

NBER WORKING PAPER SERIES

RARE BOOMS AND DISASTERS IN A MULTI-SECTOR ENDOWMENT ECONOMY

Jerry Tsai
Jessica A. Wachter

Working Paper 20062
<http://www.nber.org/papers/w20062>

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
April 2014

We thank Adlai Fisher, Joao Gomes, Leonid Kogan, Nikolai Roussanov, Harald Uhlig and seminar participants at Boston University, McGill University, University of Miami, University of Pennsylvania, at the American Finance Association Meetings and at the UBC Summer Finance Conference for helpful comments. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

NBER working papers are circulated for discussion and comment purposes. They have not been peer-reviewed or been subject to the review by the NBER Board of Directors that accompanies official NBER publications.

© 2014 by Jerry Tsai and Jessica A. Wachter. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

Rare Booms and Disasters in a Multi-sector Endowment Economy
Jerry Tsai and Jessica A. Wachter
NBER Working Paper No. 20062
April 2014
JEL No. G12

ABSTRACT

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 the high equity premium and volatility of the aggregate market. The model explains other puzzling aspects of the data such as joint patterns in time series predictability of aggregate market and value and growth returns, long periods in which growth outperforms value, and the association between directional covariance and skewness and low realized returns.

Jerry Tsai
Department of Economics
Manor Road Building
Oxford, OX1 3UQ
U.K.
jerry.tsai@economics.ox.ac.uk

Jessica A. Wachter
Department of Finance
2300 SH-DH
The Wharton School
University of Pennsylvania
3620 Locust Walk
Philadelphia, PA 19104
and NBER
jwachter@wharton.upenn.edu

1 Introduction

Among the myriad facts that characterize the cross-section of stock returns, the value premium stands out both for its empirical robustness and the difficulties it poses for theory. The value premium is the finding that stocks with high book-to-market ratios (value) have higher expected returns than stocks with low book-to-market ratios (growth). By itself, this finding would not constitute a puzzle, for it could be that value firms are more risky. Such firms would then have high expected returns in equilibrium, which would simultaneously explain both their high realized returns and their low valuations. The problem with this otherwise appealing explanation is that value stocks are not riskier according to conventional measures. Over the postwar period, which is long enough to measure second moments, value stocks have lower covariance with the market, and lower standard deviations. And while one could argue that neither definition of risk is appropriate in a complex world, the challenge still remains to find a measure of risk that does not, in equilibrium, essentially amount to covariance or standard deviation. Over a decade of theoretical research on the value premium demonstrates that this is a significant challenge indeed.

This paper proposes an explanation of the value premium that is not risk-based but rather based on rare events. We introduce a representative agent asset pricing model in which the endowment has heavy tails; it is subject to positive and negative events that are much larger than what would be expected based on a normal distribution. One of our theoretical contributions is to show a fundamental difference between how disasters and booms affect risk premia. The possibility of a disaster, of course, raises risk premia. Samples in which disasters occur with their correct size and frequency will have lower observed excess returns than samples in which disasters are relatively absent. However, the difference between these two hypothetical samples will be relatively small. The possibility of a boom also raises risk premia, because it too is a source of risk. Samples in which booms occur with their correct size and frequency will have higher observed excess returns

than samples in which booms are relatively absent. However, the difference between these two hypothetical samples will be large. It will be more than the entire premium due to the boom. We use this simple theoretical observation to account for the value premium. We define a growth sector as consisting of stocks that capture the benefits of a large consumption boom. We show that an observed value premium can result if booms were expected but did not occur.

While sample selection can account for the observed value premium, it would not, by itself explain the high volatility of growth stocks. This second finding arises if we assume, realistically, that the probability of a boom varies over time. The growth sector is exposed to time-variation in the boom probability, and so is more volatile, and has a higher covariance with the market than the value sector. However, it has lower observed returns. Moreover, while growth stocks do have higher returns in population than value stocks, the difference is small relative to the amount of additional risk. The price of boom risk is endogenously low because of decreasing marginal utility.

As the discussion above makes clear, explaining the value premium is as much about explaining second moments as first moments. Thus a key aspect of our model is generating realistic second moments by a mechanism that is favored by data. A large literature argues that stock market volatility arises primarily from time-varying risk premia (Campbell and Shiller (1988)). This is the case in our model, which implies that excess returns are predictable but that, in samples without rare events, cash flows are not. Moreover, our model implies patterns of predictability that are consistent with the data. As in the data, the price-dividend ratio predicts excess market returns, but not the return on value minus growth. The value-minus-growth return is, however, predictable by the value spread, as in the data. (Cohen, Polk, and Vuolteenaho (2003)).

In our focus on the underlying dynamics separating value and growth, our model follows a substantial literature that explicitly models the cash flow dynamics of firms, or sectors, and how these relate to risk premia in the cross-section (Ai and Kiku (2013), Berk, Green, and Naik (1999), Carlson, Fisher, and Giammarino (2004), Gârleanu, Kogan, and

Panageas (2012), Gomes, Kogan, and Zhang (2003), Kogan, Papanikolaou, and Stoffman (2013), Novy-Marx (2010), Zhang (2005)). These papers show how endogenous investment dynamics can lead to a value premium. Ultimately, however, the value premium arises in these models because of greater risk. Thus they do not explain the observed pattern in postwar variances and covariances. In contrast, a second branch of the literature relates cash flow dynamics of portfolios, as opposed to underlying firms, to risk premia (Bansal, Dittmar, and Lundblad (2005), Da (2009), Hansen, Heaton, and Li (2008), Kiku (2006)). This literature finds that dividends on the value portfolio are more correlated with a long-run component of consumption than dividends and returns on the growth portfolio. In the context of a model where risk to this long-run component is priced (Bansal and Yaron (2004)) this covariance leads to a higher premium for value. However, for this long-run component to be an important source of risk in equilibrium, it also must be present in the market portfolio, and it must be an important source of variation in these returns themselves. Again, this would seem to imply, counterfactually, that the covariance with the market return and volatility of returns would be greater for value than for growth. Moreover, if the long-run component of consumption growth is an important source of risk in the market portfolio, consumption growth should be forecastable by stock prices; however, it is not (Beeler and Campbell (2012)).

To capture the disconnect between risk and return in the cross-section, shocks associated with growth stocks should have a low price of risk. As shown by Lettau and Wachter (2007), Santos and Veronesi (2010) and Binsbergen, Brandt, and Koijen (2012), achieving this pricing poses a challenge for general equilibrium models.¹ Kogan and Papanikolaou (2013) endogenously generate a cross-section of firms through differences in investment opportunities, but, like Lettau and Wachter, they assume an exogenous stochastic discount

¹Campbell and Vuolteenaho (2004) and Lettau and Wachter (2007) consider the role of duration in generating a value premium when discount rate shocks carry a zero or negative price. Campbell and Vuolteenaho use a partial-equilibrium ICAPM while Lettau and Wachter exogenously specify the stochastic discount factor.

factor. Papanikolaou (2011) does present an equilibrium model in which investment shocks have a negative price of risk. This is achieved by assuming that the representative agent has a preference for late resolution of uncertainty. While helpful for explaining the cross-section, this assumption implies an equity premium that is counterfactually low. In our model, growth stocks are exposed to a source of risk that with a price that is close to zero. Nonetheless, our model has a reasonable equity premium.

In this paper, the presences of rare disasters generates a plausible equity premium, as in Rietz (1988), Longstaff and Piazzesi (2004), Veronesi (2004) and Barro (2006). Time-variation in disaster risk is the primary driver of stock market volatility, and in this our paper is similar to Gabaix (2012), Gourio (2012) and Wachter (2013). These papers do not study rare booms for the simple reason that they have little role in explaining the unconditional equity premium, as we show. Our rare booms are similar to technological innovations, modeled by Pástor and Veronesi (2009), and Jovanovic and Rousseau (2003).² Bekaert and Engstrom (2013) also assume a two-sided risk structure, but propose a model of the representative agent motivated by habit formation. These papers do not address the cross-section of stock returns however.

The remainder of the paper is organized as follows. Section 2 describes and solves the model. Section 3 discusses the model's quantitative implications for risk and return of value and growth firms. Section 4 looks at further implications of our model's mechanism, and how these fare in the data. Section 5 concludes.

²Pástor 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.

2 Model

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:

$$\frac{dC_t}{C_t} = \mu_{Ct} dt + \sigma dB_{Ct}, \quad (1)$$

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

$$\mu_{Ct} = \bar{\mu}_C + \mu_{1t} + \mu_{2t}, \quad (2)$$

where

$$d\mu_{jt} = -\kappa_{\mu_j} \mu_{jt} dt + Z_{jt} dN_{jt}, \quad (3)$$

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

$$d\lambda_{jt} = \kappa_{\lambda_j} (\bar{\lambda}_j - \lambda_{jt}) dt + \sigma_{\lambda_j} \sqrt{\lambda_{jt}} dB_{\lambda_{jt}}. \quad (4)$$

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 $\lambda_t = [\lambda_{1t}, \lambda_{2t}]^\top$, $\mu_t = [\mu_{1t}, \mu_{2t}]^\top$, $B_{\lambda t} = [B_{\lambda_{1t}}, B_{\lambda_{2t}}]^\top$ and $B_t = [B_{Ct}, B_{\lambda t}^\top]^\top$.³

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:

$$V_t = E_t \int_t^\infty f(C_s, V_s) ds, \quad (5)$$

where

$$f(C_t, V_t) = \beta(1 - \gamma)V_t \left(\log C_t - \frac{1}{1 - \gamma} \log((1 - \gamma)V_t) \right). \quad (6)$$

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 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 1 is in Appendix B.

Theorem 1. *Assume parameter values satisfy Assumption 1. Then the value function J takes the following form:*

$$J(W_t, \mu_t, \lambda_t) = \frac{W_t^{1-\gamma}}{1-\gamma} I(\mu_t, \lambda_t), \quad (7)$$

where

$$I(\mu_t, \lambda_t) = \exp \{ a + b_\mu^\top \mu_t + b_\lambda^\top \lambda_t \}, \quad (8)$$

³We assume throughout that $\kappa_{\mu_j}, \kappa_{\lambda_j}, \bar{\lambda}_j$ and σ_{λ_j} , for $j = 1, 2$, are strictly positive.

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 $j = 1, 2$ take the following form:

$$a = \frac{1-\gamma}{\beta} \left(\bar{\mu}_C - \frac{1}{2}\gamma\sigma^2 \right) + (1-\gamma) \log \beta + \frac{1}{\beta} b_\lambda^\top (\kappa_\lambda * \bar{\lambda}) \quad (9)$$

$$b_{\mu_j} = \frac{1-\gamma}{\kappa_{\mu_j} + \beta}, \quad (10)$$

$$b_{\lambda_j} = \frac{1}{\sigma_{\lambda_j}^2} \left(\beta + \kappa_{\lambda_j} - \sqrt{(\beta + \kappa_{\lambda_j})^2 - 2E_{\nu_j} [e^{b_{\mu_j} Z_{jt}} - 1] \sigma_{\lambda_j}^2} \right). \quad (11)$$

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. *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 (10) that $b_{\mu_j} < 0$. Because $Z_1 < 0$ and $b_{\mu_1} < 0$, $E_{\nu_1} [e^{b_{\mu_1} Z_{1t}} - 1] > 0$. Therefore,

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

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

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

and $b_{\lambda_2} < 0$. □

The riskfree rate takes a particularly simple form:

Corollary 3. *Let r_t denote the instantaneous risk-free rate in this economy, then r_t is given by*

$$r_t = \beta + \mu_{Ct} - \gamma\sigma^2. \quad (12)$$

2.3 The aggregate market

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

$$\frac{dD_t}{D_t} = \mu_{D_t} dt + \phi \sigma dB_{C_t}, \quad (13)$$

where

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

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.

2.3.1 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

$$\pi_t = \exp \left\{ \int_0^t \frac{\partial}{\partial V} f(C_s, V_s) ds \right\} \frac{\partial}{\partial C} f(C_t, V_t). \quad (14)$$

Let $H(D_t, \mu_t, \lambda_t, \tau)$ 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 B.2, given in Appendix B.4.

Corollary 4. *The solution for the function H is as follows*

$$H(D_t, \mu_t, \lambda_t, \tau) = D_t \exp \{ a_\phi(\tau) + b_{\phi\mu}(\tau)^\top \mu_t + b_{\phi\lambda}(\tau)^\top \lambda_t \}, \quad (15)$$

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 $j = 1, 2$,

$$b_{\phi\mu_j}(\tau) = \frac{\phi - 1}{\kappa_{\mu_j}} (1 - e^{-\kappa_{\mu_j}\tau}), \quad (16)$$

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

$$\frac{db_{\phi\lambda_j}}{d\tau} = \frac{1}{2}\sigma_{\lambda_j}^2 b_{\phi\lambda_j}(\tau)^2 + (b_{\lambda_j}\sigma_{\lambda_j}^2 - \kappa_{\lambda_j}) b_{\phi\lambda_j}(\tau) + E_{\nu_j} \left[e^{b_{\mu_j}Z_{jt}} \left(e^{b_{\phi\mu_j}(\tau)Z_{jt}} - 1 \right) \right] \quad (17)$$

$$\frac{da_\phi}{d\tau} = \bar{\mu}_D - \bar{\mu}_C - \beta + \gamma\sigma^2(1 - \phi) + b_{\phi\lambda}(\tau)^\top (\kappa_\lambda * \bar{\lambda}) \quad (18)$$

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

$$G(\mu_t, \lambda_t) = \int_0^\infty \exp(a_\phi(\tau) + b_{\phi\mu}(\tau)^\top \mu_t + b_{\phi\lambda}(\tau)^\top \lambda_t) d\tau. \quad (19)$$

The expressions in Corollary 4 show how prices respond to innovations in expected consumption growth and in changing disaster probabilities. Because $\phi > 1$, (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 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.⁴ 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.

⁴The derivative of (16) with respect to κ_{μ_j} equals $(\kappa_{\mu_j}\tau + 1)e^{-\kappa_{\mu_j}\tau} - 1$ which is negative, because $e^{\kappa_{\mu_j}\tau} > \kappa_{\mu_j}\tau + 1$.

