at time $s = t + \tau$:*

$$H(D_t, \lambda_t, s - t) = E_t \left[\frac{\pi_s}{\pi_t} D_s \right].$$

By Ito's Lemma, we can write

$$\frac{dH(D_t, \lambda_t, \tau)}{H(D_t, \lambda_t, \tau)} = \mu_{H(\tau),t} dt + \sigma_{H(\tau),t}^\top dB_t + \mathcal{J}_c(\pi_t H(D_t, \lambda_t, \tau)) dN_{ct} + \mathcal{J}_{cq}(\pi_t H(D_t, \lambda_t, \tau)) dN_{cq,t}.$$

for a scalar process $\mu_{H(\tau),t}$ and a vector process $\sigma_{H(\tau),t}$. Then, no-arbitrage implies that:

$$\mu_{\pi,t} + \mu_{H(\tau),t} + \sigma_{\pi,t} \sigma_{H(\tau),t}^\top + \frac{1}{\pi_t H_t(\tau)} \lambda_t^\top \bar{\mathcal{J}}^{real}(\pi_t H(D_t, \lambda_t, \tau)) = 0. \tag{A.14}$$

Proof No-arbitrage implies that $H(D_s, \lambda_s, 0) = D_s$ and that

$$\pi_t H(D_t, \lambda_t, \tau) = E_t \left[\pi_s H(D_s, \lambda_s, 0) \right].$$

To simplify notation, let $H_t = H(D_t, \lambda_t, \tau)$, $\mu_{H,t} = \mu_{H(\tau),t}$, and $\sigma_{H,t} = \sigma_{H(\tau),t}$. It follows

from Ito's Lemma that

$$\frac{dH_t}{H_{t^-}} = \mu_{H,t}dt + \sigma_{H,t}dB_t + (e^{\phi_{D,c}Z_{ct}} - 1)dN_{ct} + (e^{\phi_{D,cq}Z_{cq,t}} - 1)dN_{cq,t}.$$

Applying Ito's Lemma to $\pi_t H_t$ implies that the product can be written as

$$\pi_t H_t = \pi_0 H_0 + \int_0^t \pi_s H_s \left(\mu_{H,s} + \mu_{\pi,s} + \sigma_{\pi,s}\sigma_{H,s}^\top \right) + \int_0^t \pi_s H_s(\sigma_{H,s} + \sigma_{\pi,s})dB_s$$
$$\sum_{0 < s_{ci} \leq t} \left(\pi_{s_{ci}} H_{s_{ci}} - \pi_{s_{ci}^-} H_{s_{ci}^-} \right) + \sum_{0 < s_{cq,i} \leq t} \left(\pi_{s_{cq,i}} H_{s_{cq,i}} - \pi_{s_{cq,i}^-} H_{s_{cq,i}^-} \right), \quad \text{(A.15)}$$

where $s_{ji} = \inf\{s : N_{js} = i\}$ (namely, the time that the ith time type-j jump occurs, where $j \in \{c, cq\}$).

We use (A.15) to derive a no-arbitrage condition. The first step is to compute the expectation of the jump terms $\sum_{0 < s_{ji} \leq t} \left(\pi_{s_{ji}} H_{s_{ji}} - \pi_{s_{ji}^-} H_{s_{ji}^-} \right)$. The pure diffusion processes are not affected by the jump. Adding and subtracting the jump compensation terms from (A.15) yields:

$$\pi_t H_t = \pi_0 H_0 + \int_0^t \pi_s H_s \left(\mu_{H,s} + \mu_{\pi,s} + \sigma_{\pi,s}\sigma_{H,s}^\top + \frac{1}{\pi_s H_s} \left(\lambda_c \bar{\mathcal{J}}_c(\pi_s H_s) + \lambda_{cq} \bar{\mathcal{J}}_{cq}(\pi_s H_s) \right) \right) ds$$
$$+ \int_0^t \pi_s H_s(\sigma_{H,s} + \sigma_{\pi,s})dB_s + \sum_{0 < s_{ci} \leq t} \left(\left(\pi_{s_{ci}} H_{s_{ci}} - \pi_{s_{ci}^-} H_{s_{ci}^-} \right) - \int_0^t \pi_s H_s \lambda_c \bar{\mathcal{J}}_c(\pi_s H_s)ds \right)$$
$$+ \sum_{0 < s_{cq,i} \leq t} \left(\left(\pi_{s_{cq,i}} H_{s_{cq,i}} - \pi_{s_{cq,i}^-} H_{s_{cq,i}^-} \right) - \int_0^t \pi_s H_s \lambda_{cq} \bar{\mathcal{J}}_{cq}(\pi_s H_s)ds \right) \quad \text{(A.16)}$$

Under mild regularity conditions analogous to those given in Duffie et al. (2000), the second and the third terms on the right hand side of (A.16) are martingales. Therefore the first term on the right hand side of (A.16) must also be a martingale, and it follows that the integrand of this term must equal zero:

$$\mu_{\pi,t} + \mu_{H(\tau),t} + \sigma_{\pi,t}\sigma_{H(\tau),t}^\top + \frac{1}{\pi_t H_t(\tau)} \lambda_t^\top \bar{\mathcal{J}}^{\text{real}}(\pi_t H(D_t, \lambda_t, \tau)) = 0.$$

$$\square$$

Theorem A.1. *The function H takes an exponential form:*

$$H(D_t, \lambda_t, \tau) = D_t \exp\left\{a_\phi(\tau) + \lambda_t^\top b_{\phi\lambda}(\tau)\right\}, \tag{A.17}$$

where $b_{\phi\lambda} = [b_{\phi\lambda_c}, \ b_{\phi\lambda_{cq}}]^\top$. Function $b_{\phi\lambda_j}$ for $j \in \{c, cq\}$ solves

$$\frac{db_{\phi\lambda_j}}{d\tau} = \frac{1}{2}\sigma_{\lambda_j}^2 b_{\phi\lambda_j}(\tau)^2 + \left(b_j \sigma_{\lambda_j}^2 - \kappa_{\lambda_j}\right) b_{\phi\lambda_j}(\tau)$$
$$+ p_D E_{\nu_j}\left[e^{(\phi_{D,j} - \gamma)Z_{jt}} - e^{(1-\gamma)Z_{jt}}\right] + (1 - p_D)E_{\nu_j}\left[e^{-\gamma Z_{jt}} - e^{(1-\gamma)Z_{jt}}\right], \tag{A.18}$$

and function a_ϕ solves

$$\frac{da_\phi}{d\tau} = \mu_D - \mu - \beta + \gamma\sigma(\sigma - \sigma_D) + b_{\phi\lambda}(\tau)^\top\left(\kappa_{\lambda_j} * \bar{\lambda}_j\right). \tag{A.19}$$

The boundary conditions are $a_\phi(0) = b_{\phi\lambda_c}(0) = b_{\phi\lambda_{cq}}(0) = 0$.

Proof See proof of Theorem A.2. $\square$

Nominal asset

Similar no-arbitrage conditions can be derived for nominally denominated assets. Suppose cash-flow that follows:

$$\frac{dD_t^\$}{D_{t-}^\$} = \mu_{D^\$}\, dt + \sigma_{D^\$}\, dB_t + (e^{\phi_{D,c}^\$ Z_{ct}} - 1)\, dN_{ct}^D + (e^{\phi_{D,cq}^\$ Z_{cq,t}} - 1)\, dN_{cq,t}^D,$$

where the process N_{jt}^D is given by Assumption A.1 and $\phi_{D,c}^\$$ and $\phi_{D,cq}^\$$ are the jump multipliers for the N_{c^-} and N_{cq}–type jumps, respectively.

Lemma A.2. *Let $H^\$(D_t^\$, q_t, \lambda_t, \tau)$ denote the time t price of a single future dividend payment at time $t + \tau$:*

$$H^\$(D_t^\$, q_t, \lambda_t, s - t) = E_t\left[\frac{\pi_s^\$}{\pi_t^\$} D_s^\$\right].$$

By Ito's Lemma, we can write

$$\frac{dH^{\$}(D_t^{\$}, q_t, \lambda_t, \tau)}{H^{\$}(D_t^{\$}, q_t, \lambda_t, \tau)} = \mu_{H^{\$}(\tau),t} dt + \sigma_{H^{\$}(\tau),t}^{\top} dB_t + \mathcal{J}_c(\pi_t^{\$} H^{\$}(D_t^{\$}, q_t, \lambda_t, \tau)) dN_{ct}$$

$$+ \mathcal{J}_{cq}(\pi_t^{\$} H^{\$}(D_t^{\$}, q_t, \lambda_t, \tau)) dN_{cq,t} + \mathcal{J}_q(\pi_t^{\$} H^{\$}(D_t^{\$}, q_t, \lambda_t, \tau)) dN_{qt}.$$

for a scalar process $\mu_{H^{\$}(\tau),t}$ and a vector process $\sigma_{H^{\$}(\tau),t}$. Then, no-arbitrage implies that:

$$\mu_{\pi^{\$},t} + \mu_{H^{\$}(\tau),t} + \sigma_{\pi^{\$},t} \sigma_{H^{\$}(\tau),t}^{\top} + \frac{1}{\pi_t^{\$} H_t^{\$}(\tau)} \left(\lambda_{ct} \bar{\mathcal{J}}_c(\pi_t^{\$} H^{\$}(D_t^{\$}, q_t, \lambda_t, \tau)) \right.$$

$$\left. + \lambda_{cq,t} \left(\bar{\mathcal{J}}_{cq}(\pi_t^{\$} H^{\$}(D_t^{\$}, q_t, \lambda_t, \tau)) + \bar{\mathcal{J}}_q(\pi_t^{\$} H^{\$}(D_t^{\$}, q_t, \lambda_t, \tau)) \right) \right) = 0, \quad \text{(A.20)}$$

Proof See proof of Lemma A.1. □

Theorem A.2. *The function $H^{\$}$ takes an exponential form:*

$$H^{\$}(D_t^{\$}, q_t, \lambda_t, \tau) = D_t^{\$} \exp\left\{ a_{\phi^{\$}}(\tau) + b_{\phi^{\$}q}(\tau) q_t + b_{\phi^{\$}\lambda}(\tau)^{\top} \lambda_t \right\}, \quad \text{(A.21)}$$

where $b_{\phi^{\$}\lambda} = \left[b_{\phi^{\$}\lambda_c}, \ b_{\phi^{\$}\lambda_{cq}} \right]$. Function $b_{\phi^{\$}q}$ solves

$$\frac{db_{\phi^{\$}q}}{d\tau} = -\kappa_q b_{\phi^{\$}q}(\tau) - 1; \quad \text{(A.22)}$$

function $b_{\phi^{\$}\lambda_c}$ solves

$$\frac{db_{\phi^{\$}\lambda_c}}{d\tau} = \frac{1}{2} \sigma_{\lambda_c}^2 b_{L\lambda_c}(\tau)^2 + \left(b_c \sigma_{\lambda_c}^2 - \kappa_{\lambda_c} \right) b_{\phi^{\$}\lambda_c}(\tau)$$

$$+ p_D E_{\nu_c} \left[e^{(\phi_{D,c}^{\$} - \gamma) Z_{ct}} - e^{(1-\gamma) Z_{ct}} \right] + (1 - p_D) E_{\nu_c} \left[e^{-\gamma Z_{ct}} - e^{(1-\gamma) Z_{ct}} \right]; \quad \text{(A.23)}$$

function $b_{\phi^\$ \lambda_{cq}}$ *solves*

$$\frac{db_{\phi^\$ \lambda_{cq}}}{d\tau} = \frac{1}{2}\sigma_{\lambda_{cq}}^2 b_{\phi^\$ \lambda_{cq}}(\tau)^2 + \left(b_{cq}\sigma_{\lambda_{cq}}^2 - \kappa_{\lambda_{cq}}\right) b_{\phi^\$ \lambda_{cq}}(\tau) + E_{\nu_q}\left[e^{-b_{\phi^\$ q}(\tau))Z_{qt}} - 1\right]$$

$$+ p_D E_{\nu_{cq}}\left[e^{\left(\phi_{D,cq}^\$ - (\gamma + b_{\phi^\$ q}(\tau))\right)Z_{cq,t}} - e^{(1-\gamma)Z_{cq,t}}\right]$$

$$+ (1 - p_D)E_{\nu_{cq}}\left[e^{-(\gamma + b_{\phi^\$ q}(\tau))Z_{cq,t}} - e^{(1-\gamma)Z_{cq,t}}\right]; \quad \text{(A.24)}$$

and function a_L *solves*

$$\frac{da_{\phi^\$}}{d\tau} = \mu_D - \beta - \mu + \gamma\sigma(\sigma - \sigma_D) + \sigma_P^2 + \frac{1}{2}\sigma_q^2 b_{\phi^\$ q}(\tau)^2 + b_{\phi^\$ q}(\tau)\kappa_q\bar{q} + b_{\phi^\$ \lambda}(\tau)^\top(\kappa_\lambda * \bar{\lambda}). \quad \text{(A.25)}$$