Corollary 5. *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 (16). The results for λ_{1t} and λ_{2t} are less obvious. We give a full proof in Appendix B and discuss the intuition here. Consider the ODE (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 (17) can itself be written as a sum of two terms:

$$E_{\nu_1} \left[e^{b_{\mu_1} Z_{1t}} \left(e^{b_{\phi\mu_1}(\tau) Z_{1t}} - 1 \right) \right] =$$

$$- \underbrace{E_{\nu_1} \left[\left(e^{b_{\mu_1} Z_{1t}} - 1 \right) \left(1 - e^{b_{\phi\mu_1}(\tau) Z_{1t}} \right) \right]}_{\text{Risk premium effect}} + \underbrace{E_{\nu_1} \left[e^{b_{\phi\mu_1}(\tau) Z_{1t}} - 1 \right]}_{\text{Cash flow and riskfree rate effect}} \quad (20)$$

The first of the terms in (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.⁵ 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 (20) with a negative sign. The second term in (20) is the expected price response in the event of a disaster,⁶ representing the combined effect changes in expected cash flows and the riskfree rate. Thus the response of equity values to a change in the disaster probability is determined by a risk premium effect, and a (joint) cash flow and riskfree rate effect. Both effects operate in the same direction, namely an increase in the probability lowers valuations. For booms, on the other hand, the joint riskfree-rate and cash flow effect is positive, and dominates the risk premium effect. An increase in the probability of a boom increases valuations, in spite of the effect on the equity premium.

⁵More precisely, this is the static disaster premium for zero-coupon equity with maturity τ .

⁶Again, more precisely, it is the price response of zero-coupon equity with maturity τ .

2.3.2 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 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:

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

Further, define

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

for $j = 1, 2$, and

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

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:

$$r_t^m = \mu_{F,t} + \frac{D_t}{F_t} + \frac{1}{F_t} \lambda_t^\top \bar{\mathcal{J}}(F_t). \quad (21)$$

Corollary 6. *The equity premium relative to the risk-free rate r is*

$$r_t^m - r_t = \phi\gamma\sigma^2 - \underbrace{\sum_j \lambda_{jt} E_{\nu_j} \left[(e^{b_{\mu_j} Z_{jt}} - 1) \frac{\mathcal{J}_j(G_t)}{G_t} \right]}_{\text{static rare event premium}} - \underbrace{\sum_j \lambda_{jt} \frac{1}{G_t} \frac{\partial G}{\partial \lambda_j} b_{\lambda_j} \sigma_{\lambda_j}^2}_{\lambda\text{-premium}}. \quad (22)$$

As Corollary 6 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*:

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

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.⁷

Finally, the third term in (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 7. 1. *The static disaster and boom premiums are positive.*

⁷However, 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.

2. *The premiums for time-varying disaster and boom probabilities (the λ_j -premiums) are also positive.*

Proof To show the first statement, recall that $b_{\mu_j} < 0$ for $j = 1, 2$ (Corollary 2). 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 5), $\mathcal{J}_1(G_t) < 0$. 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). Further, $\partial G/\partial \lambda_1 < 0$ (Corollary 5). For booms ($j = 2$), each of these quantities takes the opposite sign. The result follows. \square

The intuitive content of Corollary 7 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 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

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

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 6, these expected returns can be calculated as follows:

Corollary 8. *The observed expected excess return in a sample without jumps is*

$$r_{nj,t}^m - r_t = \phi\gamma\sigma^2 - \sum_j \lambda_{jt} E_{\nu_j} \left[e^{b_{\mu_j}Z_{jt}} \frac{\mathcal{J}_j(G_t)}{G_t} \right] - \sum_j \lambda_{jt} \frac{1}{G_t} \frac{\partial G}{\partial \lambda_j} b_{\lambda_j} \sigma_{\lambda_j}^2 \quad (23)$$

This expression differs from (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 $-\sum_j \lambda_{jt} E_{\nu_j} \left[(e^{b_{\mu_j} Z_{jt}} - 1) \frac{\mathcal{J}_j(G_t)}{G_t} \right]$. 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 9. *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 $\mathcal{J}_2(G) > 0$. \square

Note that the observed disaster premium is positive, just like the true disaster premium. However, the observed boom premium is negative, the opposite sign to the true boom premium.⁸ This observation will play an important role in the theory that follows.

2.4 Growth and value sectors

Consider an asset that pays a dividend D_t^v defined by the process

$$\frac{dD_t^v}{D_t^v} = \mu_{D_t}^v dt + \phi \sigma dB_{Ct}, \quad (24)$$

where $\mu_{D_t}^v = \bar{\mu}_D + \phi \mu_{1t}$. Note that this asset is the same as the aggregate market, except it is not exposed to positive jumps. The superscript v denotes “value”. As with the market portfolio, the price follows from no-arbitrage:

$$F^v(D_t^v, \mu_t, \lambda_t) = E_t \int_0^\infty \frac{\pi_{t+\tau}}{\pi_t} D_{t+\tau}^v d\tau \quad (25)$$

As long as there are no positive jumps, the dividend on this claim (which we will call, for the moment, the value claim) and the aggregate market are identical. However, when

⁸We 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). 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 in actual data.

a positive jump takes place, the market dividend diverges permanently from the value dividend. The dividend on the value claim will henceforth grow at a lower rate than the aggregate dividend. We illustrate this in Figure 6.

A concern one might have with this definition of value is that the size of the value sector in the economy is non-stationary. Over time, the value sector grows ever smaller as a proportion of the market as a whole. This does not appear to be the case in the data. The issue of stationarity is an important one in the literature, with some authors taking the view that a non-stationary cross-section is acceptable (Martin (2013), Cochrane, Longstaff, and Santa-Clara (2008)), while others explicitly modeling share dynamics so a stationary cross-section is achieved (Santos and Veronesi (2010), Lettau and Wachter (2007)).

As it turns out, there is simple modification to this definition that does imply a stationary distribution. Consider the price-dividend ratio of the claim defined above:

$$\begin{aligned} G^v(\mu_t, \lambda_t) &= E_t \int_0^\infty \frac{\pi_{t+\tau}}{\pi_t} \frac{D_{t+\tau}^v}{D_t^v} d\tau \\ &= \int_0^\infty \exp(a_\phi^v(\tau) + b_{\phi\mu}^v(\tau)^\top \mu_t + b_{\phi\lambda}^v(\tau)^\top \lambda_t) d\tau. \end{aligned}$$

The terms in the second equation will be described in more details in the section that follows; all that matters for the moment is that this is a function of the state variables. Note that the ratio of this price-dividend ratio to the price-dividend ratio of the market will be stationary, since they are both functions of stationary variables.

At each time t , define the value sector, not as (25), but as

$$F^v(D_t, \mu_t, \lambda_t) = D_t G^v(\mu_t, \lambda_t) \tag{26}$$

This definition of the value sector is as valid as (25). Equation (26) represents the no-arbitrage value of the asset with dividend stream given by (24) when we consider dividends at times $s > t$, with the boundary condition $D_t^v = D_t$. That is, it is our previous value sector, rescaled by the D_t/D_t^v . At every time t , we can define this stream of dividends, and, by no-arbitrage, the value claim (26). Instantaneous risk premia on the value claim will be identical regardless of what definition one uses since these depend on the distribution

of future dividend growth. Returns over a finite time interval can be defined using only the price-dividend ratios and future dividend growth rates, and as such do not depend on which definition of the value claim one uses (Appendix C).⁹

Given this definition of the value sector, the growth sector is defined as the residual. Let $D_t^g = D_t - D_t^v$. Define F_t^g to be the price of the growth claim. Then, by the absence of arbitrage,

$$F_t^g = F_s - F_t^v.$$

The value of the growth sector is also shown in Figure 6.¹⁰

This change in scale is a mathematical convenience, but it also reflects a real-world property of shares, namely that the name that one attached to a particular stream of cash flows changes over time with mergers, acquisitions and liquidations. One can interpret this as a model of two firms, “value” and “growth”; in the event of a large boom, both firms are dissolved and two new firms are created, with the owners of the previous firms receiving the correct number shares of the new firms. Of course, we are not claiming that this is literally what happens, but merely that it is not necessary, either in the model or in the data, for the value of a stream of cash flows to be associated with a single entity over time.

The following sections formally describe the prices and returns on the value sector.

2.4.1 Prices

Let $H^v(D_{t,s}^v, \mu_s, \lambda_s, \tau)$ 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 (14). As in the

⁹Multiperiod returns can be defined by compounding single-period returns.

¹⁰In this stylized model, the growth claim is extreme and pays no current dividends. In reality, what we think of as growth stocks and value stocks would be a combination of the value and growth claim. Writing down such share dynamics would greatly complicate the model without adding to the intuition however. Note that in a version model where growth does pay positive dividends, dividend growth will not be higher for growth firms than for value firms unless a rare boom actually occurs. Thus our model is consistent with the results of Chen (2012), who finds that that expected dividend growth on stocks is not significantly higher than on value stocks in some data subsamples.

case of the aggregate market,

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

Furthermore,

$$F^v(D_{t,s}^v, \mu_s, \lambda_s) = \int_0^\infty H^v(D_{t,s}^v, \mu_s, \lambda_s, \tau) d\tau. \quad (27)$$

The following corollary is a special case of Theorem B.2, given in Appendix B.4.

Corollary 10. *The solution for the function H^v is as follows:*

$$H^v(D_{t,s}^v, \mu_s, \lambda_s, \tau) = D_{t,s}^v \exp \{ a_\phi^v(\tau) + b_{\phi\mu}^v(\tau)^\top \mu_s + b_{\phi\lambda}^v(\tau)^\top \lambda_s \},$$

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

$$b_{\phi\mu_1}^v(\tau) = \frac{\phi - 1}{\kappa_{\mu_1}} (1 - e^{-\kappa_{\mu_1}\tau}) \quad (28)$$

$$b_{\phi\mu_2}^v(\tau) = -\frac{1}{\kappa_{\mu_2}} (1 - e^{-\kappa_{\mu_2}\tau}), \quad (29)$$

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

$$\frac{db_{\phi\lambda_j}^v}{d\tau} = \frac{1}{2}\sigma_{\lambda_j}^2 b_{\phi\lambda_j}^v(\tau)^2 + (b_{\lambda_j}\sigma_{\lambda_j}^2 - \kappa_{\lambda_j}) b_{\phi\lambda_j}^v(\tau) + E_{\nu_j} \left[e^{b_{\mu_j} Z_{jt}} \left(e^{b_{\phi\mu_j}^v(\tau) Z_{jt}} - 1 \right) \right], \quad (30)$$

$$\frac{da_\phi^v}{d\tau} = \bar{\mu}_D - \bar{\mu}_C - \beta + \gamma\sigma^2(1 - \phi) + b_{\phi\lambda_j}^v(\tau)^\top (\kappa_\lambda * \bar{\lambda}) \quad (31)$$

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

It follows from (27) and Corollary 10 that the price-dividend ratio on the value sector is

$$G^v(\mu_t, \lambda_t) = \int_0^\infty \exp(a_\phi^v(\tau) + b_{\phi\mu}^v(\tau)^\top \mu_t + b_{\phi\lambda}^v(\tau)^\top \lambda_t) d\tau. \quad (32)$$

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

Corollary 11. *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.

2.4.2 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.¹¹

Corollary 12. *The value sector premium relative to the risk-free rate r is*

$$r_t^v - r_t = \phi\gamma\sigma^2 - \sum_j \lambda_{jt} E_{\nu_j} \left[(e^{b\mu_j Z_{jt}} - 1) \frac{\mathcal{J}_j(G_t^v)}{G_t^v} \right] - \sum_j \lambda_{jt} \frac{1}{G_t^v} \frac{\partial G^v}{\partial \lambda_j} b_{\lambda_j} \sigma_{\lambda_j}^2 \quad (33)$$

The three terms in (33) have an analogous interpretation to those for the market premium, and can also be signed.

Corollary 13. *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 14. *The observed expected excess return on the value sector in a sample without jumps is*

$$r_{\text{nj},t}^v - r_t = \phi\gamma\sigma^2 - \sum_j \lambda_{jt} E_{\nu_1} \left[e^{b\mu_j Z_{jt}} \frac{\mathcal{J}(G_t^v)}{G_t^v} \right] - \sum_j \lambda_{jt} \frac{1}{G_t^v} \frac{\partial G^v}{\partial \lambda_j} b_{\lambda_j} \sigma_{\lambda_j}^2 \quad (34)$$

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

¹¹The proofs of these results are directly analogous to those for the market, and therefore we do not repeat them.

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 will, in general, 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.

3 Quantitative results

3.1 Data

To calibrate the rare events, we use international consumption data described in detail in Barro and Ursua (2008). These data contain annual observations on real, per capita consumption for 43 countries; 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, Polk, and Vuolteenaho (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.

3.2 Calibration

The parameter set consists of the normal-times parameters $\bar{\mu}_C$, σ and $\bar{\mu}_D$, leverage ϕ , the preference parameters β and γ , the parameters determining the duration of disasters and booms (κ_{μ_1} and κ_{μ_2} respectively), the parameters determining the disaster and boom processes ($\bar{\lambda}_j$, κ_{λ_j} , and σ_{λ_j} for $j = 1, 2$) and finally the distributions of the disasters and booms themselves. Some of these parameters define latent processes for which direct measurement is difficult. The fact that these processes relate to rare events makes the

problem even harder.

For this reason, we proceed by dividing the parameters into groups and impose reasonable restrictions on the parameter space. First, the mean and standard deviation of consumption during normal times are clearly determined by $\bar{\mu}_C$ and σ . We can immediately eliminate two free parameters by setting these equal to their values in the postwar data (see Tables 1 and 3).

Second, to discipline to our calibration, we assume that consumption growth after a disaster reverts to normal at the same rate as consumption growth following a boom, namely, $\kappa_{\mu_1} = \kappa_{\mu_2}$. Further, we assume that the rare event processes are symmetric. That is, we assume that the average probability of a boom equals that of a disaster ($\bar{\lambda}_1 = \bar{\lambda}_2$), and that the processes have the same mean reversion and volatility parameters ($\kappa_{\lambda_1} = \kappa_{\lambda_2}$ and $\sigma_{\lambda_1} = \sigma_{\lambda_2}$).

Third, we calibrate the average disaster probability and the disaster distribution to international consumption data. Barro and Ursua (2008) estimate that the probability of a rare disaster in OECD countries is 2.86%.¹² We use this number as our average disaster probability, $\bar{\lambda}_1$. Following Barro and Jin (2011), we assume a power law distribution for rare events (see Gabaix (2009) for a discussion of the properties of power law distributions). Using maximum likelihood, Barro and Jin estimate a tail parameter of 6.86%. They also argue that the distribution of disasters is better characterized by a double power law, with a lower exponent for larger disasters. Incorporating this more complicated specification would lead to a fatter tail and a higher equity premium and volatility. Thus our parameter choice is conservative.¹³ Following Barro and Ursua (2008), we assume a 10% minimum

¹²We calibrate the size of the disasters to the full set of countries 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.

¹³One concern is that the consumption data on disasters and booms is international, while our stock market data is from the U.S. However, many of the facts that we seek to explain have been reported as robust features of the international data (e.g. Campbell (2003), Fama and French (1992)). We view the international data as disciplining the choice of distribution of the rare events, as the data from the U.S. is

disaster size.

Like the disaster parameters, the parameters for the boom distribution are difficult to observe directly. There are three such parameters: the minimum jump size, the power law parameter, and the mean reversion in booms κ_μ .¹⁴ As was the case in fitting the disaster distribution, we use international data to inform our choices. However, we also want to make sure that we are not assuming growth prospects that are unrealistic given market values. For this reason, we also require the model to match the relative valuations of market and growth (Table 4). Specifically, we require the model to match the relative book-to-market ratio of value (deciles 2–9) as compared with the market as a whole because, in the model, the “book” values of the market and of value are the same. However, requiring the model to match relative market valuations produces similar answers. These choices lead us to a lower minimum jump size for booms (5%) and a thinner tail (the power law parameter is 15 rather than about 7) as compared with disasters. While our procedure may seem imprecise, it is the case that variation in these parameters that keeps the relative size of the value sector the same has very little impact on asset prices. This is because the pricing effect of changes in cash flows is offset by the impact on discount rates.

The disaster distribution and the boom distribution in the model and in the data are reported in Figures 1 and 2. These figures show that, except for small disasters at the five-year horizon, our assumptions imply less extreme distributions than the data. It is the case that our power law distributions are unbounded, thus allowing for small but positive weight on events that are greater in magnitude than what has occurred in the data. We have checked that truncating the distributions has little effect on the results.

The remaining parameters are the dividend process parameters μ_D , and ϕ , preference parameters β and γ , and rare event parameters $\kappa_{\lambda_1} = \kappa_{\lambda_2}$ and $\sigma_{\lambda_1} = \sigma_{\lambda_2}$. We choose these parameters to minimize the the distance between the mean value of various statistics in a

¹⁴Our parameter restrictions imply that $\kappa_{\mu_1} = \kappa_{\mu_2}$. As we discuss in what follows, our choice implies a good empirical fit to the disaster distribution as well as the boom distribution.

sample without rare events and the corresponding statistic in the postwar data. We also impose some reasonable economic limits on the parameter choices from this search.

The first requirement is that the solution to the agent’s problem exists. From Theorem 1, it follows that parameters must satisfy

$$\frac{1}{2\sigma_{\lambda_1}^2} (\kappa_{\lambda_1} + \beta)^2 \geq E_{\nu_1} [e^{b\nu_1 Z_1} - 1]. \quad (35)$$

This is a joint restriction on the size of a disaster, on the agent’s risk aversion, on the discount rate of the agent, and on the persistence and volatility of the disaster probability process. While it might seem that we could make disasters arbitrarily large given the lack of data, thus creating an arbitrarily large risk premium, (35) shows that in this model, at least, this is not the case.¹⁵ Large disasters will make the right hand side of (35) very large. The solution will only exist if changes to λ_1 are short-lived or that the volatility of λ_1 is very small. While it is possible that this is true, such a model will not be able to account for the observed levels of stock market volatility. Restriction (35) limits the choices of disaster parameters.¹⁶

Our second requirement is that the discount rate β be greater than zero. Because of positive consumption growth and a elasticity of intertemporal substitution equal to 1, matching the low riskfree rate of 1.25 will be a challenge. We discuss this aspect of the model’s fit in more detail in a later section. We choose a small positive number for the lower bound of β , and our minimization procedure selects this as optimal on account of the riskfree rate.

Our third requirement is that ϕ be not “too” high. A value of leverage ϕ will help the model match the equity premium and volatility. If we choose to let this parameter value freely, the data will want a high value. In principle, ϕ could be determined from the ratio of normal-times volatility of dividends to normal-times volatility of consumption.

¹⁵Chen, Dou, and Kogan (2013) note this lack of identification as a disadvantage of rare events models.

¹⁶Additional parameter restrictions are discussed in Appendix A; these turn out not to matter for parameter ranges of interest.

This would give a high number of 4.7.¹⁷ However, what really matters for the model is the response of dividends to a disaster. Here, again, what little data we have suggests a high value (Longstaff and Piazzesi (2004) find that earnings fell by more than 100% during the Great Depression period). However, dividends may be expected to mean revert while for simplicity we assume a permanent change. To be conservative, we fix $\phi = 3.5$ as the upper limit of ϕ , which is in line with values considered in the literature (for example, Bansal and Yaron (2004) assume a value of 3.0, while Backus, Chernov, and Martin (2011) assume a value of 5.1). Given that high values of ϕ are helpful for the moments of equity returns, our minimization chooses this upper bound.¹⁸

The restrictions above reduce the parameter set to three. Without loss of generality, we search over μ_D , γ , κ_{λ_1} . The moments we use are average dividend growth, the equity premium, the volatility of the market return, the average price-dividend ratio and the persistence of the value spread.¹⁹ We search over parameters by simulating 1500 60-year samples and taking only those without rare events. We minimize the sum of squared differences between the mean and the data moment, normalizing by the variance from the simulations without rare events. The criterion function is minimized for average dividend growth $\mu_D = 3\%$, risk aversion $\gamma = 3$ and mean reversion $\kappa_{\lambda_1} = 0.11$. Value spread moments are reported in Table 4, while aggregate market moments are reported in Table 5.

One concern that has been raised about rare events models is their sensitivity to parameters that cannot be deduced from consumption data alone (Chen, Dou, and Kogan (2013)). This is a reasonable concern, especially if one chooses parameters that are at the limit of acceptability (say, right at the 5% confidence interval), and if there seems to be a lot of discretion in choosing parameters. Given that our choice for disaster parameters is

¹⁷However, dividend and consumption growth are not perfectly correlated in the data, suggesting that a lower number is more reasonable.

¹⁸We also consider a model in which the normal-times standard deviation of dividend growth is twice that of consumption (rather than 3.5 times, as in our benchmark calibration), but where everything else is the same. The results are very similar to what is reported here.

¹⁹Attempting to match the very high persistence of the price-dividend ratio leads to unstable results.

the same as in previous studies (which were not designed to match the value premium), in this study this concern would primarily apply to the boom parameters. Does it violate rational expectations for investors to price in the probability of a boom, and have it not occur? With our parameters, it does not, because our calibration implies that the probability of not observing a boom in a 60-year period is about 20%. Thus there is no need to assume that the post-war period is exceptional in that a boom has not been observed. Moreover, while increasing the probability of a boom, or the size of the boom if it does occur increases the value premium, it also reduces the equity premium. Thus requiring the model to match the equity premium as well as the value premium is an additional source of discipline.²⁰

3.3 Prices and expected returns as functions of the state variables

3.3.1 Prices

Prices in the model are determined by the two sets of state variables: the drift terms for consumption growth and the rare event probabilities. Mathematically, prices 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 ,

²⁰A related concern is that, given the results of Asness, Moskowitz, and Pedersen (2013), the value premium may be a population rather than a small-sample effect. However, the data on individual equities in Asness et al. come from the U.S., the U.K. and Japan. Given that a large boom would have worldwide implications, impacting at the least the major developed markets, adding data from the U.K. and from Japan does not necessarily help us in observing the correct number or size of the booms. Other data they consider are international equity indices. Our model is a natural fit for explaining these data as well, since the stock markets of some countries might be expected to outperform in the event of a large global boom; these would be “growth” according to their measure and would have lower observed returns. This would also explain the links between the value effects from the international equity indices and the individual equities.

which can be interpreted as the maturity of these claims. Panel A of Figure 3 shows the coefficients on the drift terms for the market ($b_{\phi\mu_j}(\tau)$) and for the value claim ($b_{\phi\mu_j}^v(\tau)$) as a function of τ .

Panel A shows that the coefficients for the market, $b_{\phi\mu_1}(\tau)$ 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 rare events are short-lived.

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. 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 28 with 29; the first of these terms has a ϕ while the second does not). Thus the riskfree rate effect implies that an increase in expected consumption growth arising from booms decreases the price of the value claim.

Panel B shows the functions $b_{\phi\lambda_j}(\tau)$ (which multiply λ_{jt} in the expression for the market price-dividend ratio) and $b_{\phi\lambda_j}^v(\tau)$ (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 smaller than for disasters because of decreasing marginal utility. Consider the last term in (17): $E_{\nu_j} \left[e^{b_{\phi\mu_j} Z_{jt}} \left(e^{b_{\phi\mu_j}(\tau) Z_{jt}} - 1 \right) \right]$. The magnitude of this function is determined in large part by this relatively simple expres-

sion. 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.²¹ 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 the probability of a boom decreases the price of the value claim, because of the riskfree rate effect described above.

3.3.2 Risk premia

Panel A of Figure 4 shows the equity premium (left) and the risk premium on the value claim (right) as a function of the probability of a disaster. These are defined as the expected instantaneous return on the asset less the riskfree rate (Section 2.3.2). 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 4 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.²² For the market portfolio, the static disaster premium

²¹The expression $e^{b_{\phi\mu_j(\tau)} Z_{jt}} - 1$ gives the percent change in the price of zero-coupon equity with maturity τ .

²²Nearly the entire premium for time-variation in risk (the 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

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. 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.

Panel B shows observed expected excess returns in a sample where no rare events take place (see Section 2.3.2 for a formal definition). The observed expected return is a bit higher than the full risk premium, but not by much. Thus it matters little for the equity premium whether disasters are observed in the sample or not.

Panel A of Figure 4 shows these quantities as functions of the probability of a boom. The solid line shows the full premium and the dashed line the static rare event premium. The dashed-dotted line represents the part of the static premium due to booms. 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 discussion 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 the disaster premium, as indicated by Panels A and B of Figure 4. But, as discussed in Sections 2.3.2, the boom premium switches sign, depending on whether booms are observed or not. That is, the population boom premium is more than entirely due to the realized return should a boom take place; in normal times, the investors receive a lower return than if there were no booms. Figure 5 shows that, for the market, the premium for booms lowers the equity premium by 1% 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 (negligible) booms. First, the price of risk for time-varying disasters is much larger in magnitude; b_{λ_1} is 16.1, while b_{λ_2} is -1.8. 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 $\partial G/\partial \lambda_2$).

effect on the riskfree rate, the observed value premium will exceed the premium on the aggregate market.

3.4 Simulation results

To evaluate the quantitative success of the model, we simulate monthly data for 600,000 years, and also simulate 100,000 60-year samples. For each sample, we initialize the λ_{jt} processes using a draw from the stationary distribution.²³ 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.

3.4.1 The aggregate market

Table 3 reports moments of log growth rates of consumption and dividends. There is little skewness or kurtosis in postwar annual consumption data.²⁴ Postwar dividend growth 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 55.

Table 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 5.44%; 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.

²³The stationary distribution for λ_{jt} is Gamma with shape parameter $2\kappa_j\bar{\lambda}_j/\sigma_{\lambda_j}^2$ and scale parameter $\sigma_{\lambda_j}^2/(2\kappa_j)$ (Cox, Ingersoll, and Ross (1985)).

²⁴In the definition of kurtosis that we use, three is the value for the normal distribution.

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.²⁵ However, explaining long-horizon returns in this way implicitly assumes an unrealistically long decision interval for agents.

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

²⁵Constantinides (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.

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.²⁶

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 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 (1.95% 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 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 chang-

²⁶See discussions in Gourio (2008), Nakamura, Steinsson, Barro, and Ursúa (2013) and Wachter (2013).

ing 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.²⁷ 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 is 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, Kiku, and Yaron (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 area for further research.

3.4.2 Unconditional moments of value and growth portfolios

Table 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 2.74%. 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

²⁷The 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.

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% critical value of the value premium is 4.03%, 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 2.4.2 and 3.3.2, it is because investors are willing to accept a lower return on growth in most periods, in return for an occasional very high payout.²⁸

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 traditional value premium. For example, the difference in expected returns between the top and bottom quintile portfolios is 4.3% in postwar data (Table 8). The 95% critical value in our no-jump simulations is 4%. 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 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 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

²⁸A 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.

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 3.3.1 and 3.3.2 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.

3.4.3 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.²⁹ 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 7 shows results of predictability regressions. The table reports mean coefficients and R^2 statistics from simulated samples without jumps, from the full set of samples, and in population.

²⁹Lettau and Wachter (2011) show that if a single factor drives risk premia, then population values of predictive coefficients should be proportional across asset classes.

Panel A of Table 7 shows the results of regressing the aggregate market portfolio return on the price-dividend ratio in the actual and simulated data. The model can reproduce the data finding that the price-dividend ratio predicts excess returns. This result arises primarily from the fact that a high value of the disaster probability implies a higher equity premium and a lower price-dividend ratio. It is also the case that a high value of the boom probability implies a lower observed return in samples without jumps, as well as a higher price-dividend ratio. Coefficients and R^2 statistics are smaller in a sample with jumps and without: this is both because more of the variance of stock returns arises from the greater variance of expected dividends during disasters and because the effect of the boom probability reverses (high premia are associated with high valuations) in the full set of samples. Finally, the population coefficients and R^2 statistics are smaller still because of the well-known small-sample bias.

In the data, the market return can also be predicted by the value spread, though with substantially smaller t -statistics and R^2 values (Panel B of Table 7). The model also captures the sign and the relative magnitude of this predictability. Compared with the price-dividend ratio, the value spread is driven more by the time-varying probability of a boom and less by the probability of a disaster. This explains why risk premia on the market portfolio, which is mainly driven by the disaster probability, are not captured as well by the value spread.

Panel C of Table 7 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.³⁰ The present model does, however, predict a positive

³⁰Roussanov (2014) also notes that the conditional mean of the value-minus-growth portfolio does not vary in the way that univariate models of time-varying risk aversion would predict.