The boundary conditions are $a_{\phi^\$}(0) = b_{\phi^\$ q}(0) = b_{\phi^\$ \lambda_c}(0) = b_{\phi^\$ \lambda_{cq}}(0) = 0.$

Proof It follows from Ito's Lemma that

$$\frac{dH_t^\$}{H_{t^-}^\$} = \mu_{H^\$,t}dt + \sigma_{H^\$,t}dB_t + \frac{1}{H_{t^-}^\$}\left(\mathcal{J}_c(H_t^\$) + \mathcal{J}_{cq}(H_t^\$) + \mathcal{J}_q(H_t^\$)\right),$$

where $\mu_{H^\$}$ and $\sigma_{H^\$}$ are given by

$$\mu_{H^\$,t} = \frac{1}{H^\$}\left(\frac{\partial H^\$}{\partial q}\left(\bar{q} - q_t\right) + \frac{\partial H^\$}{\partial \lambda_c}\left(\bar{\lambda}_c - \lambda_{ct}\right) + \frac{\partial H^\$}{\partial \lambda_{cq}}\left(\bar{\lambda}_c - \lambda_{cq,t}\right) - \frac{\partial H^\$}{\partial \tau}\right.$$

$$\left. + \frac{1}{2}\frac{\partial^2 H^\$}{\partial q_j^2}\sigma_q^2 + \frac{1}{2}\left(\frac{\partial^2 H^\$}{\partial \lambda_c^2}\sigma_{\lambda_c}^2 + \frac{\partial^2 H^\$}{\partial \lambda_c^2}\sigma_{\lambda_c}^2\right)\right)$$

$$= b_{\phi^\$ q}(\tau)\kappa_q\left(\bar{q} - q_t\right) + b_{\phi^\$ \lambda_c}(\tau)\kappa_{\lambda_c}\left(\bar{\lambda}_c - \lambda_{ct}\right) + b_{\phi^\$ \lambda_{cq}}(\tau)\kappa_{\lambda_{cq}}\left(\bar{\lambda}_{cq} - \lambda_{cq,t}\right)$$

$$+ \frac{1}{2}b_{\phi^\$ q}(\tau)^2\sigma_q^2 + \frac{1}{2}\left(b_{\phi^\$ \lambda_c}(\tau)^2\sigma_{\lambda_c}^2\lambda_{ct} + b_{\phi^\$ \lambda_{cq}}(\tau)^2\sigma_{\lambda_{cq}}^2\lambda_{cq,t}\right)$$

$$- \left(\frac{da_{\phi^\$}}{d\tau} + \frac{db_{\phi^\$ q}}{d\tau}q_t + \sum_j \frac{db_{\phi^\$ \lambda_j}}{d\tau}\lambda_{jt}\right),$$

$$\text{(A.26)}$$

and

$$\sigma_{H\$,t} = \frac{1}{L}\left(\frac{\partial H\$}{\partial q_t}[0,\ 0,\ \sigma_q\sqrt{q_t},\ 0,\ 0] + \frac{\partial H\$}{\partial \lambda_c}[0,\ 0,\ 0,\ \sigma_{\lambda_c}\sqrt{\lambda_{ct}},\ 0]\right. \tag{A.27}$$

$$\left. + \frac{\partial H\$}{\partial \lambda_{cq}}[0,\ 0,\ 0,\ 0,\ \sigma_{\lambda_{cq}}\sqrt{\lambda_{cq,t}}]\right)$$

$$= \left[0,\ 0,\ b_{\phi\$q}(\tau)\sigma_q\sqrt{q_t},\ b_{\phi\$\lambda_c}(\tau)\sigma_{\lambda_c}\sqrt{\lambda_{ct}},\ b_{\phi\$\lambda_{cq}}(\tau)\sigma_{\lambda_{cq}}\sqrt{\lambda_{cq,t}}\right]. \tag{A.28}$$

Furthermore,

$$\frac{\bar{\mathcal{J}}_c(\pi_t^\$ H_t^\$)}{\pi_t^\$ H_t^\$} = p_D E_{\nu_c}\left[e^{(\phi_{D,c}^\$ - \gamma)Z_{ct}} - 1\right] + (1 - p_D)E_{\nu_c}\left[e^{-\gamma Z_{ct}} - 1\right], \tag{A.29}$$

$$\frac{\bar{\mathcal{J}}_{cq}(\pi_t^\$ H_t^\$)}{\pi_t^\$ H_t^\$} = p_D E_{\nu_{cq}}\left[e^{\left(\phi_{D,cq}^\$ - (\gamma + b_{\phi\$q}(\tau))\right)Z_{cq,t}} - 1\right] + (1 - p_D)E_{\nu_{cq}}\left[e^{-(\gamma + b_{\phi\$q}(\tau))Z_{cq,t}} - 1\right], \tag{A.30}$$

and

$$\frac{\bar{\mathcal{J}}_q(\pi_t^\$ H_t^\$)}{\pi_t^\$ H_t^\$} = E_{\nu_q}\left[e^{-b_{\phi\$q}(\tau))Z_{qt}} - 1\right]. \tag{A.31}$$

Recall that $\lambda_q = \lambda_{cq}$. Substituting (A.26) – (A.30) along with (A.11) and (A.12) into the no-arbitrage condition (A.20) implies that functions $a_{\phi\$}$, $b_{\phi\$q}$, $b_{\phi\$\lambda_c}$, and $b_{\phi\$\lambda_{cq}}$ solve the following ordinary differential equation:

$$b_{\phi\$q}(\tau)\kappa_q\left(\bar{q} - q_t\right) + b_{\phi\$\lambda_c}(\tau)\kappa_{\lambda_c}\left(\bar{\lambda}_c - \lambda_{ct}\right) + b_{\phi\$\lambda_{cq}}(\tau)\kappa_{\lambda_{cq}}\left(\bar{\lambda}_{cq} - \lambda_{cq,t}\right)$$

$$+ \frac{1}{2}b_{\phi\$q}(\tau)^2\sigma_q^2 + \frac{1}{2}\left(b_{\phi\$\lambda_c}(\tau)^2\sigma_{\lambda_c}^2\lambda_{ct} + b_{\phi\$\lambda_c}(\tau)^2\sigma_{\lambda_c}^2\lambda_{ct}\right) - \beta - \mu + \gamma\sigma^2 - q_t + \sigma_P^2$$

$$+ b_{\phi\$\lambda_c}(\tau)b_j\sigma_{\lambda_c}^2\lambda_{ct} + b_{\phi\$\lambda_{cq}}(\tau)b_j\sigma_{\lambda_{cq}}^2\lambda_{cq,t} + p_D\lambda_{ct}E_{\nu_c}\left[e^{(\phi_{D,c}^\$ - \gamma)Z_{ct}} - e^{(1-\gamma)Z_{ct}}\right]$$

$$+ (1 - p_D)\lambda_{ct}E_{\nu_c}\left[e^{-\gamma Z_{ct}} - e^{(1-\gamma)Z_{ct}}\right] + p_D\lambda_{cq,t}E_{\nu_{cq}}\left[e^{\left(\phi_{D,cq}^\$ - (\gamma + b_{\phi\$q}(\tau))\right)Z_{cq,t}} - e^{(1-\gamma)Z_{cq,t}}\right]$$

$$+ (1 - p_D)\lambda_{cq,t}E_{\nu_{cq}}\left[e^{-(\gamma + b_{\phi\$q}(\tau))Z_{cq,t}} - e^{(1-\gamma)Z_{cq,t}}\right] + \lambda_{cq,t}E_{\nu_q}\left[e^{-b_{\phi\$q}(\tau))Z_{qt}} - 1\right]$$

$$- \left(\frac{da_{\phi\$}}{d\tau} + \frac{db_{\phi\$q}}{d\tau}q_t + \frac{db_{\phi\$\lambda_c}}{d\tau}\lambda_{ct} + \frac{db_{\phi\$\lambda_{cq}}}{d\tau}\lambda_{cq,t}\right) = 0. \quad \text{(A.32)}$$

Collecting q_t terms results in the following ordinary differential equation:

$$\frac{db_{\phi^\$q}}{d\tau} = -\kappa_q b_{\phi^\$q}(\tau) - 1;$$

collecting terms multiplying λ_c results in the following ordinary differential equation for $b_{\phi^\$\lambda_c}$

$$\frac{db_{\phi^\$\lambda_c}}{d\tau} = \frac{1}{2}\sigma_{\lambda_c}^2 b_{L\lambda_c}(\tau)^2 + \left(b_c\sigma_{\lambda_c}^2 - \kappa_{\lambda_c}\right) b_{\phi^\$\lambda_c}(\tau)$$
$$+ p_D E_{\nu_c}\left[e^{(\phi_{D,c}^\$ - \gamma)Z_{ct}} - e^{(1-\gamma)Z_{ct}}\right] + (1-p_D)E_{\nu_c}\left[e^{-\gamma Z_{ct}} - e^{(1-\gamma)Z_{ct}}\right];$$

collecting terms multiplying λ_{cq} results in the following ordinary differential equation for $b_{\phi^\$\lambda_{cq}}$

$$\frac{db_{\phi^\$\lambda_{cq}}}{d\tau} = \frac{1}{2}\sigma_{\lambda_{cq}}^2 b_{\phi^\$\lambda_{cq}}(\tau)^2 + \left(b_{cq}\sigma_{\lambda_{cq}}^2 - \kappa_{\lambda_{cq}}\right) b_{\phi^\$\lambda_{cq}}(\tau) + E_{\nu_q}\left[e^{-b_{\phi^\$q}(\tau))Z_{qt}} - 1\right]$$
$$+ p_D E_{\nu_{cq}}\left[e^{\left(\phi_{D,cq}^\$ - (\gamma + b_{\phi^\$q}(\tau))\right)Z_{cq,t}} - e^{(1-\gamma)Z_{cq,t}}\right]$$
$$+ (1-p_D)E_{\nu_{cq}}\left[e^{-(\gamma + b_{\phi^\$q}(\tau))Z_{cq,t}} - e^{(1-\gamma)Z_{cq,t}}\right];$$

and collecting constant terms results in the following ordinary differential equation for a_L:

$$\frac{da_{\phi^\$}}{d\tau} = \mu_D - \beta - \mu + \gamma\sigma(\sigma - \sigma_D) + \sigma_P^2 + \frac{1}{2}\sigma_q^2 b_{\phi^\$q}(\tau)^2 + b_{\phi^\$q}(\tau)\kappa_q\bar{q} + b_{\phi^\$\lambda}(\tau)^\top(\kappa_\lambda * \bar{\lambda}).$$

The boundary conditions are $a_{\phi^\$}(0) = b_{\phi^\$q}(0) = b_{\phi^\$\lambda_c}(0) = b_{\phi^\$\lambda_{cq}}(0) = 0.$ □

A.1.3. Nominal bond pricing

Proof of Corollary 1.5

$$y_t^{\$,(\tau)} = \frac{1}{\tau}\log\left(\frac{f_t^\$}{L_t^{\$,(\tau)}}\right),$$

where $L_t^{\$,(\tau)}$ is given by (1.14), then the results immediately follows. □

Proof of Theorem 1.6 By the no-arbitrage condition (A.20) and the definition of $\mu_{\pi^\$}$ (A.12), we can rewrite the premium in population (1.22) as

$$r_t^{\$,(\tau)} - r_t^\$ = -\sigma_{\pi^\$,t}\sigma_{L,t}^\top - \lambda_{ct}\left(\frac{\bar{\mathcal{J}}_c\left(\pi_t^\$ L_t^\$\right)}{\pi_t^\$ L_t^\$} - \frac{\bar{\mathcal{J}}_c(\pi_t^\$)}{\pi_t^\$} - \frac{\bar{\mathcal{J}}_c(L_t^\$)}{L_t^\$}\right)$$

$$- \lambda_{cq,t}\left(\frac{\bar{\mathcal{J}}_{cq}\left(\pi_t^\$ L_t^\$\right)}{\pi_t^\$ L_t^\$} - \frac{\bar{\mathcal{J}}_{cq}(\pi_t^\$)}{\pi_t^\$} - \frac{\bar{\mathcal{J}}_{cq}(L_t^\$)}{L_t^\$}\right) - \lambda_{qt}\left(\frac{\bar{\mathcal{J}}_q\left(\pi_t^\$ L_t^\$\right)}{\pi_t^\$ L_t^\$} - \frac{\bar{\mathcal{J}}_q(\pi_t^\$)}{\pi_t^\$} - \frac{\bar{\mathcal{J}}_q(L_t^\$)}{L_t^\$}\right).$$