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.³¹

One might think that the reason that the value-minus-growth return cannot be predicted by the price-dividend ratio is that it is not very predictable. This is, however, not the case. Panel D of Table 7 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 10%, 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.³²

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

³¹ The mean 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.

³²In 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 mean coefficient is again positive because of small-sample bias.

kind presented here.

4 Further implications

4.1 When does growth outperform value?

Our model predicts that in samples when booms are not realized, value stocks will on average exhibit greater returns than growth stocks. The model also predicts that there are periods when this is not the case; namely, when growth outperforms value. These will be periods when either booms are realized, or when there are positive shocks to the probability of a boom. Both could be expected to occur during periods of substantial technological innovation.

How does this prediction fare in the data? In this section we examine the performance of value and growth during a period that is indisputably characterized by these shocks, namely the 1990s.

Table 8 shows statistics for portfolios formed by sorting stocks into deciles on the basis of the book-to-market ratio. We examine results for the full CRSP universe, as well as for the top size decile, as our model is arguably more appropriately thought of as a model for large stocks. As is well known, value outperforms growth by a substantial margin over the postwar sample. However, during the 1990s, the greatest performance belongs to the lowest book-to-market decile, namely the most extreme growth portfolio (Panel B). Not only does the top growth portfolio exhibit higher returns during this period, it also has a much higher beta than usual, while value has in fact a lower beta. This is evidence in favor of the principal mechanism in the model: that growth returns are more exposed to boom risk than value stocks.

This exercise naturally raises the question of the differential performance of value and growth during disasters, and during periods when the probability of a disaster increases. To keep this paper of manageable length, we have not introduced differential exposure of

value and growth dividends to disasters. Considering varying exposure to these shocks, however, would be within the spirit of the model. We therefore look at the differential exposure and performance of value and growth during disaster periods.

Panels C and D of Table 8 report the expected returns and betas to value and growth portfolios during the Great Depression and the financial crisis of 2008 respectively. Perhaps surprisingly, value does not appear to underperform growth during these periods. During the Great Depression, the average return on value was about the same as growth, and, while the average return on the top book-to-market decile was very low in 2008, returns on deciles near the top were much less extreme. The picture changes when one looks at beta. Despite having the highest beta over the full period, growth has a relatively low beta during these disaster times. This suggests that growth may indeed be less exposed to disaster risk than value, though why this was not reflected in the performance of value stocks during these periods remains an open question.

Figure 7 shows the returns on the top and bottom quintiles of value and growth, as well as the market during the full sample period. The shaded areas of the figure represent the time intervals evaluated in the last three panels of Table 8. As is well known (Ang and Chen (2007)), value returns exhibit a stronger sensitivity to the market return over the first half of the sample than in the second. It is possible that this reflects a greater sensitivity to the risk of a disaster, or it could indicate a shift in the underlying characteristics of value and growth. The figure also shows that the high beta on value during the Great Depression period was mostly driven by exposure to positive market changes rather than negative ones. However, the higher beta on growth during the 1990s clearly came about both because of greater exposure overall, and specifically greater exposure to positive market changes.

4.2 Downside and upside risk

Ang, Chen, and Xing (2006) define downside β to be the covariance dividend by the variance, where these moments are computed using only those observations when the market

return is below its mean. Likewise, upside β is the measure when the covariance and variance are computed for observations when the market return is above its mean. They show that, in the cross-section stocks with higher downside β s have higher returns, an empirical finding that is also shown by Lettau, Maggiori, and Weber (2013) to hold across asset classes.

Ang, Chen, and Xing (2006) also define relative upside and relative downside β to be the one-sided β measure minus the traditional β . They find that stocks with high relative downside β have higher mean returns while stocks with high relative upside β have lower mean returns. Because there are many sources of heterogeneity in stocks that are not captured in the present study, these relative β results seem most relevant.

That there would be a relation between one-sided risk and rare events is not obvious. Performance during disasters or large booms represents an extreme one-sided risk, while the literature focuses on exposure to differential exposure to up- and down-moves during normal market conditions.

However, in the present model there is a connection because exposure to rare events drives normal-times variation. An asset that is more exposed to disaster risk will also tend to fall during market declines, because a market decline over a finite time interval is more likely to be caused by an increase in the probability of a disaster than a decrease in the probability of a boom. This effect arises from the correlation between the volatility parameters of these processes and the probabilities. Moreover, while changes in boom probabilities in general have a smaller effect on the market return than disaster probabilities, large upward price movements will be disproportionately caused by changes in the boom probability.

Our theory predicts that those assets with high relative downside betas will have high observed expected returns, as in the data. These assets have more exposure to crashes than to booms, even though their overall exposure to changes in the market is similar. Likewise, assets with high relative upside betas have low observed expected returns. These assets have more covariance with boom risk. They have lower expected returns because boom

risk has a low price. Indeed, Table 9 shows that growth assets have high relative upside β s while value assets have high relative downside β s.

4.3 Skewness in the time series and cross-section

This paper predicts that positive skewness should reduce returns, both in the time series and the cross-section. Several recent papers argue that proxies for disaster risk predict future returns (Kelly and Jiang (2013), Manela and Moreira (2013)). Colacito, Ghysels, and Meng (2013) shows that skewness in analysts forecasts, which takes on both negative and positive values, predicts returns with a negative sign. Like other papers on disaster risk, this paper predicts that a greater risk of disaster should be associated with a higher equity premium. It also predicts, consistent with Colacito et al., that a higher chance of a boom will be associated with a lower premium, in a sample where booms do not occur. Figure 8 shows the skewness conditional on the probability of disasters and booms. Indeed, skewness is decreasing in the disaster probability and increasing in the boom probability. Expected returns (in a sample without jumps) go in the other direction.

Measuring skewness whether in the time series or the cross-section, is a challenge. Several recent papers, however, are able to calculate ex ante return skewness using option prices on individual stocks (Conrad, Dittmar, and Ghysels (2013), Chang, Christoffersen, and Jacobs (2013)). These papers show that higher skewness is associated with lower returns in the cross-section, another prediction of this paper. Conrad, Dittmar, and Ghysels (2013) also finds that stocks with higher valuation ratios have higher skewness, again consistent with the results in this paper. Conrad et al. show that stocks with higher book-to-market ratios have lower coskewness, which is the relevant measure in this paper because booms are market-wide. They also report that stocks with higher overall skewness have higher price-to-earnings ratios. These facts are consistent with the finding in this paper that high valuations are tied to the small probability of very high returns.

5 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, 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.

Appendix

A Required conditions on the parameters

Assumption 1.

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

Assumption 2.

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

Assumption 3.

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

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 1 is required for the solution for the value function to exist. This restriction rules out parameters that can lead to infinitely negative utility. Note that for booms, (assuming relative risk aversion is greater than 1) this restriction is satisfied automatically since the right hand side is negative. For disasters, Assumption 1 is an important restriction on the persistence, volatility and size of the disaster process. We discuss this in greater detail in Section 3.2.

Assumptions 2 and 3 guarantee convergence of prices, which are given by integrating (essentially, summing) expected dividends into the infinite future. Assumption 2 ensures that $b_{\phi_{\lambda_2}}(\tau)$ converges as τ approaches infinity; namely that the effect of the boom on future dividends cannot explode as the horizon increases.³³ Assumption 3 states that the asymptotic slope of $a_{\phi}(\tau)$ is negative. This is the dynamic analogue of the condition that

³³Note that no extra assumptions are required for the convergence of $b_{\phi_{\lambda_1}}(\tau)$ because $Z_1 < 0$ and hence $e^{\frac{\phi-1}{\kappa_{\mu_1}} Z_1} < 1$. Nor are extra assumptions required for the value function expression $b_{\phi_{\lambda_2}}^v(\tau)$ to converge since 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.

the growth rate be less than the discount rate in the static Gordon growth model. This condition pertains to the market portfolio. If it is satisfied, the analogous condition for the value claim is satisfied automatically.³⁴

B Detailed derivation of the model

This Appendix derives the results given in the main text. The derivations generalize those in Wachter (2013), 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.³⁵

B.1 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

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

Further, define

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

³⁴Specifically, define

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

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

³⁵Using log-linearization, Eraker and Shaliastovich (2008) and Benzoni, Collin-Dufresne, and Goldstein (2011) find approximate solutions to related continuous-time jump-diffusion models when the EIS is not equal to one.

for $j = 1, 2$, and

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

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 $[\sigma_{\lambda_1}^2, \sigma_{\lambda_2}^2]^\top$.

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 use the notation $\partial h / \partial \lambda$ to denote the 1×2 vector $[\partial^2 h / \partial \lambda_1^2, \partial^2 h / \partial \lambda_2^2]$.

B.2 The value function

Proof of Theorem 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

$$\frac{dS_t}{S_t} = \frac{dC_t}{C_t} = \mu_{C_t} dt + \sigma dB_{C_t}. \quad (\text{B.1})$$

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

$$dW_t = (W_t \alpha_t (\mu_{C_t} - r_t + l^{-1}) + W_t r_t - c_t) dt + W_t \alpha_t \sigma dB_{C_t},$$

where r_t denote the instantaneous risk-free rate. Optimal consumption and portfolio choice must satisfy the following Hamilton-Jacobi-Bellman equation:

$$\sup_{\alpha_t, c_t} \left\{ \frac{\partial J}{\partial W} (W_t \alpha_t (\mu_{C_t} - r_t + l^{-1}) + W_t r_t - c_t) + \frac{\partial J}{\partial \lambda} (\kappa_\lambda * (\bar{\lambda} - \lambda_t)) - \frac{\partial J}{\partial \mu} (\kappa_\mu * \mu_t) + \frac{1}{2} \frac{\partial^2 J}{\partial W^2} W_t^2 \alpha_t^2 \sigma^2 + \frac{1}{2} \left(\frac{\partial^2 J}{\partial \lambda^2} \right)^\top (\sigma_\lambda^2 * \lambda_t) + \lambda_t^\top \bar{\mathcal{J}}(J(W_t, \mu_t, \lambda_t)) + f(c_t, V) \right\} = 0, \quad (\text{B.2})$$

where, as defined in Appendix B.1,

$$\begin{aligned}\frac{\partial^2 J}{\partial \lambda^2} &= \begin{bmatrix} \frac{\partial^2 J}{\partial \lambda_1^2} & \frac{\partial^2 J}{\partial \lambda_2^2} \end{bmatrix}^\top \\ \sigma_\lambda^2 &= [\sigma_{\lambda_1}^2, \sigma_{\lambda_2}^2]^\top.\end{aligned}$$

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

$$\begin{aligned}\frac{\partial J}{\partial W} W_t \mu_{Ct} + \frac{\partial J}{\partial \lambda} (\kappa_\lambda * (\bar{\lambda} - \lambda_t)) - \frac{\partial J}{\partial \mu} (\kappa_\mu * \mu_t) + \frac{1}{2} \frac{\partial^2 J}{\partial W^2} W_t^2 \sigma^2 \\ + \frac{1}{2} \left(\frac{\partial^2 J}{\partial \lambda^2} \right)^\top (\sigma_\lambda^2 * \lambda_t) + \lambda_t^\top \bar{\mathcal{J}}(J(W_t, \mu_t, \lambda_t)) + f(C_t, V) = 0.\end{aligned}\quad (\text{B.3})$$

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

$$\begin{aligned}f(C_t, V_t) &= f(W_t l^{-1}, J(W_t, \mu_t, \lambda_t)) \\ &= \beta W_t^{1-\gamma} I(\mu_t, \lambda_t) \left(\log \beta - \frac{\log I(\mu_t, \lambda_t)}{1-\gamma} \right).\end{aligned}\quad (\text{B.4})$$

Substituting (B.4) and (7) into (B.3)

$$\begin{aligned}\mu_{Ct} + (1-\gamma)^{-1} I^{-1} \frac{\partial I}{\partial \lambda} (\kappa_\lambda * (\bar{\lambda} - \lambda_t)) - (1-\gamma)^{-1} I^{-1} \frac{\partial I}{\partial \mu} (\kappa_\mu * \mu_t) - \frac{1}{2} \gamma \sigma^2 \\ + \frac{1}{2} (1-\gamma)^{-1} I^{-1} \left(\frac{\partial^2 I}{\partial \lambda^2} \right)^\top (\sigma_\lambda^2 * \lambda_t) + (1-\gamma)^{-1} \lambda_t^\top \bar{\mathcal{J}}(I(\mu_t, \lambda_t)) \\ + \beta \left(\log \beta - \frac{\log I(\mu_t, \lambda_t)}{1-\gamma} \right) = 0.\end{aligned}$$

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

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

Collecting the constant terms:

$$a = \frac{1 - \gamma}{\beta} \left(\bar{\mu}_C - \frac{1}{2} \gamma \sigma^2 \right) + (1 - \gamma) \log \beta + \sum_j b_{\lambda_j} \frac{\kappa_{\lambda_j}}{\beta} \bar{\lambda}_j.$$

□

Proof of Corollary 3 The risk-free rate is obtained by taking the derivative of the HJB (B.2) with respect to α_t , evaluating at $\alpha_t = 1$ and setting it equal to 0. The result immediately follows. □

B.3 The state-price density

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

$$\pi_t = \exp \left\{ \int_0^t \frac{\partial}{\partial V} f(C_s, V_s) ds \right\} \frac{\partial}{\partial C} f(C_t, V_t). \quad (\text{B.5})$$

Note that the exponential term is deterministic. From (6), we obtain

$$\frac{\partial}{\partial C} f(C_t, V_t) = \beta (1 - \gamma) \frac{V_t}{C_t}.$$

The equilibrium condition $V_t = J(\beta^{-1} C_t, \mu_t, \lambda_t)$, together with the form of the value function (7), implies

$$\frac{\partial}{\partial C} f(C_t, V_t) = \beta^\gamma C_t^{-\gamma} I(\mu_t, \lambda_t). \quad (\text{B.6})$$

Applying Ito's Lemma to (B.6) implies

$$\frac{d\pi_t}{\pi_t^-} = \mu_{\pi t} dt + \sigma_{\pi t} dB_t + \sum_j \frac{\mathcal{J}_j(\pi_t)}{\pi_t^-} dN_{jt}, \quad (\text{B.7})$$

where

$$\sigma_{\pi t} = \left[-\gamma \sigma, \quad b_{\lambda_1} \sigma_{\lambda_1} \sqrt{\lambda_{1t}}, \quad b_{\lambda_2} \sigma_{\lambda_2} \sqrt{\lambda_{2t}} \right], \quad (\text{B.8})$$

and

$$\mathcal{J}_j(\pi_t) = e^{b_{\mu_j} Z_{jt}} - 1, \quad (\text{B.9})$$

for $j = 1, 2$. It also follows from no-arbitrage that

$$\mu_{\pi t} = -r_t - \lambda_t^\top \frac{\bar{\mathcal{J}}(\pi_t)}{\pi_t} \quad (\text{B.10})$$

$$\begin{aligned} &= -r_t - \sum_j \lambda_{jt} E_{\nu_j} [e^{b_{\mu_j} Z_{jt}} - 1] \\ &= -\beta - \mu_{Ct} + \gamma\sigma^2 - \sum_j \lambda_{jt} E_{\nu_j} [e^{b_{\mu_j} Z_{jt}} - 1]. \end{aligned} \quad (\text{B.11})$$

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 (B.7). The first element of (B.8) 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 (B.8).