From (A.10), we know that for $j \in \{c, cq\}$,

$$\frac{\bar{\mathcal{J}}_j(\pi_t^\$)}{\pi_t^\$} = E_{\nu_j}\left[e^{-\gamma Z_{jt}} - 1\right],$$

and $\frac{\bar{\mathcal{J}}_q(\pi_t^\$)}{\pi_t^\$} = 0$. Furthermore, recall that the N_q type of jump (inflation spike) does not affect $\pi^\$$, therefore; $\frac{\bar{\mathcal{J}}_q(\pi_t^\$ L_t^\$)}{\pi_t^\$ L_t^\$} = \frac{\bar{\mathcal{J}}_q(L_t^\$)}{L_t^\$}$. From (A.29) – (A.30) we know that

$$\frac{\bar{\mathcal{J}}_c\left(\pi_t^\$ L_t^\$\right)}{\pi_t^\$ L_t^\$} = p_D E_{\nu_c}\left[e^{(1-\gamma)Z_{ct}} - 1\right] + (1 - p_D)E_{\nu_c}\left[e^{-\gamma Z_{ct}} - 1\right],$$

$$\frac{\bar{\mathcal{J}}_{cq}\left(\pi_t^\$ L_t^\$\right)}{\pi_t^\$ L_t^\$} = p_D E_{\nu_{cq}}\left[e^{1-(\gamma+b_{L^\$ q}(\tau))Z_{cq,t}} - 1\right] + (1 - p_D)E_{\nu_{cq}}\left[e^{-(\gamma+b_{L^\$ q}(\tau))Z_{cq,t}} - 1\right].$$

Furthermore,

$$\frac{\bar{\mathcal{J}}_c\left(L_t^\$\right)}{L_t^\$} = p_D E_{\nu_c}\left[e^{Z_{ct}} - 1\right],$$

$$\frac{\bar{\mathcal{J}}_{cq}\left(L_t^\$\right)}{L_t^\$} = p_D E_{\nu_{cq}}\left[e^{1-b_{L^\$ q}(\tau)Z_{cq,t}} - 1\right] + (1 - p_D)E_{\nu_{cq}}\left[e^{-b_{L^\$ q}(\tau)Z_{cq,t}} - 1\right].$$

Together with (A.11) and (A.28), we obtain:

$$r_t^{\$(\tau)} - r_t^\$ = -\lambda_t^\top \left(b_{L^\$\lambda}(\tau) * b * \sigma_\lambda^2 \right) + \lambda_c p_D\, E_{\nu_c} \left[(e^{-\gamma Z_{ct}} - 1)(1 - e^{Z_{ct}}) \right]$$

$$+ \lambda_{cq} \Bigg((1 - p_D) E_{\nu_{cq}} \left[(e^{-\gamma Z_{cq,t}} - 1)(1 - e^{-b_{L^\$q}(\tau) Z_{cq,t}}) \right]$$

$$+ p_D E_{\nu_{cq}} \left[(e^{-\gamma Z_{cq,t}} - 1)(1 - e^{(1 - b_{L^\$q}(\tau)) Z_{cq,t}}) \right] \Bigg).$$

□

A.2. Appendix for Rare Booms and Disasters in a Multi-sector Endowment Economy

A.2.1. Required conditions on the parameters

Assumption A.2.

$$\left(\kappa_{\lambda_j} + \beta\right)^2 \geq 2\sigma_{\lambda_j}^2 E_{\nu_1}\left[e^{b_{\mu_j}Z_j} - 1\right] \qquad j = 1, 2.$$

Assumption A.3.

$$(b_{\lambda_2}\sigma_{\lambda_2}^2 - \kappa_\lambda 2)^2 \geq 2\sigma_{\lambda_2}^2 E_{\nu_2}\left[e^{b_{\mu_2}Z_2}\left(e^{\frac{\phi-1}{\kappa_\mu 2}Z_2} - 1\right)\right].$$

Assumption A.4.

$$\bar{\mu}_D - \bar{\mu}_C - \beta + \gamma\sigma^2\left(1 - \phi\right) - \sum_j \frac{\kappa_{\lambda_j}\bar{\lambda}_j}{\sigma_{\lambda_j}^2}\left(\zeta_{\phi_j} - \kappa_{\lambda_j} + b_{\lambda_j}\sigma_{\lambda_j}^2\right) < 0,$$

where

$$\zeta_{\phi_j} = \sqrt{(b_{\lambda_j}\sigma_{\lambda_j}^2 - \kappa_{\lambda_j})^2 - 2E_{\nu_j}\left[e^{b_{\mu_j}Z_j}\left(e^{\frac{\phi-1}{\kappa_{\mu_j}}Z_j} - 1\right)\right]\sigma_{\lambda_j}^2}.$$

Assumption A.2 is required for the solution for $J(W_t, \mu_t, \lambda_t)$ to be real-valued. Assumption A.3 is required for $b_{\phi\lambda_2}(\tau)$ to converge as τ approaches infinity. Without this assumption, the price-dividend ratio market does not have a finite solution. Note that the analogous condition for $j = 1$ is satisfied automatically because $Z_1 < 0$ and hence $e^{\frac{\phi-1}{\kappa_\mu 1}Z_1} < 1$. Furthermore, the analogous condition for the value claim is satisfied automatically; this condition replaces $e^{\frac{\phi-1}{\kappa_\mu 2}Z_2}$ with $e^{-\frac{1}{\kappa_\mu 2}Z_2}$ which is less than one. Assumption A.4 states that the asymptotic slope of $a_\phi(\tau)$ is negative. This is required for convergence of the price-dividend ratio on the market. If this condition is satisfied, the analogous condition for the value function is satisfied automatically.[1]

---

[1]Specifically, define

$$\zeta_{\phi_2}^v = \sqrt{(b_{\lambda_2}\sigma_{\lambda_2}^2 - \kappa_\lambda 2)^2 - 2E_{\nu_2}\left[e^{b_{\mu_2}Z_2}\left(e^{-\frac{1}{\kappa_\mu 2}Z_2} - 1\right)\right]\sigma_{\lambda_2}^2}$$

A.2.2. Detailed derivation of the model

This Appendix derives the results given in the main text. The derivations generalize those in Wachter (2012), where there is a single disaster probability, and the shocks are to realized consumption growth. In what follows, there are two time-varying jump probabilities, and, more importantly, the jumps are in expected consumption growth. Like the results in the earlier paper, the derivations here assume that the EIS parameter is equal to one, and, based on this assumption, lead to solutions that are in closed-form up to a system of ordinary differential equations.[2]

Notation

Let X_t be a pure diffusion process, and let μ_{jt}, $j = 1, 2$ be defined as above. Consider a scalar, real-valued function $h(\mu_{1t}, \mu_{2t}, X_t)$. Define

$$
\begin{aligned}
\mathcal{J}_1(h(\mu_{1t}, \mu_{2t}, X_t)) &= h(\mu_1 + Z_1, \mu_2, X_t) \\
\mathcal{J}_2(h(\mu_{1t}, \mu_{2t}, X_t)) &= h(\mu_1, \mu_2 + Z_2, X_t)
\end{aligned}
$$

Further, define

$$
\bar{\mathcal{J}}_j(h(\mu_{1t}, \mu_{2t}, X_t)) = E_{\nu_j} \mathcal{J}_j(h(\mu_{1t}, \mu_{2t}, X))
$$

for $j = 1, 2$, and

$$
\bar{\mathcal{J}}(h(\mu_{1t}, \mu_{2t}, X_t)) = \left[\bar{\mathcal{J}}_1(h(\mu_{1t}, \mu_{2t}, X_t)), \; \bar{\mathcal{J}}_2(h(\mu_{1t}, \mu_{2t}, X_t)) \right]^\top .
$$

In what follows, we will use the notation $*$ to denote element-by-element multiplication for two vectors of equal length. We will use x^2 notation for a vector x to denote the square of each element in x. For example, σ_λ^2 will denote the vector $[\sigma_{\lambda_1}^2, \, \sigma_{\lambda_2}^2]^\top$.

Then $\zeta_{\phi_2}^v > \zeta_{\phi_2}$.

[2]Using log-linearization, Eraker and Shaliastovich (2008) and Benzoni et al. (2011) find approximate solutions to related continuous-time jump-diffusion models when the EIS is not equal to one.

Finally, because the process λ are independent, the second cross-partial derivatives do not enter into equations that determine the price. Given a function $h(\lambda, X)$, we will will use the notation $\partial h / \partial \lambda$ to denote the 1×2 vector $[\partial^2 h / \partial \lambda_1^2, \; \partial^2 h / \partial \lambda_2^2]$.

The value function

Proof of Theorem 2.1 Let S denote the value of a claim to aggregate consumption, and conjecture that the price-dividend ratio for the consumption claim is constant:

$$\frac{S_t}{C_t} = l,$$

for some constant l. This relation implies that S_t satisfies

$$\frac{dS_t}{S_t} = \frac{dC_t}{C_t} = \mu_{Ct} \, dt + \sigma \, dB_{Ct}. \qquad (A.33)$$

Consider an agent who allocates wealth between S and the risk-free asset. Let α_t be the fraction of wealth in the risky asset S_t, and let c_t be the agent's consumption. The wealth process is then given by

$$dW_t = \left(W_t \alpha_t \left(\mu_{Ct} - r_t + l^{-1} \right) + W_t r_t - c_t \right) dt + W_t \alpha_t \sigma dB_{ct},$$

where r_t denote the instantaneous risk-free rate. Optimal consumption and portfolio choice must satisfy the following Hamilton-Jacobi-Bellman equation:

$$\sup_{\alpha_t, c_t} \left\{ \frac{\partial J}{\partial W} \left(W_t \alpha_t \left(\mu_{Ct} - r_t + l^{-1} \right) + W_t r_t - c_t \right) + \frac{\partial J}{\partial \lambda} \left(\kappa_\lambda * \left(\bar{\lambda} - \lambda_t \right) \right) - \frac{\partial J}{\partial \mu} \left(\kappa_\mu * \mu_t \right) \right.$$

$$\left. + \frac{1}{2} \frac{\partial^2 J}{\partial W^2} W_t^2 \alpha_t^2 \sigma^2 + \frac{1}{2} \left(\frac{\partial^2 J}{\partial \lambda^2} \right)^\top \left(\sigma_\lambda^2 * \lambda_t \right) + \lambda_t^\top \bar{\mathcal{J}}(J(W_t, \mu_t, \lambda_t)) + f\left(c_t, V \right) \right\} = 0, \quad (A.34)$$

where, as defined in Appendix A.2.2,

$$\frac{\partial^2 J}{\partial \lambda^2} = \left[\frac{\partial^2 J}{\partial \lambda_1^2}, \frac{\partial^2 J}{\partial \lambda_2^2}\right]^\top$$

$$\sigma_\lambda^2 = \left[\sigma_{\lambda_1}^2, \sigma_{\lambda_2}^2\right]^\top.$$

In equilibrium, $\alpha_t = 1$ and $c_t = C_t = W_t l^{-1}$. Substituting these policy functions into (A.34) implies

$$\frac{\partial J}{\partial W} W_t \mu_{Ct} + \frac{\partial J}{\partial \lambda} \left(\kappa_\lambda * (\bar{\lambda} - \lambda_t)\right) - \frac{\partial J}{\partial \mu} \left(\kappa_\mu * \mu_t\right) + \frac{1}{2}\frac{\partial^2 J}{\partial W^2} W_t^2 \sigma^2$$

$$+ \frac{1}{2}\left(\frac{\partial^2 J}{\partial \lambda^2}\right)^\top (\sigma_\lambda^2 * \lambda_t) + \lambda_t^\top \bar{\mathcal{J}}(J(W_t, \mu_t, \lambda_t)) + f(C_t, V) = 0. \quad \text{(A.35)}$$

By the envelope condition $\partial f / \partial C = \partial J / \partial W$, we obtain $\beta = l^{-1}$. Given that the consumption-wealth ratio equals β^{-1}, it follows that

$$f(C_t, V_t) = f\left(W_t l^{-1}, J\left(W_t, \mu_t, \lambda_t\right)\right)$$

$$= \beta W_t^{1-\gamma} I\left(\mu_t, \lambda_t\right)\left(\log \beta - \frac{\log I\left(\mu_t, \lambda_t\right)}{1-\gamma}\right). \quad \text{(A.36)}$$

Substituting (A.36) and (2.7) into (A.35)

$$\mu_{Ct} + (1-\gamma)^{-1} I^{-1} \frac{\partial I}{\partial \lambda} \left(\kappa_\lambda * (\bar{\lambda} - \lambda_t)\right) - (1-\gamma)^{-1} I^{-1} \frac{\partial I}{\partial \mu} \left(\kappa_\mu * \mu_t\right) - \frac{1}{2}\gamma\sigma^2$$

$$+ \frac{1}{2}(1-\gamma)^{-1} I^{-1} \left(\frac{\partial^2 I}{\partial \lambda^2}\right)^\top (\sigma_\lambda^2 * \lambda_t) + (1-\gamma)^{-1} \lambda_t^\top \bar{\mathcal{J}}(I(\mu_t, \lambda_t))$$

$$+ \beta \left(\log \beta - \frac{\log I\left(\mu_t, \lambda_t\right)}{1-\gamma}\right) = 0.$$

Note that $\mu_{Ct} = \bar{\mu}_C + \mu_{1t} + \mu_{2t}$.

Collecting coefficients on μ_{jt} results in the following equation for b_{μ_j}:

$$1 - (1-\gamma)^{-1} b_{\mu_j} \kappa_{\mu_j} - \beta(1-\gamma)^{-1} b_\mu = 0,$$

solving this equation yields

$$b_{\mu_j} = \frac{1 - \gamma}{\kappa_{\mu_j} + \beta}.$$

Collecting coefficients on λ_{jt} yields

$$b_{\lambda_j} = \frac{\beta + \kappa_{\lambda_j}}{\sigma_{\lambda_j}^2} - \sqrt{\left(\frac{\beta + \kappa_{\lambda_j}}{\sigma_{\lambda_j}^2}\right)^2 - \frac{2E_{\nu_j}\left[e^{b_{\mu_j} Z_{jt}} - 1\right]}{\sigma_{\lambda_j}^2}}.$$

Collecting the constant terms:

$$a = \frac{1 - \gamma}{\beta}\left(\bar{\mu}_C - \frac{1}{2}\gamma\sigma^2\right) + (1 - \gamma)\log\beta + \sum_j b_{\lambda_j}\frac{\kappa_{\lambda_j}}{\beta}\bar{\lambda}_j.$$

$\square$

Proof of Corollary 2.2 The risk-free rate is obtained by taking the derivative of the HJB (A.34) with respect to α_t, evaluating at $\alpha_t = 1$ and setting it equal to 0. The result immediately follows. $\square$

The state-price density

Duffie and Skiadas (1994) show that the state-price density π_t equals

$$\pi_t = \exp\left\{\int_0^t \frac{\partial}{\partial V} f(C_s, V_s)\, ds\right\} \frac{\partial}{\partial C} f(C_t, V_t). \tag{A.37}$$

Note that the exponential term is deterministic. From (2.6), we obtain

$$\frac{\partial}{\partial C} f(C_t, V_t) = \beta(1 - \gamma)\frac{V_t}{C_t}.$$

The equilibrium condition $V_t = J\left(\beta^{-1}C_t, \mu_t, \lambda_t\right)$, together with the form of the value function (2.7), implies

$$\frac{\partial}{\partial C} f(C_t, V_t) = \beta^\gamma C_t^{-\gamma} I(\mu_t, \lambda_t). \tag{A.38}$$

Applying Ito's Lemma to (A.38) implies

$$\frac{d\pi_t}{\pi_{t^-}} = \mu_{\pi t}dt + \sigma_{\pi t}dB_t + \sum_j \frac{\mathcal{J}_j(\pi_t)}{\pi_{t^-}}dN_{jt},\tag{A.39}$$

where

$$\sigma_{\pi t} = \left[-\gamma\sigma,\ b_{\lambda_1}\sigma_{\lambda_1}\sqrt{\lambda_{1t}},\ b_{\lambda_2}\sigma_{\lambda_2}\sqrt{\lambda_{2t}}\right],\tag{A.40}$$

and

$$\mathcal{J}_j(\pi_t) = e^{b_{\mu_j}Z_{jt}} - 1,\tag{A.41}$$

for $j = 1, 2$. It also follows from no-arbitrage that

$$\mu_{\pi t} = -r_t - \lambda_t^\top \frac{\bar{\mathcal{J}}(\pi_t)}{\pi_t}\tag{A.42}$$

$$= -r_t - \sum_j \lambda_{jt}E_{\nu_j}\left[e^{b_{\mu_j}Z_{jt}} - 1\right]$$

$$= -\beta - \mu_{Ct} + \gamma\sigma^2 - \sum_j \lambda_{jt}E_{\nu_j}\left[e^{b_{\mu_j}Z_{jt}} - 1\right].\tag{A.43}$$

In the event of a disaster, marginal utility (as represented by the state-price density) jumps upward, and in the event of a boom the marginal utility jumps downward, as can be seen by the term multiplying the Poisson process in (A.39). The first element of (A.40) implies that the standard diffusion risk in consumption is priced; more interestingly, changes in λ_{jt} are also priced as reflected by the new element of (A.40).

Pricing the general equity claim

We first consider the price of a general form of the dividend stream. The dividend stream on the aggregate market and the dividend stream for value will be special cases. Suppose dividends evolve according to

$$\frac{dD_t}{D_t} = \mu_{Dt}\,dt + \sigma_D\,dB_{Ct},\tag{A.44}$$

where

$$\mu_{Dt} = \bar{\mu}_D + \phi_{D,1}\mu_{1t} + \phi_{D,2}\mu_{2t},$$

$\phi_{D,j}$ denotes the jump multiplier for the type-j jump.

Lemma A.3. *Let $H\left(D_t, \mu_t, \lambda_t, \tau\right)$ denote the time t price of a single future dividend payment at time $t + \tau$:*

$$H(D_t, \mu_t, \lambda_t, \tau) = E_t\left[\frac{\pi_{t+\tau}}{\pi_t} D_{t+\tau}\right].$$

By Ito's Lemma, we can write

$$\frac{dH_t}{H_t} = \mu_{H(\tau),t}dt + \sigma_{H(\tau),t}dB_t + \sum_j \mathcal{J}_j(H_t)dN_{jt}.$$

for a scalar process $\mu_{H(\tau),t}$ and a vector process $\sigma_{H(\tau),t}$, where $H_t = H(D_t, \mu_t, \lambda_t, \tau)$. Then no-arbitrage implies that

$$\mu_{\pi,t} + \mu_{H(\tau),t} + \sigma_{\pi,t}\sigma_{H(\tau),t}^\top + \frac{1}{\pi_t H_t}\lambda_t^\top \bar{\mathcal{J}}(\pi_t H_t) = 0. \tag{A.45}$$

Proof No-arbitrage implies that $H(D_s, \lambda_s, \mu_s, 0) = D_s$ and that

$$\pi_t H(D_t, \lambda_t, \mu_t, \tau) = E_t\left[\pi_s H(D_s, \lambda_s, \mu_s, 0)\right].$$

For the remainder of the argument, we simplify notation by writing $H_t = H(D_t, \mu_t, \lambda_t, \tau)$, $\mu_{H,t} = \mu_{H(\tau),t}$ and $\sigma_{H,t} = \sigma_{H(\tau),t}$. Ito's Lemma applied to $\pi_t H_t$ implies

$$\pi_t H_t = \pi_0 H_0 + \int_0^t \pi_s H_s\left(\mu_{H,s} + \mu_{\pi,s} + \sigma_{\pi,s}\sigma_{H,s}^\top\right) + \int_0^t \pi_s H_s(\sigma_{H,s} + \sigma_{\pi,s})dB_s$$

$$+ \sum_j \sum_{0 < s_{ij} \le t}\left(\pi_{s_{ij}} H_{s_{ij}} - \pi_{s_{ij}^-} H_{s_{ij}^-}\right), \quad \text{(A.46)}$$

where $s_{ij} = q\{s : N_{js} = i\}$ (namely, the time that the ith type j jump occurs). Adding and

subtracting the jump compensation term from (A.46) yields:

$$\pi_t H_t = \pi_0 H_0 + \int_0^t \pi_s H_s \left(\mu_{H,s} + \mu_{\pi,s} + \sigma_{\pi,s}\sigma_{H,s}^\top + \sum_j \lambda_j \frac{\bar{\mathcal{J}}_j(\pi_s H_s)}{\pi_s H_s} \right) ds$$

$$+ \int_0^t \pi_s H_s (\sigma_{H,s} + \sigma_{\pi,s}) dB_s$$

$$+ \sum_j \left(\sum_{0 < s_{ij} \le t} \left(\pi_{s_{ij}} H_{s_{ij}} - \pi_{s_{ij}^-} H_{s_{ij}^-} \right) - \int_0^t \pi_s H_s \lambda_j \bar{\mathcal{J}}_j(\pi_s H_s) ds \right). \quad \text{(A.47)}$$

Under regularity conditions analogous to those given in Duffie et al. (2000) the second and the third integrals on the right hand side of (A.47) are martingales. Therefore the first integral on the right hand side of (A.47) must also be a martingale, and it follows that the integrand of this term must equal zero. □

Theorem A.3. *The function H takes an exponential form:*

$$H(D_t, \mu_t, \lambda_t, \tau) = D_t \exp \left\{ a_\phi(\tau) + b_{\phi\mu}(\tau)^\top \mu_t + b_{\phi\lambda}(\tau)^\top \lambda_t \right\}, \quad \text{(A.48)}$$

where $b_{\phi\mu} = [b_{\phi\mu_1}, \ b_{\phi\mu_2}]^\top$ and $b_{\phi\lambda} = [b_{\phi\lambda_1}, \ b_{\phi\lambda_2}]^\top$ and

$$\frac{db_{\phi\mu_j}}{d\tau} = -\kappa_{\mu_j} b_{\phi_j\mu} + (\phi_{D,j} - 1), \quad \text{(A.49)}$$

$$\frac{db_{\phi\lambda_j}}{d\tau} = \frac{1}{2}\sigma_{\lambda_j}^2 b_{\phi\lambda_j}(\tau)^2 + \left(b_{\lambda_j}\sigma_{\lambda_j}^2 - \kappa_{\lambda_j} \right) b_{\phi\lambda_j}(\tau) + E_{\nu_j} \left[e^{b_{\mu_j} Z_{jt}} \left(e^{b_{\phi\mu_j}(\tau)Z_{jt}} - 1 \right) \right], \quad \text{(A.50)}$$

$$\frac{da_\phi}{d\tau} = \bar{\mu}_D - \bar{\mu}_C - \beta + \gamma\sigma(\sigma - \sigma_D) + b_{\phi\lambda}(\tau)^\top (\kappa_\lambda * \bar{\lambda}). \quad \text{(A.51)}$$

The boundary conditions are $b_{\phi\mu_j}(0) = b_{\phi\lambda_j}(0) = a_\phi(0) = 0$.