B.4 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

$$\frac{dD_t}{D_t} = \mu_{Dt} dt + \sigma_D dB_{Ct}, \quad (\text{B.12})$$

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 B.1. *Let $H(D_t, \mu_t, \lambda_t, \tau)$ 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

$$\mu_{\pi,t} + \mu_{H(\tau),t} + \sigma_{\pi,t}\sigma_{H(\tau),t}^\top + \frac{1}{\pi_t H_t} \lambda_t^\top \bar{\mathcal{J}}(\pi_t H_t) = 0. \quad (\text{B.13})$$

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

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

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

$$\begin{aligned} \pi_t H_t = \pi_0 H_0 + \int_0^t \pi_s H_s (\mu_{H,s} + \mu_{\pi,s} + \sigma_{\pi,s}\sigma_{H,s}^\top) + \int_0^t \pi_s H_s (\sigma_{H,s} + \sigma_{\pi,s})dB_s \\ + \sum_j \sum_{0 < s_{ij} \leq t} (\pi_{s_{ij}} H_{s_{ij}} - \pi_{s_{ij}^-} H_{s_{ij}^-}), \end{aligned} \quad (\text{B.14})$$

where $s_{ij} = \inf\{s : N_{js} = i\}$ (namely, the time that the i th type j jump occurs). Adding and subtracting the jump compensation term from (B.14) yields:

$$\begin{aligned} \pi_t H_t = \pi_0 H_0 + \int_0^t \pi_s H_s \left(\mu_{H,s} + \mu_{\pi,s} + \sigma_{\pi,s}\sigma_{H,s}^\top + \sum_j \lambda_j \frac{\bar{\mathcal{J}}_j(\pi_s H_s)}{\pi_s H_s} \right) ds \\ + \int_0^t \pi_s H_s (\sigma_{H,s} + \sigma_{\pi,s})dB_s \\ + \sum_j \left(\sum_{0 < s_{ij} \leq t} (\pi_{s_{ij}} H_{s_{ij}} - \pi_{s_{ij}^-} H_{s_{ij}^-}) - \int_0^t \pi_s H_s \lambda_j \bar{\mathcal{J}}_j(\pi_s H_s) ds \right). \end{aligned} \quad (\text{B.15})$$

Under regularity conditions analogous to those given in Duffie, Pan, and Singleton (2000) the second and the third integrals on the right hand side of (B.15) are martingales. Therefore the first integral on the right hand side of (B.15) must also be a martingale, and it follows that the integrand of this term must equal zero. \square

Theorem B.2. *The function H takes an exponential form:*

$$H(D_t, \mu_t, \lambda_t, \tau) = D_t \exp \{ a_\phi(\tau) + b_{\phi\mu}(\tau)^\top \mu_t + b_{\phi\lambda}(\tau)^\top \lambda_t \}, \quad (\text{B.16})$$

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

$$\frac{db_{\phi\mu_j}}{d\tau} = -\kappa_{\mu_j} b_{\phi\mu_j} + (\phi_{D,j} - 1), \quad (\text{B.17})$$

$$\frac{db_{\phi\lambda_j}}{d\tau} = \frac{1}{2} \sigma_{\lambda_j}^2 b_{\phi\lambda_j}(\tau)^2 + (b_{\lambda_j} \sigma_{\lambda_j}^2 - \kappa_{\lambda_j}) b_{\phi\lambda_j}(\tau) + E_{\nu_j} \left[e^{b_{\mu_j} Z_{jt}} \left(e^{b_{\phi\mu_j}(\tau) Z_{jt}} - 1 \right) \right], \quad (\text{B.18})$$

$$\frac{da_\phi}{d\tau} = \bar{\mu}_D - \bar{\mu}_C - \beta + \gamma \sigma (\sigma - \sigma_D) + b_{\phi\lambda}(\tau)^\top (\kappa_\lambda * \bar{\lambda}). \quad (\text{B.19})$$

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

$$\frac{\bar{J}_j(\pi_t H_t)}{\pi_t H_t} = E_{\nu_j} \left[e^{(b_{\mu_j} + b_{\mu_j} \phi(\tau)) Z_{jt}} - 1 \right], \quad (\text{B.20})$$

$$\begin{aligned} \mu_{H(\tau),t} &= \frac{1}{H} \left(\frac{\partial H}{\partial D} \mu_{Dt} + \frac{\partial H}{\partial \lambda} (\kappa_\lambda * (\bar{\lambda} - \lambda_t)) - \frac{\partial H}{\partial \mu} (\kappa_\mu * \mu_t) \right. \\ &\quad \left. - \frac{\partial H}{\partial \tau} + \frac{1}{2} \left(\frac{\partial^2 H}{\partial \lambda^2} \right) (\sigma_\lambda^2 * \lambda_t) \right) \end{aligned} \quad (\text{B.21})$$

$$\begin{aligned} &= \mu_{Dt} + b_{\phi\lambda}(\tau)^\top (\kappa_\lambda * (\bar{\lambda} - \lambda_t)) + b_{\phi\mu}(\tau)^\top (\kappa_\mu * \mu_t) \\ &\quad - \left(\frac{da_\phi}{d\tau} + \lambda_t^\top \frac{db_{\phi\lambda}}{d\tau} + \mu_t^\top \frac{db_{\phi\mu}}{d\tau} \right) + \frac{1}{2} (b_{\phi\lambda}(\tau)^2)^\top (\sigma_\lambda^2 * \lambda_t), \end{aligned} \quad (\text{B.22})$$

and

$$\begin{aligned} \sigma_{H(\tau),t} &= \frac{1}{H} \left(\frac{\partial H}{\partial D} \mu_D [\sigma_D, 0, 0] + \frac{\partial H}{\partial \lambda_1} [0, \sigma_{\lambda_1} \sqrt{\lambda_{1t}}, 0] + \frac{\partial H}{\partial \lambda_2} [0, 0, \sigma_{\lambda_2} \sqrt{\lambda_{2t}}] \right) \\ &= \left[\sigma_D, b_{\phi\lambda_1}(\tau) \sigma_{\lambda_1} \sqrt{\lambda_{1t}}, b_{\phi\lambda_2}(\tau) \sigma_{\lambda_2} \sqrt{\lambda_{2t}} \right]. \end{aligned} \quad (\text{B.23})$$

Substituting (B.20), (B.22) and (B.23) along with (B.8) and (B.11) into the no-arbitrage condition (B.13) implies

$$\begin{aligned} &\mu_{Dt} + b_{\phi\lambda}(\tau)^\top (\kappa_\lambda * (\bar{\lambda} - \lambda_t)) + b_{\phi\mu}(\tau)^\top (\kappa_\mu * \mu_t) + (b_{\phi\lambda}(\tau)^2)^\top (\sigma_\lambda^2 * \lambda_t) \\ &\quad - \beta - \mu_{Ct} + \gamma \sigma^2 - \gamma \sigma \sigma_D + \sum_j \lambda_{jt} E_{\nu_j} \left[e^{(b_{\mu_j} + b_{\mu_j} \phi(\tau)) Z_{jt}} - e^{b_{\mu_j} Z_{jt}} \right] \\ &\quad - \left(\frac{da_\phi}{d\tau} + \lambda_t^\top \frac{db_{\phi\lambda}}{d\tau} + \mu_t^\top \frac{db_{\phi\mu}}{d\tau} \right) = 0. \end{aligned}$$

Notice that, by definition, $\mu_{D_t} - \mu_{C_t} = (\bar{\mu}_D - \bar{\mu}_C) + \sum_j (\phi_{D,j} - 1)\mu_{jt}$. Matching the terms multiplying μ_j implies (B.17), matching the terms multiplying λ_j implies (B.18) and matching the constant terms implies (B.19). \square

Let $F_t = F(D_t, \mu_t, \lambda_t)$ denote the time t price of the claim to the dividend stream defined by (B.12).

Lemma B.3. *No-arbitrage implies*

$$\mu_{\pi,t} + \mu_{F,t} + \frac{D_t}{F_t} + \sigma_{\pi,t}\sigma_{F,t}^\top + \sum_j \lambda_{jt} \frac{\bar{\mathcal{J}}_j(\pi_t F_t)}{\pi_t F_t} = 0, \quad (\text{B.24})$$

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

$$F(D_t, \mu_t, \lambda_t)\mu_{F,t} = \int_0^\infty \left(H(\tau)\mu_{D_t} + \sum_j H_{\lambda_j}(\tau)(\bar{\lambda}_j - \lambda_j) + \sum_j H_{\mu_j}(\tau)\mu_j + \frac{1}{2} \sum_j H_{\lambda_j \lambda_j}(\tau) \right) d\tau,$$

where H_{μ_j} , H_{λ_j} and $H_{\lambda_j \lambda_j}$ denote partial derivatives. It then follows from the equation for $\mu_{H(\tau),t}$ (B.21) that

$$F(D_t, \mu_t, \lambda_t)\mu_{F,t} = \int_0^\infty \left(H(D_t, \lambda_t, \mu_t, \tau)\mu_{H(\tau),t} - \frac{\partial}{\partial \tau} H(D_t, \mu_t, \lambda_t, \tau) \right) d\tau. \quad (\text{B.25})$$

In short, (B.25) holds because H is a function of τ but F is not. Because $\lim_{\tau \rightarrow \infty} H(D_t, \mu_t, \lambda_t, \tau) = 0$,

$$- \int_0^\infty \frac{\partial}{\partial \tau} H(D_t, \mu_t, \lambda_t, \tau) d\tau = H(D_t, \mu_t, \lambda_t, 0) = D_t.$$

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

$$\bar{\mathcal{J}}(\pi_t F(D_t, \mu_t, \lambda_t)) = \int_0^\infty \bar{\mathcal{J}}(\pi_t H(D_t, \mu_t, \lambda_t, \tau)) d\tau$$

The result then follows from the no-arbitrage relation for H , (B.13). \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 B.4. *Let r_t^e denote the instantaneous expected return on the general equity claim.*

Then

$$r_t^e - r_t = -\sigma_{\pi,t} \sigma_{F,t}^\top - \sum_j \lambda_{jt} E_{\nu_j} \left[\frac{\mathcal{J}_j(F_t)}{F_t} \frac{\mathcal{J}_j(\pi_t)}{\pi_t} \right]. \quad (\text{B.26})$$

Proof It follows from the definition of r_t^e (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, μ_{π_t} can be written in terms of r_t and a jump term as in (B.10). 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 (B.24) in Lemma B.3. \square

B.5 Further results on equity pricing

The following is an intermediate step in the proof of Corollary 5:

Lemma B.5.

$$\lim_{\tau \rightarrow \infty} b_{\phi \lambda_j}(\tau) = -\frac{1}{\sigma_{\lambda_j}^2} \left(\zeta_{\phi_j} - \kappa_{\lambda_j} + b_{\lambda_j} \sigma_{\lambda_j}^2 \right), \quad (\text{B.27})$$

where

$$\zeta_{\phi_j} = \sqrt{(b_{\lambda_j} \sigma_{\lambda_j}^2 - \kappa_{\lambda_j})^2 - 2E_{\nu_j} \left[e^{\left(b_{\mu_j} + \frac{\phi-1}{\kappa_{\mu_j}} \right) Z_j} - e^{b_{\mu_j} Z_j} \right] \sigma_{\lambda_j}^2}. \quad (\text{B.28})$$

Moreover, $\lim_{\tau \rightarrow \infty} b_{\phi \lambda_1}(\tau) < 0$ and $\lim_{\tau \rightarrow \infty} b_{\phi \lambda_2}(\tau) > 0$.

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 (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 (B.27).³⁶

To prove that the limits have the signs given in the Lemma, note that $Z_1 < 0$ implies that

$$E_{\nu_1} \left[e^{(b_{\mu_1} + \frac{\phi-1}{\kappa_{\mu_1}})Z_1} - e^{b_{\mu_1}Z_1} \right] < 0.$$

Therefore,

$$\zeta_{\phi_1} > |b_{\lambda_1}\sigma_{\lambda_1}^2 - \kappa_{\lambda_1}|.$$

Now, note that $Z_2 > 0$ implies that

$$E_{\nu_1} \left[e^{(b_{\mu_1} + \frac{\phi-1}{\kappa_{\mu_1}})Z_1} - e^{b_{\mu_1}Z_1} \right] > 0.$$

The parameter assumptions imply that ζ_{ϕ_2} is real-valued. As shown in Corollary 2, $b_{\lambda_2} < 0$, and that

$$\zeta_{\phi_2} < |b_{\lambda_2}\sigma_{\lambda_2}^2 - \kappa_{\lambda_2}|$$

In both cases the result on the sign follows. \square

Proof of Corollary 5 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 $\lim_{\tau \rightarrow \infty} b_{\phi\lambda_1}(\tau) < 0$ by Lemma B.5. Therefore, it suffices to show that $b_{\phi\lambda_1}(\tau)$ is a monotonic function of τ .