Proof Let $H_t = H(D_t, \mu_t, \lambda_t, \tau)$. It follows from Ito's Lemma that

$$\frac{\bar{\mathcal{J}}_j(\pi_t H_t)}{\pi_t H_t} = E_{\nu_j} \left[e^{\left(b_{\mu_j} + b_{\mu_j\phi}(\tau) \right) Z_{jt}} - 1 \right], \quad \text{(A.52)}$$

$$\mu_{H(\tau),t} = \frac{1}{H}\left(\frac{\partial H}{\partial D}\mu_{Dt} + \frac{\partial H}{\partial \lambda}\left(\kappa_\lambda * \left(\bar{\lambda} - \lambda_t\right)\right) - \frac{\partial H}{\partial \mu}\left(\kappa_\mu * \mu_t\right)\right.$$
$$\left. - \frac{\partial H}{\partial \tau} + \frac{1}{2}\left(\frac{\partial^2 H}{\partial \lambda^2}\right)\left(\sigma_\lambda^2 * \lambda_t\right)\right) \quad \text{(A.53)}$$

$$= \mu_{Dt} + b_{\phi\lambda}(\tau)^\top\left(\kappa_\lambda * \left(\bar{\lambda} - \lambda_t\right)\right) + b_{\phi\mu}(\tau)^\top\left(\kappa_\mu * \mu_t\right)$$
$$- \left(\frac{da_\phi}{d\tau} + \lambda_t^\top\frac{db_{\phi\lambda}}{d\tau} + \mu_t^\top\frac{db_{\phi\mu}}{d\tau}\right) + \frac{1}{2}\left(b_{\phi\lambda}(\tau)^2\right)^\top\left(\sigma_\lambda^2 * \lambda_t\right), \quad \text{(A.54)}$$

and

$$\sigma_{H(\tau),t} = \frac{1}{H}\left(\frac{\partial H}{\partial D}\mu_D[\sigma_D,\ 0,\ 0] + \frac{\partial H}{\partial \lambda_1}[0,\ \sigma_{\lambda_1}\sqrt{\lambda_{1t}},\ 0] + \frac{\partial H}{\partial \lambda_2}[0,\ 0,\ \sigma_{\lambda_2}\sqrt{\lambda_{2t}}]\right)$$
$$= \left[\sigma_D,\ b_{\phi\lambda_1}(\tau)\sigma_{\lambda_1}\sqrt{\lambda_{1t}},\ b_{\phi\lambda_2}(\tau)\sigma_{\lambda_2}\sqrt{\lambda_{2t}}\right]. \quad \text{(A.55)}$$

Substituting (A.52), (A.54) and (A.55) along with (A.40) and (A.43) into the no-arbitrage condition (A.45) implies

$$\mu_{Dt} + b_{\phi\lambda}(\tau)^\top\left(\kappa_\lambda * \left(\bar{\lambda} - \lambda_t\right)\right) + b_{\phi\mu}(\tau)^\top\left(\kappa_\mu * \mu_t\right) + \left(b_{\phi\lambda}(\tau)^2\right)^\top\left(\sigma_\lambda^2 * \lambda_t\right)$$
$$- \beta - \mu_{Ct} + \gamma\sigma^2 - \gamma\sigma\sigma_D + \sum_j \lambda_{jt}E_{\nu_j}\left[e^{\left(b_{\mu_j} + b_{\phi\mu_j}(\tau)\right)Z_{jt}} - e^{b_{\mu_j}Z_{jt}}\right]$$
$$- \left(\frac{da_\phi}{d\tau} + \lambda_t^\top\frac{db_{\phi\lambda}}{d\tau} + \mu_t^\top\frac{db_{\phi\mu}}{d\tau}\right) = 0.$$

Notice that, by definition, $\mu_{Dt} - \mu_{Ct} = (\bar{\mu}_D - \bar{\mu}_C) + \sum_j(\phi_{D,j} - 1)\mu_{jt}$. Matching the terms multiplying μ_j implies (A.49), matching the terms multiplying λ_j implies (A.50) and matching the constant terms implies (A.51). □

Let $F_t = F(D_t, \mu_t, \lambda_t)$ denote the time t price of the claim to the dividend stream defined by (A.44).

Lemma A.4. *No-arbitrage implies*

$$\mu_{\pi,t} + \mu_{F,t} + \frac{D_t}{F_t} + \sigma_{\pi,t}\sigma_{F,t}^\top + \sum_j \lambda_{jt}\frac{\bar{\mathcal{J}}_j(\pi_t F_t)}{\pi_t F_t} = 0, \quad \text{(A.56)}$$

where $\mu_{F,t}$ and $\sigma_{F,t}$ denote the drift and diffusion term of the F_t process, respectively.

Proof By definition,

$$F(D_t, \mu_t, \lambda_t) = \int_0^\infty H(D_t, \mu_t, \lambda_t, \tau) \, d\tau.$$

For notational simplicity, we abbreviate $H(D, \lambda, \mu, \tau)$ as $H(\tau)$. It follows from Ito's Lemma applied to $F(D_t, \mu_t, \lambda_t)$ that

$$F(D_t, \mu_t, \lambda_t)\mu_{F,t} = \int_0^\infty \left(H(\tau)\mu_{Dt} + \sum_j H_{\lambda_j}(\tau)(\bar{\lambda}_j - \lambda_j) + \sum_j H_{\mu_j}(\tau)\mu_j + \frac{1}{2}\sum_j H_{\lambda_j\lambda_j}(\tau) \right) d\tau,$$

where H_{μ_j}, H_{λ_j} and $H_{\lambda_j\lambda_j}$ denote partial derivatives. It then follows from the equation for $\mu_{H(\tau),t}$ (A.53) that

$$F(D_t, \mu_t, \lambda_t)\mu_{F,t} = \int_0^\infty \left(H(D_t, \lambda_t, \mu_t, \tau)\mu_{H(\tau),t} - \frac{\partial}{\partial\tau}H(D_t, \mu_t\lambda_t, \tau) \right) d\tau. \qquad (A.57)$$

In short, (A.57) holds because H is a function of τ but F is not.

Because $\lim_{\tau\to\infty} H(D_t, \mu_t, \lambda_t, \tau) = 0$,

$$-\int_0^\infty \frac{\partial}{\partial\tau}H(D_t, \mu_t, \lambda_t, \tau) \, d\tau = H(D_t, \mu_t, \lambda_t, 0) = D_t.$$

Ito's Lemma also implies

$$F(D_t, \mu_t, \lambda_t)\sigma_{F,t} = \int_0^\infty H(D_t, \mu_t, \lambda_t, \tau)\sigma_{H(\tau),t} \, d\tau$$

and

$$\bar{J}(\pi_t F(D_t, \mu_t, \lambda_t)) = \int_0^\infty \bar{J}(\pi_t H(D_t, \mu_t, \lambda_t, \tau)) \, d\tau$$

The result then follows from the no-arbitrage relation for H, (A.45). $\qquad\qquad \square$

Given a stream of cash flows D_t and its price F_t, define the expected return on this claim to be

$$r_t^e = \mu_{F,t} + \frac{D_t}{F_t} + \frac{1}{F_t}\lambda_t^\top \bar{\mathcal{J}}(F_t).$$

Theorem A.4. *Let r_t^e denote the instantaneous expected return on the general equity claim. Then*

$$r_t^e - r_t = -\sigma_{\pi,t}\sigma_{F,t}^\top - \sum_j \lambda_{jt} E_{\nu_j}\left[\frac{\mathcal{J}_j(F_t)}{F_t}\frac{\mathcal{J}_j(\pi_t)}{\pi_t}\right]. \tag{A.58}$$

Proof It follows from the definition of r_t^e (2.21) that

$$\mu_{F,t} + \frac{D_t}{F_t} = r_t^e - \frac{1}{F_t}\lambda_t^\top \bar{\mathcal{J}}(F_t).$$

Further, $\mu_{\pi t}$ can be written in terms of r_t and a jump term as in (A.42). Finally,

$$E_{\nu_j}\left[\frac{\mathcal{J}_j(F_t)}{F_t}\frac{\mathcal{J}_j(\pi_t)}{\pi_t}\right] = \bar{\mathcal{J}}_j(F_t\pi_t) - \bar{\mathcal{J}}_j(F_t) - \bar{\mathcal{J}}_j(\pi_t)$$

for $j = 1, 2$. The result that follows from rearranging (A.56) in Lemma A.4. $\square$

Further results on equity pricing

The following is an intermediate step in the proof of Corollary 2.4:

Lemma A.5.

$$\lim_{\tau\to\infty} b_{\phi\lambda_j}(\tau) = -\frac{1}{\sigma_{\lambda_j}^2}\left(\zeta_{\phi_j} - \kappa_{\lambda_j} + b_{\lambda_j}\sigma_{\lambda_j}^2\right), \tag{A.59}$$

where

$$\zeta_{\phi_j} = \sqrt{(b_{\lambda_j}\sigma_{\lambda_j}^2 - \kappa_{\lambda_j})^2 - 2E_{\nu_j}\left[e^{\left(b_{\mu_j}+\frac{\phi-1}{\kappa\mu_j}\right)Z_j} - e^{b_{\mu_j}Z_j}\right]\sigma_{\lambda_j}^2}. \tag{A.60}$$

Moreover, $\lim_{\tau\to\infty} b_{\phi\lambda_1}(\tau) < 0$ and $\lim_{\tau\to\infty} b_{\phi\lambda_2}(\tau) > 0$.

Proof Let $\bar{b}_{\phi\lambda_j}$ denote the limit, should it exist. In the limit, small changes in τ do not change $b_{\phi\lambda_j}(\tau)$. Taking the limit of both sides of (2.17) implies that $\bar{b}_{\phi\lambda_j}$ must satisfy the

quadratic equation

$$0 = \frac{1}{2}\sigma_{\lambda_j}^2 \bar{b}_{\phi\lambda_j}^2 + (b_{\lambda_j}\sigma_{\lambda_j}^2 - \kappa_{\lambda_j})\bar{b}_{\phi\lambda_j} + E_{\nu_j}\left[e^{\left(b_{\mu_j} + \frac{\phi-1}{\kappa_{\mu j}}\right)Z_{jt}} - e^{b_{\mu_j}Z_{jt}}\right]$$

This equation has two solutions; as for the value function, the solution corresponding to the negative root has the more reasonable economic properties and is given in (A.59).[3]

To prove that the limits have the signs given in the Lemma, note that $Z_1 < 0$ implies that

$$E_{\nu_1}\left[e^{\left(b_{\mu_1} + \frac{\phi-1}{\kappa_{\mu 1}}\right)Z_1} - e^{b_{\mu_1}Z_1}\right] < 0.$$

Therefore,

$$\zeta_{\phi_1} > |b_{\lambda_1}\sigma_{\lambda_1}^2 - \kappa_\lambda 1|.$$

Now, note that $Z_2 > 0$ implies that

$$E_{\nu_1}\left[e^{\left(b_{\mu_1} + \frac{\phi-1}{\kappa_{\mu 1}}\right)Z_1} - e^{b_{\mu_1}Z_1}\right] > 0.$$

The parameter assumptions imply that ζ_{ϕ_2} is real-valued. As shown in Corollary 2.1, $b_{\lambda_2} < 0$, and that

$$\zeta_{\phi_2} < |b_{\lambda_2}\sigma_{\lambda_2}^2 - \kappa_\lambda 2|$$

In both cases the result on the sign follows. $\qquad\qquad\square$

Proof of Corollary 2.4 The result for μ_{jt} follows immediately from the form of $b_{\phi\mu_j}(\tau)$. For λ_{1t}, first note that $b_{\phi\lambda_1}(0) = 0$ and $\lim_{\tau\to\infty} b_{\phi\lambda_1}(\tau) < 0$ by Lemma A.5. Therefore, it suffices to show that $b_{\phi\lambda_1}(\tau)$ is a monotonic function of τ.

---

[3]We have verified that (A.59) does indeed correspond to the limit when the ordinary differential equation (2.17) is solved numerically.

Assume, by contradiction that $db_{\phi\lambda_1}(\tau)/d\tau = 0$ for some τ, τ^*. Then, by (2.17),

$$b_{\phi\lambda_1}(\tau^*) = \frac{1}{\sigma_{\lambda_1}^2}\left(\sqrt{(b_{\lambda_1}\sigma_{\lambda_1}^2 - \kappa_\lambda 1)^2 - 2E_{\nu_1}\left[e^{\left(b_{\mu_1}+b_{\phi\mu_1}(\tau^*)\right)Z_1} - e^{b_{\mu_1}Z_1}\right]\sigma_{\lambda_1}^2} - \kappa_\lambda 1 + b_{\lambda_1}\sigma_{\lambda_1}^2\right)$$

$$(A.61)$$

However, differentiating (A.61) with respect to τ implies $db_{\phi\lambda_1}(\tau^*)/d\tau \neq 0$. Therefore, $db_{\phi\lambda_1}(\tau)/d\tau$ must be nonzero for all finite τ, and, because (2.17) implies that the derivative is a continuous function, it must be either (weakly) positive or negative. It follows that $b_{\phi\lambda_1}(\tau)$ is monotonic, and, by the argument given above, it must be negative and decreasing in τ. Analogous reasoning holds for $j = 2$. $\qquad\square$

Proof of Corollary 2.5 It follows from Ito's Lemma and the definition of G that

$$\sigma_{F,t} = \left[\phi\sigma_D, \ \frac{1}{G}\frac{\partial G}{\partial\lambda_1}\sigma_{\lambda_1}\sqrt{\lambda_{1t}}, , \ \frac{1}{G}\frac{\partial G}{\partial\lambda_2}\sigma_{\lambda_2}\sqrt{\lambda_{2t}}\right].$$

Because dividends are not subject to jumps

$$\frac{\mathcal{J}_j(F_t)}{F_t} = \frac{\mathcal{J}_j(G_t)}{G_t}$$

for $j = 1, 2$. The result follows from substituting these expressions and the corresponding expressions for the state-price density π_t (given in (A.40) and (A.41)) into (A.58) of Theorem A.4. $\qquad\square$

Note that the proof of Corollary 2.11 follows along similar lines.

A.2.3. Return simulation

For each asset, the realized return between time t and $t + \Delta t$ is defined as

$$R_{t+\Delta t} = \frac{F_{t+\Delta t} + \int_t^{t+\Delta t} D_s\, ds}{F_t}$$

see (Duffie, 2001, Chapter 6.L). For assets that pay a dividend in each period, namely the aggregate market and the value sector, this return can be computed based on the series of price-dividend ratios and payouts. Using the approximation $D_{t+\Delta t}\Delta t \approx \int_t^{t+\Delta t} D_s\, ds$, it follows that

$$
\begin{aligned}
R_{t,t+\Delta t} &\approx \frac{F_{t+\Delta t} + D_{t+\Delta t}\Delta t}{F_t} \\
&= \frac{\frac{F_{t+\Delta t}}{D_{t+\Delta t}} + \Delta t}{\frac{F_t}{D_t}} \frac{D_{t+\Delta t}}{D_t} \\
&= \frac{G(\mu_{t+\Delta t}, \lambda_{t+\Delta t}) + \Delta t}{G(\lambda_t)} \frac{D_{t+\Delta t}}{D_t}.
\end{aligned}
$$

Computing the return on the growth sector requires a different approach. For $u \geq s \geq t$, let $R^g_{t,s,u}$ denote the return between s and u on the growth sector formed at time t. Because value and growth must add up to the aggregate market,

$$
R^m_{t,t+\Delta t} = \frac{F^v_{t,t}}{F_t} R^v_{t,t+\Delta t} + \left(1 - \frac{F^v_{t,t}}{F_t}\right) R^g_{t,t,t+\Delta t}.
$$

Rearranging, it follows that one-period returns on the growth sector equal

$$
R^g_{t,t,t+\Delta t} = \frac{1}{1 - \frac{F^v_{t,t}}{F_t}}\left(R^m_{t,t+\Delta t} - \frac{F^v_{t,t}}{F_t} R^v_{t,t+\Delta t}\right). \tag{A.62}
$$

Because the price of the value sector formed at time t relative to the aggregate market is given by

$$
\frac{F^v_{t,t}}{F_t} = \frac{G^v(\mu_t, \lambda_t)}{G(\mu_t, \lambda_t)},
$$

it is straightforward to compute the return (A.62) on the growth sector.

164

A.3. Appendix for Dynamic Asset Allocation with Learning

A.3.1. Feedforward Neural Network

This paper solves the portfolio choice problem by standard backward induction, and uses a feedforward neural network to approximate the value function. This method is first introduced to the portfolio choice literature by Skoulakis (2007). Theoretically, it has been shown that one-hidden-layer feedforward neural networks can uniformly approximate any continuous multivariate function to any desired degree of accuracy. In the empirical section, I follow that paper and use two hidden layers to provide more flexibility.

A neural network with two hidden layer defined on $\mathbb{R}^d$ is of the form

$$F\left(x; \alpha, B, \theta, \gamma\right) = \sum_{m=1}^{M} \alpha_m g\left(\beta^\top x + \theta_m\right) + \gamma$$

where $x \in \mathbb{R}^d$, $g(\cdot)$ is the activation function, $\alpha \in \mathbb{R}^M$, $B = [\beta_1, \cdots, \beta_M]^\top$, $\theta \in \mathbb{R}^M$, and $\gamma \in \mathbb{R}$. M refers to the number of nodes in the hidden layer.

In this paper, MATLAB's Neural Network Toolbox is used for all calculations. I choose two hidden layers, fifty hidden nodes, Tan-Sigmoid function as the activation function g, and gradient descent with adaptive learning rate back-propagation.

For more detail on implementing feedforward neural network methodology in this problem, see the Appendix of Skoulakis (2007).

A.3.2. Bayesian Updating Framework

This Appendix follows Skoulakis (2007) to derive the state variables and their laws of motion.

First recall that the posterior distribution can be characterized by $Z^\top Z$, $Z^\top Y$, and $Y^\top Y$. Then follow the definition of Z and Y, they can be decomposed as:

$$Z^\top Z = \begin{bmatrix} t & \sum_{\tau=0}^{t-1} x_\tau \\ \sum_{\tau=0}^{t-1} x_\tau & \sum_{\tau=0}^{t-1} x_\tau^2 \end{bmatrix} = \begin{bmatrix} t & tm_3(t) \\ tm_3(t) & tm_4(t) \end{bmatrix},$$

$$Z^\top Y = \begin{bmatrix} \sum_{\tau=1}^{t} r_\tau & \sum_{\tau=1}^{t} x_\tau \\ \sum_{\tau=1}^{t} x_{\tau-1} r_\tau & \sum_{\tau=1}^{t} x_{\tau-1} x_\tau \end{bmatrix} = \begin{bmatrix} tm_1(t) & tm_3(t) - x_0 + m_8(t) \\ tm_7(t) & tm_5(t) \end{bmatrix},$$

$$Y^\top Y = \begin{bmatrix} \sum_{\tau=1}^{t} r_\tau^2 & \sum_{\tau=1}^{t} r_\tau x_\tau \\ \sum_{\tau=1}^{t} r_\tau x_\tau & \sum_{\tau=1}^{t} x_\tau^2 \end{bmatrix} = \begin{bmatrix} tm_2(t) & tm_6(t) \\ tm_{6(t)} & tm_4(t) - x_0^2 + m_8(t)^2 \end{bmatrix},$$

where $m_1(t) = \sum_{\tau=1}^{t} r_\tau = \bar{r}_t$, $m_2(t) = \frac{1}{t}\sum_{\tau=1}^{t} r_\tau^2$, $m_3(t) = \sum_{\tau=0}^{t-1} x_\tau = \bar{x}_t$, $m_4(t) = \frac{1}{t}\sum_{\tau=0}^{t-1} x_\tau^2$, $m_5(t) = \frac{1}{t}\sum_{\tau=1}^{t} x_{\tau-1} x_\tau$, $m_6(t) = \frac{1}{t}\sum_{\tau=1}^{t} r_\tau x_\tau$, $m_7(t) = \frac{1}{t}\sum_{\tau=1}^{t} x_{\tau-1} r_\tau$, $m_8(t) = x_t$ are all that is required to characterize the posterior distribution of the parameters. These variables are updated according to the following:

$$m_1(t_{k+1}) = \frac{t_k}{t_{k+1}} m_1(t_k) + \frac{1}{t_{k+1}} R_{1,k+1}$$

$$m_2(t_{k+1}) = \frac{t_k}{t_{k+1}} m_2(t_k) + \frac{1}{t_{k+1}} R_{2,k+1}$$

$$m_3(t_{k+1}) = \frac{t_k}{t_{k+1}} m_3(t_k) + \frac{1}{t_{k+1}} \left[m_8(t_k) + Q_{1,k+1} \right]$$

$$m_4(t_{k+1}) = \frac{t_k}{t_{k+1}} m_4(t_k) + \frac{1}{t_{k+1}} \left[m_8(t_k)^2 + Q_{2,k+1} \right]$$

$$m_5(t_{k+1}) = \frac{t_k}{t_{k+1}} m_5(t_k) + \frac{1}{t_{k+1}} \left[m_8(t_k) x_{t_k+1} + F_{k+1} \right]$$

$$m_6(t_{k+1}) = \frac{t_k}{t_{k+1}} m_6(t_k) + \frac{1}{t_{k+1}} G_{k+1}$$

$$m_7(t_{k+1}) = \frac{t_k}{t_{k+1}} m_7(t_k) + \frac{1}{t_{k+1}} \left[m_8(t_k) r_{t_k+1} + H_{k+1} \right]$$

$$m_8(t_{k+1}) = x_{t_{k+1}}$$

where

$$R_{m,k+1} = r^m_{t_k+1} + \cdots + r^m_{t_{k+1}}, \ m = 1,2$$

$$Q_{m,k+1} = x^m_{t_k+1} + \cdots + x^m_{t_{k+1}}, \ m = 1,2$$

$$F_{k+1} = x_{t_k+1}x_{t_k+2} + \cdots + x_{t_{k+1}-1}x_{t_{k+1}}$$

$$G_{k+1} = r_{t_k+1}x_{t_k+1} + \cdots + r_{t_{k+1}}x_{t_{k+1}}$$

$$H_{k+1} = x_{t_k+1}r_{t_k+2} + \cdots + x_{t_{k+1}-1}r_{t_{k+1}}.$$

Following Skoulakis (2007), I work with a transformation of these variables $m(t)$. Define

$$s_1(t) = \bar{r}_t = m_1(t)$$

$$s_2(t) = \frac{1}{t}\sum_{\tau=1}^{t}(r_\tau - \bar{r}_t)^2 = m_2(t) - m_1(t)^2$$

$$s_3(t) = \bar{x}_t = m_3(t)$$

$$s_4(t) = \frac{1}{t}\sum_{\tau=0}^{t-1}(x_\tau - \bar{x}_t)^2 = m_4(t) - m_3(t)^2$$

$$s_5(t) = \frac{\frac{1}{t}\sum_{\tau=1}^{t}(x_{\tau-1} - \bar{x}_t)(x_\tau - \bar{x}_t)}{\frac{1}{t}\sum_{\tau=0}^{t}(x_\tau - \bar{x}_t)^2} = \frac{m_5(t) - m_3(t)\left[m_3(t) - \frac{x_0 - m_8(t)}{t}\right]}{s_4(t)}$$

$$s_6(t) = \frac{\frac{1}{t}\sum_{\tau=1}^{t}(r_\tau - \bar{r}_t)(x_\tau - \bar{x}_t)}{\sqrt{\frac{1}{t}\sum_{\tau=1}^{t}(r_\tau - \bar{r}_t)^2}\sqrt{\frac{1}{t}\sum_{\tau=0}^{t-1}(x_\tau - \bar{x}_t)^2}} = \frac{m_6(t) - m_1(t)\left[m_3(t) - \frac{x_0 - m_8(t)}{t}\right]}{\sqrt{s_2(t)s_4(t)}}$$

$$s_7(t) = \frac{\frac{1}{t}\sum_{\tau=1}^{t}(r_\tau - \bar{r}_t)(x_{\tau-1} - \bar{x}_t)}{\sqrt{\frac{1}{t}\sum_{\tau=1}^{t}(r_\tau - \bar{r}_t)^2}\sqrt{\frac{1}{t}\sum_{\tau=0}^{t-1}(x_\tau - \bar{x}_t)^2}} = \frac{m_7(t) - m_1(t)m_3(t)}{\sqrt{s_2(t)s_4(t)}}$$

$$s_8(t) = m_8(t).$$

Knowing $s(t)$, one can then recover $m(t)$ by

$$m_1(t) = s_1(t)$$

$$m_2(t) = s_2(t) + s_1(t)^2$$

$$m_3(t) = s_3(t)$$

$$m_4(t) = s_4(t) + s_3(t)^2$$

$$m_5(t) = s_4(t)s_5(t) + s_3(t)\left[s_3(t) - \frac{x_0 - s_8(t)}{t}\right]$$

$$m_6(t) = \sqrt{s_2(t)s_4(t)}\,s_6(t) + s_1(t)\left[s_3(t) - \frac{x_0 - s_8(t)}{t}\right]$$

$$m_7(t) = \sqrt{s_2(t)s_4(t)}\,s_7(t) + s_1(t)s_3(t)$$

$$m_8(t) = s_8(t).$$

Thus the laws of motion of the state variables $s(t)$ can be derived:

$$s_1(t_{k+1}) = \frac{t_k}{t_{k+1}} s_1(t_k) + \frac{1}{t_{k+1}} R_{1,k+1}$$

$$s_2(t_{k+1}) = \frac{t_k}{t_{k+1}}\left[s_2(t_k) + s_1(t_k)^2\right] + \frac{1}{t_{k+1}} R_{2,k+1} - s_1(t_{k+1})^2$$