Assume, by contradiction that $db_{\phi\lambda_1}(\tau)/d\tau = 0$ for some τ, τ^* . Then, by (17),

$$b_{\phi\lambda_1}(\tau^*) = \frac{1}{\sigma_{\lambda_1}^2} \left(\sqrt{(b_{\lambda_1}\sigma_{\lambda_1}^2 - \kappa_{\lambda_1})^2 - 2E_{\nu_1} \left[e^{(b_{\mu_1} + b_{\phi\mu_1}(\tau^*))Z_1} - e^{b_{\mu_1}Z_1} \right] \sigma_{\lambda_1}^2 - \kappa_{\lambda_1} + b_{\lambda_1}\sigma_{\lambda_1}^2} \right) \quad (\text{B.29})$$

However, differentiating (B.29) 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 (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 $j = 2$. \square

³⁶We have verified that (B.27) does indeed correspond to the limit when the ordinary differential equation (17) is solved numerically.

Proof of Corollary 6 It follows from Ito's Lemma and the definition of G that

$$\sigma_{F,t} = \left[\phi\sigma_D, \frac{1}{G} \frac{\partial G}{\partial \lambda_1} \sigma_{\lambda_1} \sqrt{\lambda_{1t}}, \frac{1}{G} \frac{\partial G}{\partial \lambda_2} \sigma_{\lambda_2} \sqrt{\lambda_{2t}} \right]. \quad (\text{B.30})$$

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 (B.8) and (B.9)) into (B.26) of Theorem B.4. \square

Note that the proof of Corollary 12 follows along similar lines.

C 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

$$\begin{aligned} R_{t,t+\Delta t} &\approx \frac{F_{t+\Delta t} + D_{t+\Delta t}\Delta t}{F_t} \\ &= \frac{\frac{F_{t+\Delta t}}{D_{t+\Delta t}} + \Delta t \frac{D_{t+\Delta t}}{D_t}}{\frac{F_t}{D_t}} \\ &= \frac{G(\mu_{t+\Delta t}, \lambda_{t+\Delta t}) + \Delta t \frac{D_{t+\Delta t}}{D_t}}{G(\lambda_t)}. \end{aligned}$$

Computing the return on the growth sector requires a different approach. For $u \geq s \geq t$, let $R_{t,s,u}^g$ 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,

$$R_{t,t+\Delta t}^m = \frac{F_{t,t}^v}{F_t} R_{t,t+\Delta t}^v + \left(1 - \frac{F_{t,t}^v}{F_t} \right) R_{t,t,t+\Delta t}^g.$$

Rearranging, it follows that one-period returns on the growth sector equal

$$R_{t,t+\Delta t}^g = \frac{1}{1 - \frac{F_{t,t}^v}{F_t}} \left(R_{t,t+\Delta t}^m - \frac{F_{t,t}^v}{F_t} R_{t,t+\Delta t}^v \right). \quad (\text{C.1})$$

Because the price of the value sector formed at time t relative to the aggregate market is given by

$$\frac{F_{t,t}^v}{F_t} = \frac{G^v(\mu_t, \lambda_t)}{G(\mu_t, \lambda_t)},$$

it is straightforward to compute the return (C.1) on the growth sector.

References

- Abel, Andrew, 1999, Risk premia and term premia in general equilibrium, *Journal of Monetary Economics* 43, 3–33.
- Ai, Hengjie, and Dana Kiku, 2013, Growth to value: Option exercise and the cross section of equity returns, *Journal of Financial Economics* 107, 325–349.
- Ang, Andrew, and Joseph Chen, 2007, CAPM over the long run: 1926-2001, *Journal of Empirical Finance* 14, 1–40.
- Ang, A., J. Chen, and Y. Xing, 2006, Downside Risk, *Review of Financial Studies* 19, 1191–1239.
- Asness, Clifford S., Tobias J. Moskowitz, and Lasse Heje Pedersen, 2013, Value and Momentum Everywhere: Value and Momentum Everywhere, *The Journal of Finance* 68, 929–985.
- Backus, David, Mikhail Chernov, and Ian Martin, 2011, Disasters Implied by Equity Index Options, *The Journal of Finance* 66, 1969–2012.
- Bansal, Ravi, Robert F. Dittmar, and Christian T. Lundblad, 2005, Consumption, dividends, and the cross-section of equity returns, *Journal of Finance* 60, 1639–1672.
- Bansal, Ravi, Dana Kiku, and Amir Yaron, 2012, An Empirical Evaluation of the Long-Run Risks Model for Asset Prices, *Critical Finance Review* 1, 183–221.
- Bansal, Ravi, and Amir Yaron, 2004, Risks for the long-run: A potential resolution of asset pricing puzzles, *Journal of Finance* 59, 1481–1509.
- Barro, Robert J., 2006, Rare disasters and asset markets in the twentieth century, *Quarterly Journal of Economics* 121, 823–866.
- Barro, Robert J., and Tao Jin, 2011, On the size distribution of macroeconomic disasters, *Econometrica* 79, 1567–1589.

- Barro, Robert J., and Jose F. Ursua, 2008, Macroeconomic crises since 1870, *Brookings Papers on Economic Activity* no. 1, 255–350.
- Beeler, Jason, and John Y. Campbell, 2012, The Long-Run Risks Model and Aggregate Asset Prices: An Empirical Assessment, *Critical Finance Review* 1, 141–182.
- Bekaert, Geert, and Eric Engstrom, 2013, Asset Return Dynamics under Bad Environment Good Environment Fundamentals, National Bureau of Economic Research, Working paper 15222.
- Benzoni, Luca, Pierre Collin-Dufresne, and Robert S. Goldstein, 2011, Explaining asset pricing puzzles associated with the 1987 market crash, *Journal of Financial Economics* 101, 552 – 573.
- Berk, Jonathan B., Richard C. Green, and Vasant Naik, 1999, Optimal investment, growth options, and security returns, *Journal of Finance* 54, 1553–1607.
- Binsbergen, Jules H. van, Michael W. Brandt, and Ralph S. Koijen, 2012, On the timing and pricing of dividend, *American Economic Review* 102, 1596–1618.
- Breenden, Douglas T., 1979, An intertemporal asset pricing model with stochastic consumption and investment opportunities, *Journal of Financial Economics* 7, 265–296.
- Campbell, John Y., 2003, Consumption-based asset pricing, in G. Constantinides, M. Harris, and R. Stulz, eds.: *Handbook of the Economics of Finance, vol. 1b* (Elsevier Science, North-Holland).
- Campbell, John Y., and Robert J. Shiller, 1988, The dividend-price ratio and expectations of future dividends and discount factors, *Review of Financial Studies* 1, 195–228.
- Campbell, John Y., and Tuomo Vuolteenaho, 2004, Bad beta, good beta, *American Economic Review* 94, 1249–1275.

- Carlson, Murray, Adlai Fisher, and Ron Giammarino, 2004, Corporate investment and asset price dynamics: Implications for the cross-section of returns, *Journal of Finance* 59, 2577–2603.
- Chang, Bo Young, Peter Christoffersen, and Kris Jacobs, 2013, Market skewness risk and the cross section of stock returns, *Journal of Financial Economics* 107, 46–68.
- Chen, Hui, Winston Wei Dou, and Leonid Kogan, 2013, Measuring the “Dark Matter” in Asset Pricing Models, Working paper, MIT.
- Chen, Huafeng (Jason), 2012, Do Cash Flows of Growth Stocks Really Grow Faster?, Working paper, University of British Columbia.
- Cochrane, John H., 2011, Presidential Address: Discount Rates, *The Journal of Finance* 66, 1047–1108.
- Cochrane, John H., Francis A. Longstaff, and Pedro Santa-Clara, 2008, Two Trees, *Review of Financial Studies* 21, 347–385.
- Cohen, Randolph B., Christopher Polk, and Tuomo Vuolteenaho, 2003, The value spread, *Journal of Finance* 58, 609–641.
- Colacito, Riccardo, Eric Ghysels, and Jinghan Meng, 2013, Skewness in Expected Macro Fundamentals and the Predictability of Equity Returns: Evidence and Theory, Working Paper.
- Conrad, Jennifer, Robert F. Dittmar, and Eric Ghysels, 2013, Ex Ante Skewness and Expected Stock Returns, *Journal of Finance* 68, 85–124.
- Constantinides, George M., 2008, Discussion of “Macroeconomic crises since 1870”, *Brookings Papers on Economic Activity* no. 1, 336–350.
- Cox, John C., Jonathan C. Ingersoll, and Stephen A. Ross, 1985, A theory of the term structure of interest rates, *Econometrica* 53, 385–408.

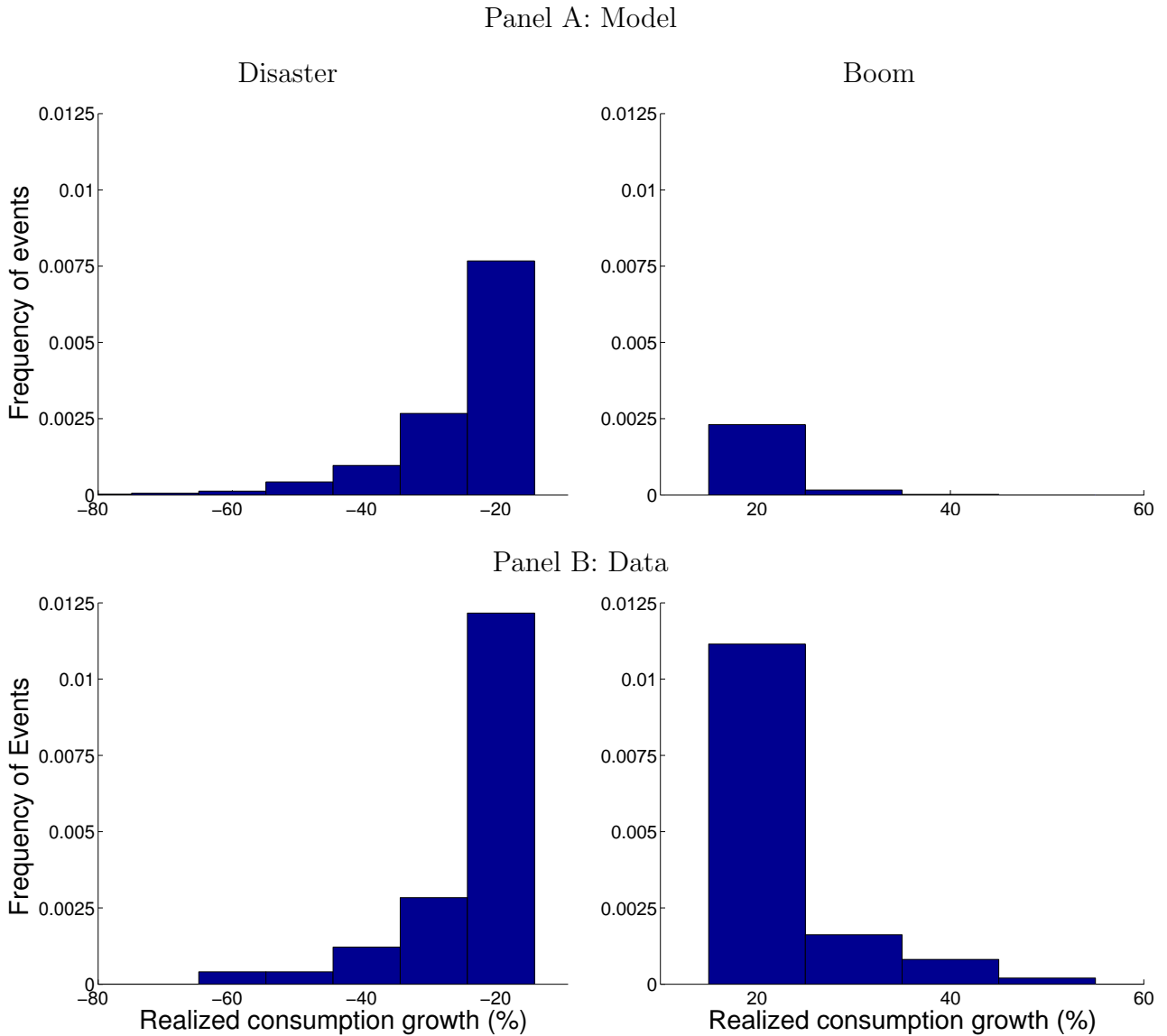
- Da, Zhi, 2009, Cash Flow, Consumption Risk, and the Cross-section of Stock Returns, *The Journal of Finance* 64, 923–956.
- Duffie, Darrell, 2001, *Dynamic Asset Pricing Theory*. (Princeton University Press Princeton, NJ) 3 edn.
- Duffie, Darrell, and Larry G Epstein, 1992, Asset pricing with stochastic differential utility, *Review of Financial Studies* 5, 411–436.
- Duffie, Darrell, Jun Pan, and Kenneth Singleton, 2000, Transform analysis and asset pricing for affine jump-diffusions, *Econometrica* 68, 1343–1376.
- Duffie, Darrell, and Costis Skiadas, 1994, Continuous-time asset pricing: A utility gradient approach, *Journal of Mathematical Economics* 23, 107–132.
- Epstein, Larry, and Stan Zin, 1989, Substitution, risk aversion and the temporal behavior of consumption and asset returns: A theoretical framework, *Econometrica* 57, 937–969.
- Eraker, Bjorn, and Ivan Shaliastovich, 2008, An equilibrium guide to designing affine pricing models, *Mathematical Finance* 18, 519–543.
- Fama, Eugene F., and Kenneth R. French, 1992, The cross-section of expected returns, *Journal of Finance* 47, 427–465.
- Gabaix, Xavier, 2009, Power laws in economics and finance, *Annual review of economics* 1, 255–293.
- Gabaix, Xavier, 2012, An exactly solved framework for ten puzzles in macro-finance, *Quarterly Journal of Economics* 127, 645–700.
- Gârleanu, Nicolae, Leonid Kogan, and Stavros Panageas, 2012, Displacement risk and asset returns, *Journal of Financial Economics* 105, 491–510.
- Gomes, Joao, Leonid Kogan, and Lu Zhang, 2003, Equilibrium cross section of returns, *Journal of Political Economy* 111, 693–732.