$$s_3(t_{k+1}) = \frac{t_k}{t_{k+1}} s_3(t_k) + \frac{1}{t_{k+1}}\left[s_8(t_k) + Q_{1,k+1}\right]$$

$$s_4(t_{k+1}) = \frac{t_k}{t_{k+1}}\left[s_4(t_k) + s_3(t_k)^2\right] + \frac{1}{t_{k+1}}\left[s_8(t_k)^2 + Q_{2,k+1}\right] - s_3(t_{k+1})^2$$

$$s_5(t_{k+1}) = \frac{\frac{t_k}{t_{k+1}}\left[s_4(t_k)s_5(t_k) + s_3(t_k)\left(s_3(t_k) - \frac{x_0 - s_8(t_k)}{t_k}\right)\right] + \frac{1}{t_{k+1}}\left[s_8(t_k)x_{t_{k+1}} + F_{k+1}\right]}{s_4(t_{k+1})}$$

$$- \frac{s_3(t_{k+1})\left[s_3(t_{k+1}) - \frac{x_0 - x_{t_{k+1}}}{t_{k+1}}\right]}{s_4(t_{k+1})}$$

$$s_6(t_{k+1}) = \frac{\frac{t_k}{t_{k+1}}\sqrt{s_2(t_k)s_4(t_k)}\,s_6(t_k) + s_1(t_k)\left[s_3(t_k) - \frac{x_0 - s_8(t_k)}{t_k}\right] + \frac{1}{t_{k+1}} G_{k+1}}{\sqrt{s_2(t_{k+1})s_4(t_{k+1})}}$$

$$- \frac{s_1(t_{k+1})\left[s_3(t_{k+1}) - \frac{x_0 - x_{t_{k+1}}}{t_{k+1}}\right]}{\sqrt{s_2(t_{k+1})s_4(t_{k+1})}}$$

$$s_7(t_{k+1}) = \frac{\frac{t_k}{t_{k+1}}\left[\sqrt{s_2(t_k)s_4(t_k)}\,s_7(t_k) + s_1(t_k)s_3\right] + \frac{1}{t_{k+1}}\left[s_8(t_k)r_{t_{k+1}} + H_{k+1}\right]}{\sqrt{s_2(t_{k+1})s_4(t_{k+1})}}$$

$$- \frac{s_1(t_{k+1})s_3(t_{k+1})}{\sqrt{s_2(t_{k+1})s_4(t_{k+1})}}$$

$$s_8(t_{k+1}) = x_{t_{k+1}}.$$

A.3.3. Exact Likelihood

In this paper I assume that the first observation of predictor variable x_0 is non-stochastic and conveys no information about the parameters. This assumption discard potential information x_0 may convey. Furthermore, the specification for the prior potentially allows for a non-stationary process for the predictor as ρ could be greater than 1. In order to address these issues, I consider the exact Bayesian model, following Stambaugh (1999).

In particular, I treat x_0 as being drawn from the stationary distribution of Equation (4):

$$x_0 \sim N\left(\frac{\theta}{1-\rho}, \frac{\sigma_v^2}{1-\rho^2}\right).$$

The likelihood function becomes

$$L(b, \Sigma; D) = (2\pi|\Sigma|)^{-\left(\frac{T}{2}\right)} \exp\left\{-\frac{1}{2}(z - Zb)^{\top}(\Sigma^{-1} \otimes I_T)(z - Zb)\right\} \times$$
$$\left(\frac{1-\rho^2}{2\pi\sigma_v^2}\right)^{\frac{1}{2}} \exp\left\{-\frac{1-\rho^2}{2\sigma_v^2}\left(x_0 - \frac{\theta}{1-\rho}\right)^2\right\},$$

imposing the assumption that ρ is between -1 and 1 for stationarity, the prior becomes

$$p(b, \Sigma) \propto |\Sigma|^{-\frac{3}{2}}, \ \rho \in (-1, 1).$$

Combining the two equations yields the posterior

$$p(b, \Sigma|D) \propto |\Sigma|^{-\frac{T+3}{2}} exp\left\{-\frac{1}{2}(z - Zb)'\left(\Sigma^{-1} \otimes I\right)(z - Zb)\right\} \times$$
$$\times \left(\frac{1-\rho^2}{\sigma_v^2}\right)^{1/2} exp\left\{-\frac{1}{2}\left(\frac{1-\rho^2}{\sigma_v^2}\right)\left(x_0 - \frac{\theta}{1-\rho}\right)^2\right\}.$$

To generate draws from this posterior, I implement the Metropolis-Hasting Algorithm (see Chib and Greenberg (1995)).

For $j = 1 : I$

1. Generate $b_j \sim MN\left(\hat{b}, \; (\Sigma_j \otimes X'X)^{-1}\right)$, accept this draw with probability

$$\alpha = min\left\{1, \; \frac{\left(\frac{1-\rho_j^2}{\sigma_v^2}\right)^{\frac{1}{2}} exp\left\{-\frac{1}{2}\left(\frac{1-\rho_j^2}{\sigma_v^2}\right)\left(x_0 - \frac{\theta_j}{1-\rho_j}\right)^2\right\}}{\left(\frac{1-\rho_{j-1}^2}{\sigma_v^2}\right)^{\frac{1}{2}} exp\left\{-\frac{1}{2}\left(\frac{1-\rho_{j-1}^2}{\sigma_v^2}\right)\left(x_0 - \frac{\theta_{j-1}}{1-\rho_j}\right)^2\right\}}\right\}$$

2. Then generate

$$\Sigma_j \sim IW\left(T+1, \; S + \left(B_j + \hat{B}\right)' X'X \left(B_j + \hat{B}\right)\right),$$

   and accept this new draw with probability

$$\alpha = min\left\{1, \; \frac{\left(\sigma_j^{11}\right)^{-\frac{1}{2}} exp\left\{-\frac{1}{2}\left(\frac{1-\rho^2}{\sigma_{vj}^2}\right)\left(x_0 - \frac{\theta}{1-\rho}\right)\right\}}{\left(\sigma_{j-1}^{11}\right)^{-\frac{1}{2}} exp\left\{-\frac{1}{2}\left(\frac{1-\rho^2}{\sigma_{vj-1}^2}\right)\left(x_0 - \frac{\theta}{1-\rho}\right)\right\}}\right\},$$

   where $\sigma_v^2 = \sigma^{11}|\Sigma^{-1}|^{-1}$.

With the exact likelihood specification, learning still introduces a large negative hedging demand, and the results are similar to those obtained using the simpler specification.

BIBLIOGRAPHY

A. Abel. Risk premia and term premia in general equilibrium. *Journal of Monetary Economics*, 43(1):3–33, February 1999.

H. Ai and D. Kiku. Growth to value: Option exercise and the cross-section of equity returns. *Journal of Financial Economics*, 107(2):325–349, February 2013.

A. Ang and G. Bekaert. Stock return predictability: Is it there? *Review of Financial Studies*, 20(3):651–707, May 2007.

A. Ang and M. Piazzesi. A no-arbitrage vector autoregression of term structure dynamics with macroeconomic and latent variables. *Journal of Monetary Economics*, 50(4):745–787, May 2003.

A. Ang, S. Dong, and M. Piazzesi. No-arbitrage taylor rules. Working paper, 2007.

G. S. Bakshi and Z. Chen. Inflation, asset prices, and the term structure of interest rates in monetary economies. *Review of Financial Studies*, 9(1):241–275, Spring 1996.

P. Balduzzi and A. W. Lynch. Transaction costs and predictability: some utility cost calculations. *Journal of Financial Economics*, 52(1):47–78, April 1999.

R. Bansal and I. Shaliastovich. A long-run risks explanation of predictability puzzles in bond and currency markets. *Review of Financial Studies*, 26(1):1–33, January 2013.

R. Bansal and A. Yaron. Risks for the long-run: A potential resolution of asset pricing puzzles. *Journal of Finance*, 59(4):1481–1509, August 2004.

R. Bansal, D. Kiku, and A. Yaron. Long run risks, the macroeconomy and asset prices. *American Economic Review Papers and Proceedings*, 100(2):542–546, May 2010.

R. Bansal, D. Kiku, and A. Yaron. An empirical evaluation of the long-run risks model for asset prices. *Critical Finance Review*, 1(1):183–221, 2012.

N. Barberis. Investing for the long run when returns are predictable. *Journal of Finance*, 55(1):225–264, February 2000.

R. J. Barro. Rare disasters and asset markets in the twentieth century. *The Quarterly Journal of Economics*, 121(3):823–866, August 2006.

R. J. Barro and T. Jin. On the size distribution of macroeconomic disasters. working paper, Harvard University, February 2011.

R. J. Barro and J. F. Ursua. Macroeconomic crises since 1870. *Brookings Papers on Economic Activity*, pages 255–350, Spring 2008.

V. S. Bawa, S. J. Brown, and R. W. Klein. *Estimation risk and optimal portfolio choice.* North-Holland, Amsterdam, The Netherlands, 1979.

J. Beeler and J. Y. Campbell. The long-run risks model and aggregate asset prices: An empirical assessment. *Critical Finance Review*, 1(1):141–182, 2012.

G. Bekaert and E. Engstrom. Asset return dynamics under bad environment-good environment fundamentals. working paper, Columbia University, 2010.

G. Bekaert, R. J. Hodrick, and D. A. Marshall. Peso problem explanations for term structure anomalies. *Journal of Monetary Economics*, 48(2):241–270, October 2001.

G. Bekaert, E. Engstrom, and S. R. Grenadier. Stock and bond returns with moody investors. *Journal of Empirical Finance*, 17(5):867–894, December 2010.

L. Benzoni, P. Collin-Dufresne, and R. S. Goldstein. Explaining asset pricing puzzles associated with the 1987 market crash. *Journal of Financial Economics*, 101(3):552 – 573, September 2011.

J. B. Berk, R. C. Green, and V. Naik. Optimal investment, growth options, and security returns. *Journal of Finance*, 54(5):1553–1607, October 1999.

H. S. Bhamra, L.-A. Kuehn, and I. A. Strebulaev. The levered equity risk premium and credit spreads: A unified framework. *Review of Financial Studies*, 23(2):645–703, February 2010.

R. Bikbov and M. Chernov. No-arbitrage macroeconomic determinants of the yield curve. *Journal of Econometrics*, 159(1):166–182, November 2010.

J. Boudoukh, R. Michaely, M. Richardson, and M. R. Roberts. On the importance of measuring payout yield: Implications for empirical asset pricing. *Journal of Finance*, 62 (2):877–915, April 2007.

M. W. Brandt. Estimating portfolio and consumption choice: A conditional euler equations approach. *Journal of Finance*, 54(5):1609–1645, October 1999.

M. W. Brandt. Portfolio choice problems. In Y. Ait-Sahalia and L. P. Hansen, editors, *Handbook of Financial Econometrics, Volume 1: Tools and Techniques.* North-Holland, Amsterdam, The Netherlands, 2010.

M. W. Brandt, A. Goyal, P. Santa-Clara, and J. R. Stroud. A simulation approach to dynamics portfolio choice with an application to learning about return predictability. *Review of Financial Studies*, 18(3):831–873, Fall 2005.

D. T. Breenden. An intertemporal asset pricing model with stochastic consumption and investment opportunities. *Journal of Financial Economics*, 7(3):265–296, September 1979.

M. Brennan. The role of learning in dynamic portfolio decisions. *European finance review*, 1(3):295–306, 1998.

M. J. Brennan, E. S. Schwartz, and R. Lagnado. Strategic asset allocation. *Journal of Economic Dynamics and Control*, 21(8-9):1377–1403, June 1997.

S. J. Brown. Optimal portfolio choice under uncertainty: A bayesian approach. In V. S. Bawa, S. J. Brown, and R. W. Klein, editors, *Estimation Risk and Optimal Portfolio Choice*. North-Holland, Amsterdam, The Netherlands, 1979.

A. Buraschi and A. Jiltsov. Habit formation and macroeconomic models of the term structure of interest rates. *Journal of Finance*, 62(6):3009–3063, December 2007.

L. E. Calvet and A. J. Fisher. Multifrequency news and stock returns. *Journal of Financial Economics*, 86(1):178–212, October 2007.

L. E. Calvet and A. J. Fisher. Multifrequency jump-diffusions: An equilibrium approach. *Journal of Mathematical Economics*, 44(2):207–226, January 2008.

J. Y. Campbell. Consumption-based asset pricing. In G. Constantinides, M. Harris, and R. Stulz, editors, *Handbook of the Economics of Finance, vol. 1b*, pages 803–887. Elsevier Science, North-Holland, 2003.

J. Y. Campbell and J. H. Cochrane. By force of habit: A consumption-based explanation of aggregate stock market behavior. *Journal of Political Economy*, 107(2):205–251, April 1999.

J. Y. Campbell and R. J. Shiller. The dividend-price ratio and expectations of future dividends and discount factors. *Review of Financial Studies*, 1(3):195–228, July 1988.

J. Y. Campbell and R. J. Shiller. Yield spreads and interest rate movements: A bird's eye view. *Review of Economic Studies*, 58(3):495–514, May 1991.

J. Y. Campbell and L. M. Viceira. Consumption and portfolio decisions when expected returns are time-varying. *Quarterly Journal of Economics*, 114(2):433–495, May 1999.

J. Y. Campbell and L. M. Viceira. *Strategic Asset Allocation: Portfolio Choice for Long-Term Investors*. Oxford University Press, Oxford, UK, 2002.

J. Y. Campbell and T. Vuolteenaho. Bad beta, good beta. *American Economic Review*, 94 (5):1249–1275, December 2004.

J. Y. Campbell, C. Polk, and T. Vuolteenaho. Growth or glamour? fundamentals and systematic risk in stock returns. *Review of Financial Studies*, 23(1):305 – 344, January 2010.

M. Carlson, A. Fisher, and R. Giammarino. Corporate investment and asset price dynam-

ics: Implications for the cross-section of returns. *Journal of Finance*, 59(6):2577–2603, December 2004.

S. Chib and E. Greenberg. Understanding the metropolis-hastings algorithm. *American Statistician*, 49(4):327–335, November 1995.

J. H. Cochrane. Explaining the variance of price-dividend ratios. *Review of Financial Studies*, 5(2):243–280, April 1992.

J. H. Cochrane. Presidential address: Discount rates. *The Journal of Finance*, 66(4):1047–1108, August 2011.

J. H. Cochrane and M. Piazzesi. Bond risk premia. *American Economic Review*, 95(1):138–160, March 2005.

R. B. Cohen, C. Polk, and T. Vuolteenaho. The value spread. *Journal of Finance*, 58(2):609–641, April 2003.

P. Collin-Dufresne and R. S. Goldstein. Do bonds span the fixed income markets? theory and evidence for unspanned stochastic volatility. *Journal of Finance*, 57(4):1685–1730, August 2002.

P. Collin-Dufresne, R. S. Goldstein, and C. S. Jones. Can interest rate volatility be extracted from the cross section of bond yields? *Journal of Financial Economics*, 94(1):47–66, October 2009.

G. M. Constantinides. Discussion of 'macroeconomic crises since 1870'. *Brookings Papers on Economic Activity*, pages 336–350, Spring 2008.

J. C. Cox, J. C. Ingersoll, and S. A. Ross. A theory of the term structure of interest rates. *Econometrica*, 53(2):385–408, March 1985.

Q. Dai and K. Singleton. Expectations puzzles, time-varying risk premia, and affine models of the term structure. *Journal of Financial Economics*, 63(3):415–442, March 2002.

R. M. Dammon, C. S. Spatt, and H. H. Zhang. Optimal consumption and investment with capital gains taxes. *Review of Financial Studies*, 14(3):583–616, July 2001.

S. R. Das and R. K. Sundaram. A numerical algorithm for consumption-investment problems. *International Journal of Intelligent Systems in Accounting, Finance, and Management*, 11(2):55–69, April 2002.

G. R. Duffee. Term premia and interest rate forecasts in affine models. *Journal of Finance*, 57(1):369–443, February 2002.

G. R. Duffee. Term structure estimation without using latent factors. *Journal of Financial Economics*, 79(3):507–536, March 2006.

G. R. Duffee. Bond pricing and the macroeconomy. forthcoming, *Handbook of the Economics of Finance*, 2012.

D. Duffie. *Dynamic Asset Pricing Theory*. Princeton University Press, Princeton, NJ, 3 edition, 2001.

D. Duffie and L. G. Epstein. Asset pricing with stochastic differential utility. *Review of Financial Studies*, 5(3):411–436, July 1992.

D. Duffie and C. Skiadas. Continuous-time asset pricing: A utility gradient approach. *Journal of Mathematical Economics*, 23(2):107–132, March 1994.

D. Duffie, J. Pan, and K. Singleton. Transform analysis and asset pricing for affine jump-diffusions. *Econometrica*, 68(6):1343–1376, November 2000.

P. Ehling, M. F. Gallmeyer, C. Heyerdahl-Larsen, and P. K. Illeditsch. Beliefs about inflation and the term structure of interest rates. Working paper, 2012.

L. G. Epstein and S. E. Zin. Substitution, risk aversion and the temporal behavior of consumption and asset returns: A theoretical framework. *Econometrica*, 57(4):937–969, July 1989.

B. Eraker and I. Shaliastovich. An equilibrium guide to designing affine pricing models. *Mathematical Finance*, 18(4):519–543, October 2008.

E. F. Fama and R. R. Bliss. The information in long-maturity forward rates. *American Economic Review*, 77(4):680–692, September 1987.

E. F. Fama and K. R. French. Business conditions and expected returns on stocks and bonds. *Journal of Financial Economics*, 25(1):23–49, November 1989.

E. F. Fama and K. R. French. The cross-section of expected stock returns. *Journal of Finance*, 47(2):427–465, June 1992.

E. F. Fama and W. G. Schwert. Asset returns and inflation. *Journal of Financial Economics*, 5(2):115–146, November 1977.

X. Gabaix. Power laws in economics and finance. *Annual review of economics*, 1:255–293, 2009.

X. Gabaix. An exactly solved framework for ten puzzles in macro-finance. *Quarterly Journal of Economics*, 127(2):645–700, March 2012.

N. B. Gârleanu, L. Kogan, and S. Panageas. Displacement risk and asset returns. *Journal of Financial Economics*, 105(3):491–510, September 2012.

J. Gomes, L. Kogan, and L. Zhang. Equilibrium cross section of returns. *Journal of Political Economy*, 111(4):693–732, August 2003.

F. Gourio. Time-series predictability in the disaster model. *Finance Research Letters*, 5(4): 191–203, 2008.

F. Gourio. Disaster risk and business cycles. *American Economic Review*, 2012.

L. P. Hansen. Dynamic valuation decomposition within stochastic economies. *Econometrica*, 80(3):911–967, May 2012.

L. P. Hansen, J. C. Heaton, and N. Li. Consumption strikes back? Measuring long run risk. *Journal of Political Economy*, 116(2):260–302, 2008.

R. J. Hodrick. Dividend yields and expected stock returns: Alternative procedures for inference and measurement. *Review of Financial Studies*, 5:357–386, 1992.

H. Jeffreys. *Theory of Probability.* Oxford University Press, Clarenden, 1961.

B. Jovanovic and P. L. Rousseau. Two technological revolutions. *Journal of the European Economic Association*, 1(2/3):pp. 419–428, 2003.

C. Julliard and A. Ghosh. Can rare events explain the equity premium puzzle? *Review of Financial Studies*, (3037-3076), 2012.

S. Kandel and R. F. Stambaugh. On the predictability of stock returns: An asset allocation perspective. *Journal of Finance*, 51:385–424, 1996.

D. B. Keim and R. F. Stambaugh. Predicting returns in the stock and bond markets. *Journal of Financial Economics*, 17:357–390, 1986.

T. S. Kim and E. Omberg. Dynamic nonmyopic portfolio behavior. *Review of Financial Studies*, 9:141–161, 1996.

R. W. Klein and V. S. Bawa. The effect of estimation risk on optimal portfolio choice. *Journal of Financial Economics*, 3:215–231, 1976.

L. Kogan and R. Uppal. Risk aversion and optimal portfolio policies in partial and general equilibrium economies. Working Paper, MIT, 2001.

M. Lettau and S. C. Ludvigson. Consumption, aggregate wealth and expected stock returns. *Journal of Finance*, 56:815–849, 2001.

M. Lettau and J. A. Wachter. Why is long-horizon equity less risky? a duration-based explanation of the value premium. *Journal of Finance*, 62:55–92, February 2007.

M. Lettau and J. A. Wachter. The term structures of equity and interest rates. *Journal of Financial Economics*, 101(1):90–113, 2011.

M. Lettau, S. C. Ludvigson, and J. A. Wachter. The declining equity premium: What role does macroeconomic risk play? *Review of Financial Studies*, 21:1653–1687, 2008.

176

J. Lewellen. Predicting returns with financial ratios. *Journal of Financial Economics*, 74: 209–235, November 2004.

R. Litterman and J. A. Scheinkman. Common factors affecting bond returns. *Journal of Fixed Income*, pages 54–61, June 1991.

J. Liu. Portfolio selection in stochastic environments. *Review of Financial Studies*, 20(1): 1–39, 2007.

F. A. Longstaff. Arbitrage and the expectations hypothesis. *Journal of Finance*, 55:989–994, 2000.

F. A. Longstaff and M. Piazzesi. Corporate earnings and the equity premium. *Journal of Financial Economics*, 74:401–421, 2004.

R. Mehra and E. Prescott. The equity risk premium: A solution? *Journal of Monetary Economics*, 22:133–136, 1988.

R. C. Merton. Optimum consumption and portfolio rules in a continuous-time model. *Journal of Economic Theory*, 3:373–413, 1971.

E. Nakamura, J. Steinsson, R. Barro, and J. Ursua. Crises and recoveries in an empirical model of consumption disasters. Working paper, Columbia University and Harvard University, February 2011.

R. Novy-Marx. Operating leverage. *Review of Finance*, 15(1):103–134, 2010.

L. Pastor and P. Veronesi. Learning in financial markets. *Annual Review of Financial Economics*, 1(1):361–381, 2009.

M. Piazzesi and M. Schneider. Equilibrium yield curves. *NBER Macroeconomics Annual*, 21:389–442, 2006.

T. A. Rietz. The equity risk premium: A solution. *Journal of Monetary Economics*, 22: 117–131, 1988.

G. D. Rudebusch and T. Wu. A macro-finance model of the term structure, monetary policy, and the economy. *Economic Journal*, 118:906–926, 2008.

P. A. Samuelson. Lifetime portfolio selection by dynamic stochastic programming. *The Review of Economics and Statistics*, 51(3):239–46, August 1969.

T. Santos and P. Veronesi. Habit formation, the cross section of stock returns and the cash-flow risk puzzle. *Journal of Financial Economics*, 97(3):385–413, September 2010.

M. Schroder and C. Skiadas. Optimal consumption and portfolio selection with stochastic differential utility. *Journal of Economic Theory*, 89:68–126, 1999.

R. J. Shiller. Stock prices and social dynamics. *Brookings Papers on Economic Activity*, 2: 457–498, 1984.

G. Skoulakis. Dynamic portfolio choice with bayesian learning. Working paper, University of Maryland, March 2007.

R. F. Stambaugh. The information in forward rates: Implications for models of the term structure. *Journal of Financial Economics*, 21(1):41–70, 1988.

R. F. Stambaugh. Predictive regressions. *Journal of Financial Economics*, 54:375–421, 1999.

L. H. Summers. Does the stock market rationally reflect fundamental values? *Journal of Finance*, 41:591–601, 1986.

P. Veronesi. The peso problem hypothesis and stock market returns. *Journal of Economic Dynamics and Control*, pages 707–725, 2004.

J. A. Wachter. Portfolio and consumption decisions under mean-reverting returns: An exact solution for complete markets. *Journal of Financial and Quantitative Analysis*, 37:63–91, 2002.

J. A. Wachter. A consumption-based model of the term structure of interest rates. *Journal of Financial Economics*, 79:365–399, February 2006.

J. A. Wachter. Asset allocation. *Annual Review of Financial Economics*, 2(1):175–206, 2010.

J. A. Wachter. Can time-varying risk of rare disasters explain aggregate stock market volatility? forthcoming, *Journal of Finance*, 2012.

P. Weil. Nonexpected utility in macroeconomics. *Quarterly Journal of Economics*, 105: 29–42, February 1990.

Y. Xia. Learning about predictability: The effects of parameter uncertainty on dynamic asset allocation. *Journal of Finance*, 56:205–246, 2001.

A. Zellner and V. K. Chetty. Prediction and decision problems in regression models from the bayesian point of view. *Journal of the American Statistical Association*, 69(429): 608–616, June 1965.

L. Zhang. The value premium. *Journal of Finance*, 60:67–103, February 2005.