- Gourio, François, 2008, Time-series predictability in the disaster model, *Finance Research Letters* 5, 191–203.
- Gourio, François, 2012, Disaster Risk and Business Cycles, *American Economic Review* 102, 2734–2766.
- Hansen, Lars Peter, 2012, Dynamic valuation decomposition within stochastic economies, *Econometrica* 80, 911–967.
- Hansen, Lars Peter, John C. Heaton, and Nan Li, 2008, Consumption strikes back? Measuring long run risk, *Journal of Political Economy* 116, 260–302.
- Jovanovic, Boyan, and Peter L. Rousseau, 2003, Two Technological Revolutions, *Journal of the European Economic Association* 1, pp. 419–428.
- Julliard, Christian, and Anisha Ghosh, 2012, Can Rare Events Explain the Equity Premium Puzzle?, *The Review of Financial Studies* 25, 3037–3076.
- Kelly, Bryan, and Hao Jiang, 2013, Tail Risk and Asset Prices, Working Paper.
- Kiku, Dana, 2006, Is the value premium a puzzle?, Working paper, University of Pennsylvania.
- Kogan, Leonid, and Dimitris Papanikolaou, 2013, Firm characteristics and stock returns: The role of investment-specific shocks, *Review of Financial Studies* 26, 2718–2759.
- Kogan, Leonid, Dimitris Papanikolaou, and Noah Stoffman, 2013, Technological innovation: Winners and losers, National Bureau of Economic Research Working Paper 18671.
- Lettau, Martin, Matteo Maggiori, and Michael Weber, 2013, Conditional Risk Premia in Currency Markets and Other Asset Classes, NBER Working Paper # 18844.
- Lettau, Martin, and Jessica A. Wachter, 2007, Why is long-horizon equity less risky? A duration-based explanation of the value premium, *Journal of Finance* 62, 55–92.

- Lettau, Martin, and Jessica A. Wachter, 2011, The term structures of equity and interest rates, *Journal of Financial Economics* 101, 90 – 113.
- Longstaff, Francis A., 2000, Arbitrage and the expectations hypothesis, *Journal of Finance* 55, 989–994.
- Longstaff, Francis A., and Monika Piazzesi, 2004, Corporate earnings and the equity premium, *Journal of Financial Economics* 74, 401–421.
- Manela, Asaf, and Alan Moreira, 2013, News implied volatility and disaster concerns, Working paper, Washington University and Yale University.
- Martin, Ian, 2013, The Lucas Orchard, *Econometrica* 81, 55–111.
- Mehra, Rajnish, and Edward Prescott, 1988, The Equity Risk Premium: A Solution?, *Journal of Monetary Economics* 22, 133–136.
- Nakamura, Emi, Jón Steinsson, Robert Barro, and José Ursúa, 2013, Crises and Recoveries in an Empirical Model of Consumption Disasters, *American Economic Journal: Macroeconomics* 5, 35–74.
- Novy-Marx, Robert, 2010, Operating Leverage, *Review of Finance* 15, 103–134.
- Papanikolaou, Dimitris, 2011, Investment Shocks and Asset Prices, *Journal of Political Economy* 119, 639–685.
- Pástor, Ľuboš, and Pietro Veronesi, 2009, Technological Revolutions and Stock Prices, *The American Economic Review* 99, pp. 1451–1483.
- Rietz, Thomas A., 1988, The equity risk premium: A solution, *Journal of Monetary Economics* 22, 117–131.
- Roussanov, Nikolai, 2014, Composition of wealth, conditioning information, and the cross-section of stock returns, *Journal of Financial Economics* 111, 352–380.

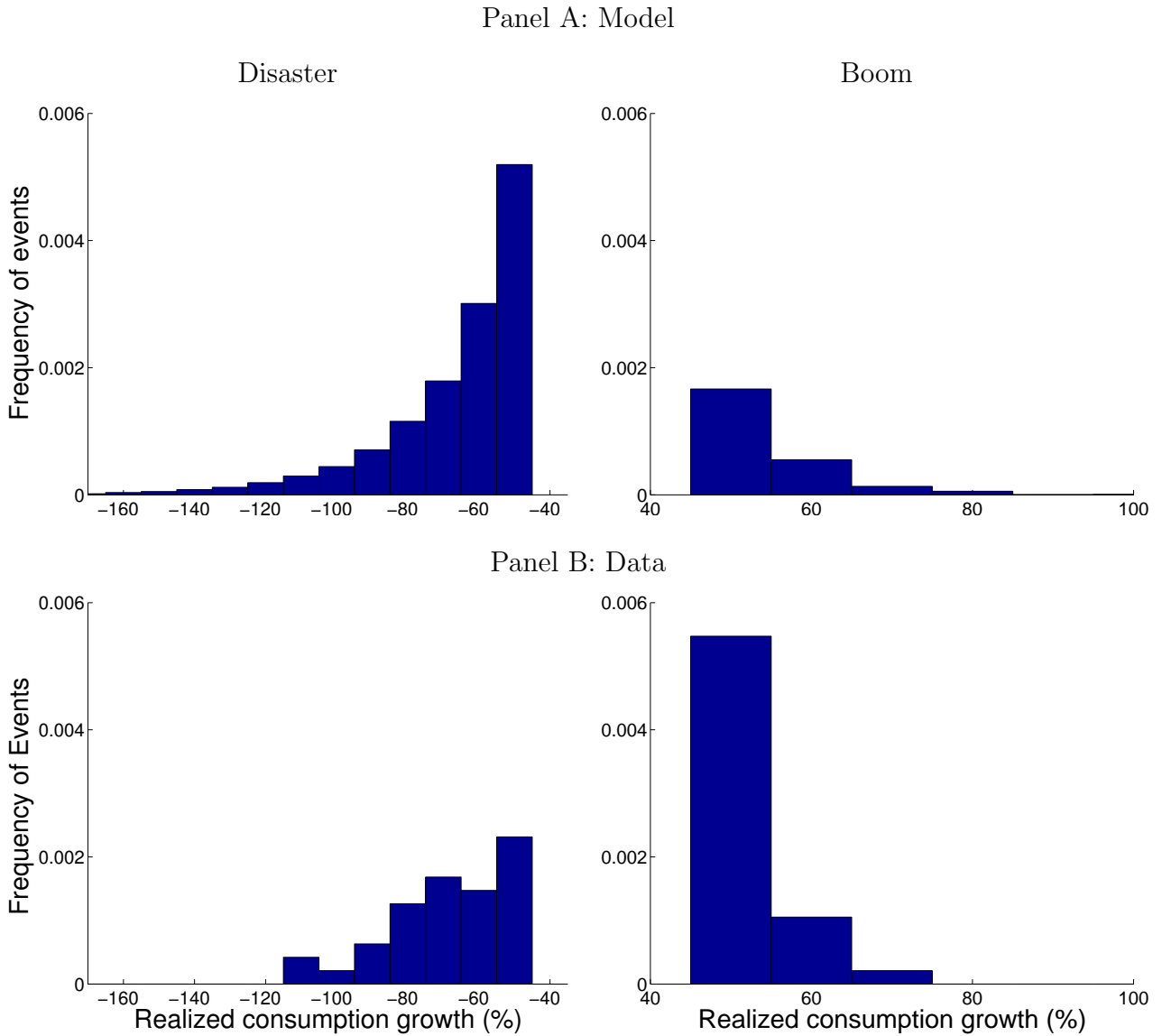
- Santos, Tano, and Pietro Veronesi, 2010, Habit formation, the cross section of stock returns and the cash-flow risk puzzle, *Journal of Financial Economics* 97, 385–413.
- Stambaugh, Robert F., 1999, Predictive regressions, *Journal of Financial Economics* 54, 375–421.
- Veronesi, Pietro, 2004, The Peso problem hypothesis and stock market returns, *Journal of Economic Dynamics and Control* 28, 707–725.
- Wachter, Jessica A., 2013, Can Time-Varying Risk of Rare Disasters Explain Aggregate Stock Market Volatility?, *The Journal of Finance* 68, 987–1035.
- Weil, Philippe, 1990, Nonexpected utility in macroeconomics, *Quarterly Journal of Economics* 105, 29–42.
- Zhang, Lu, 2005, The value premium, *Journal of Finance* 60, 67–103.

Figure 1: Tails of the one-year consumption growth rate distribution



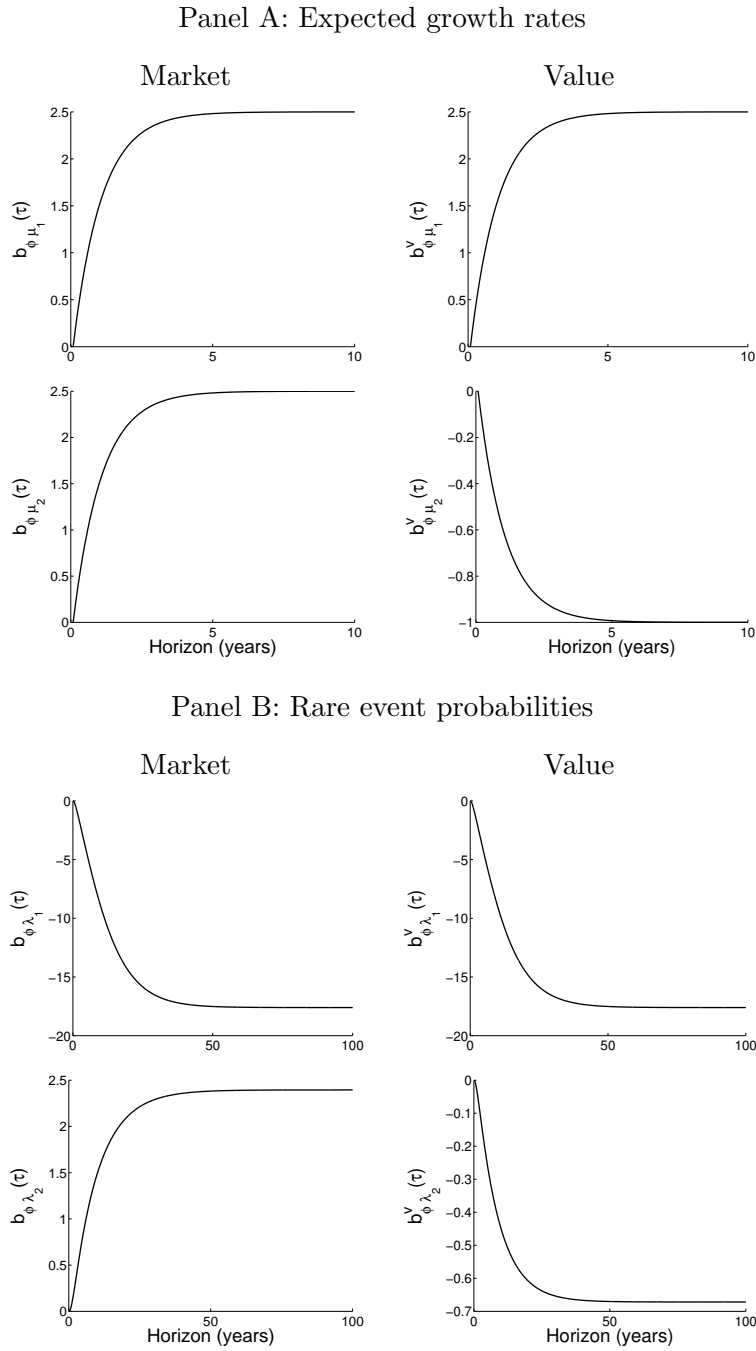
Note: This figure shows histograms of one-year consumption growth rates. The right panel considers growth rates above 15%. The left panel considers growth rates below -15%. The frequency is calculated by the number of observations within a range, divided by the total number of observations in the sample. 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: Tails of the five-year consumption growth rate distribution



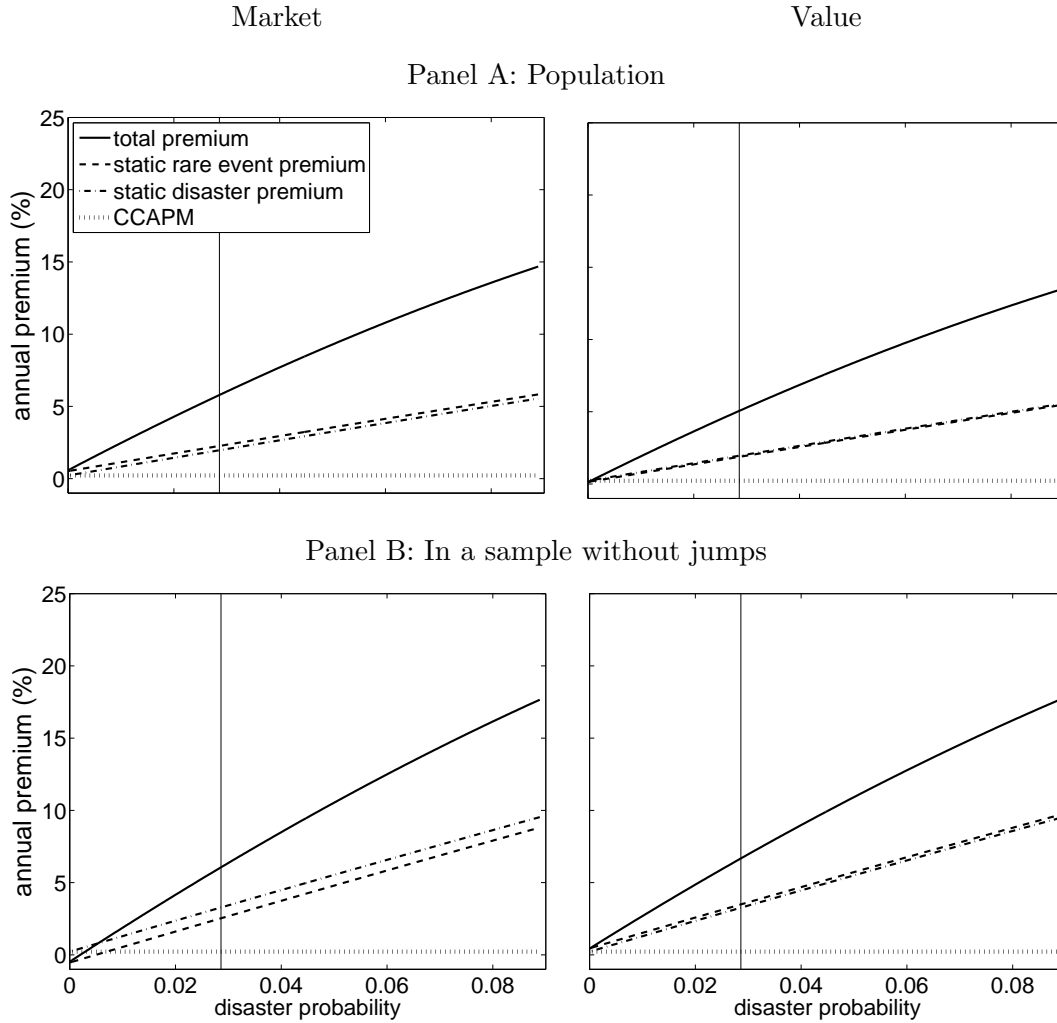
Notes: This figure shows histograms of five-year consumption growth rates. The right panel considers growth rates above 45%. The left panel considers growth rates below -45%. Panel A shows results from simulated data from the model. Panel B shows results from the data. Data are from Barro and Ursua (2008).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

Figure 3: Solution for the price-dividend ratio



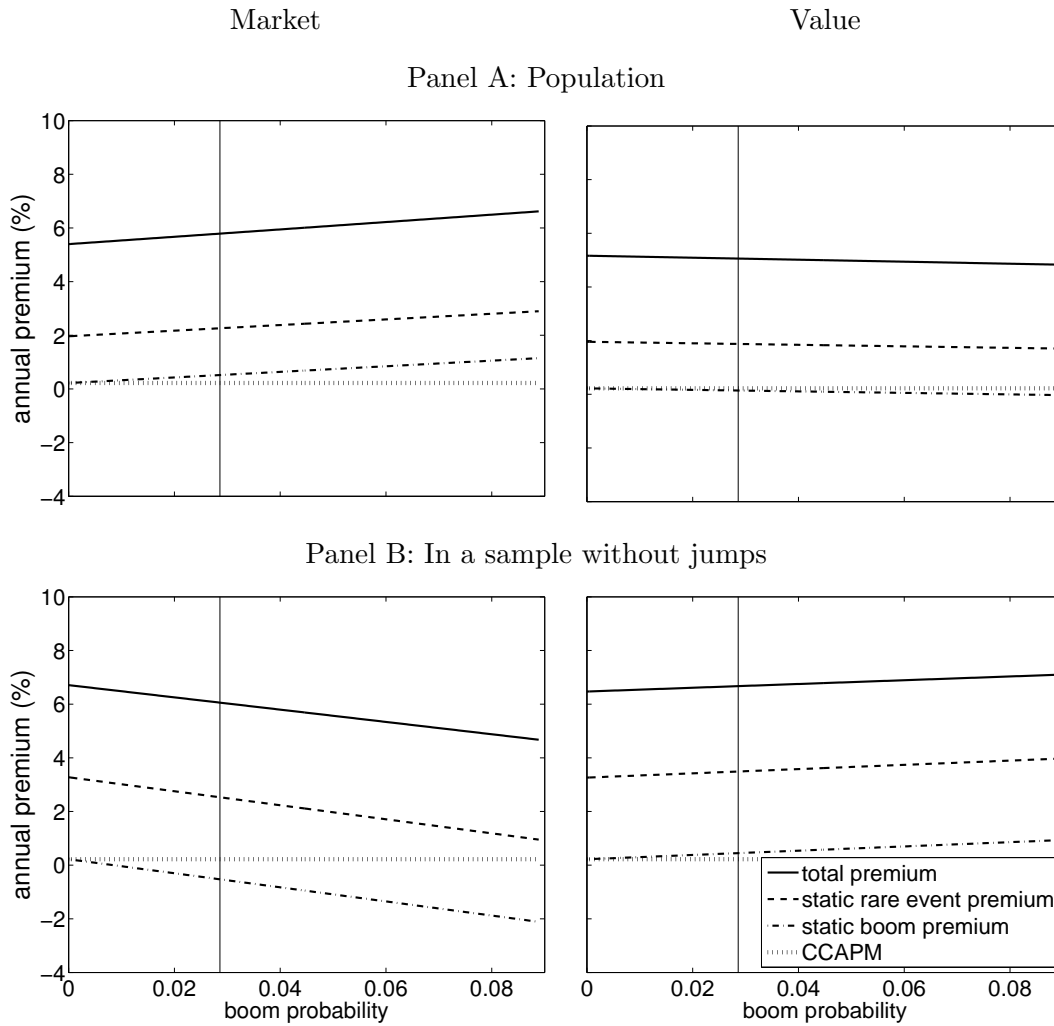
Notes: Panel A shows the coefficients multiplying the disaster component (μ_{1t}) and the boom component (μ_{2t}) of the expected growth rate in dividends. Panel B shows the coefficients multiplying λ_{1t} (the probability of a disaster) and λ_{2t} (the probability of a boom). The left panel shows results for the market; the right shows results for the value premium. The scales on the left and the right differ.

Figure 4: Risk premiums as a function of disaster probability



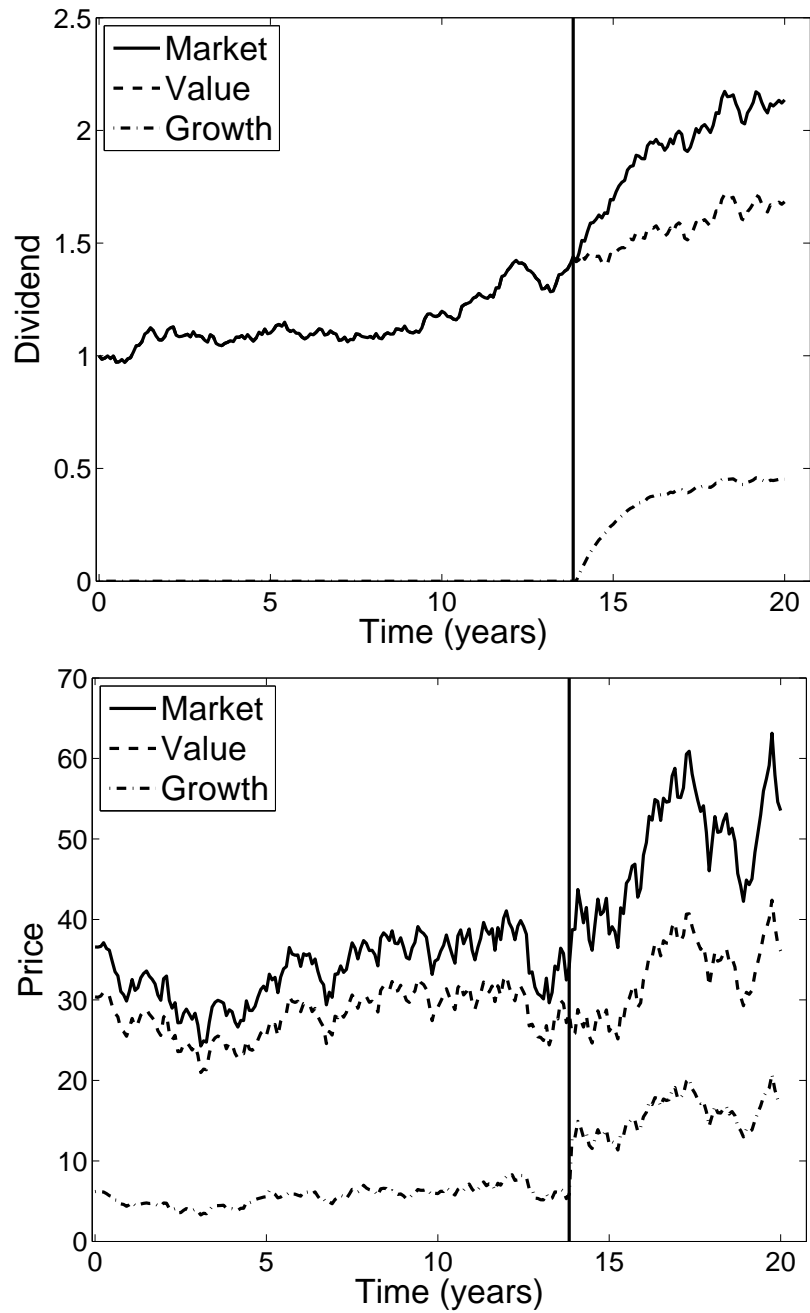
Notes: Panel A shows population risk premia as functions of the disaster probability with the boom probability fixed at its mean. The vertical line represents the mean of the disaster probability. Panel B shows the what the observed expected excess return would be in a sample without jumps.

Figure 5: Risk premiums as a function of boom probability



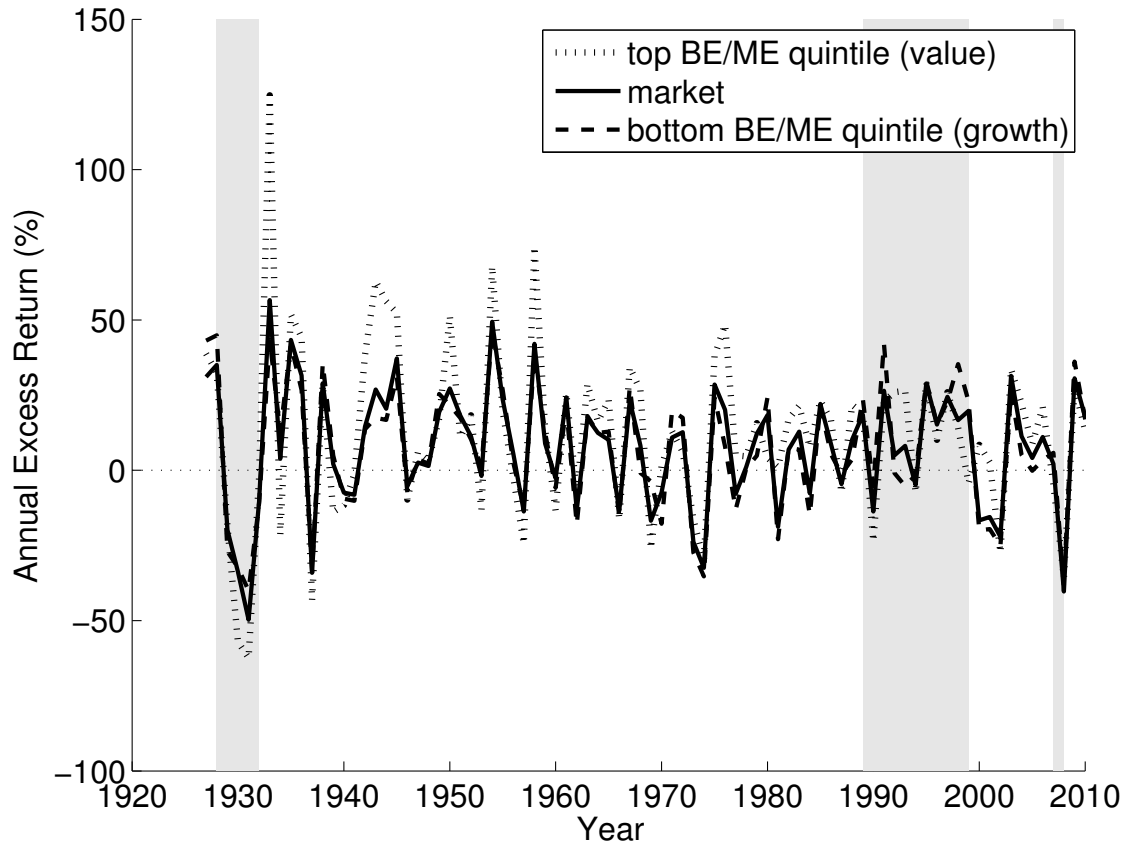
Notes: Panel A shows population risk premia as functions of the boom probability with the disaster probability fixed at its mean. The vertical line represents the mean of the boom probability. Panel B shows the what the observed expected excess return would be in a sample without jumps.

Figure 6: Sample simulation path of asset dividends and prices



Notes: This figure shows results from a time series simulated from the model that includes a boom. The top figure shows dividends (initialized at one), while the bottom panel shows prices.

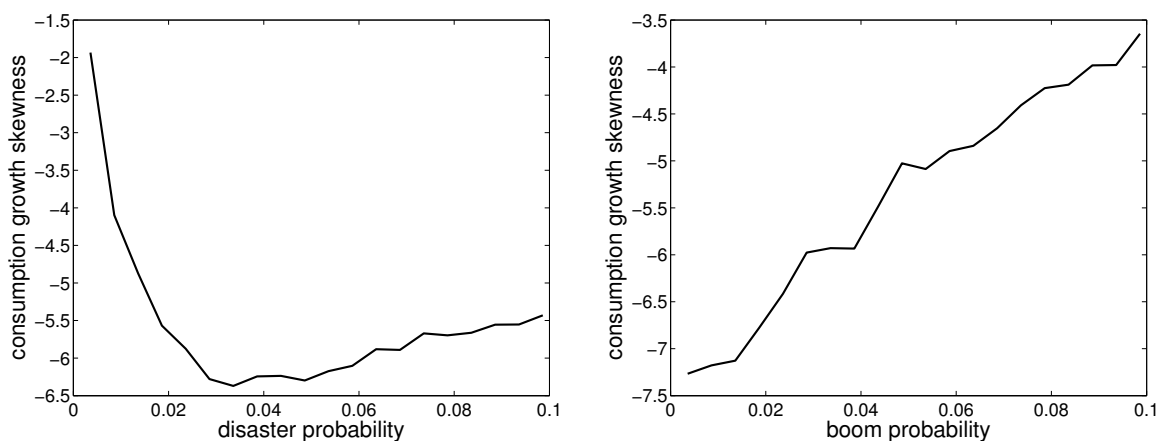
Figure 7: Value and growth returns



Notes: This figure plots the historical value and growth excess returns, along with the market excess return. The value sector is defined as the top book equity to market equity ratio quintile, the growth sector is the bottom book equity to market equity ratio quintile, and the market is the value-weighted market excess returns. All excess returns are measured relative to the 30-day Tbill. Data are annual from 1927 to 2010. The shaded areas are 1929 – 1932, 1990 – 2000, and 2008.

Figure 8: Consumption growth skewness

Panel A: Betas with respect to disaster probabilities



Notes: This figure shows the consumption growth skewness conditional on different rare event probability. The one on the left plots the skewness at different disaster probability λ_1 , with the boom probability equals to its mean. The one on the right plots the skewness at different boom probability λ_1 , with the disaster probability equals to its mean. We calculate it by simulating 500,000 years of consumption growth at a monthly frequency for each value of the disaster probabilities. We aggregate the monthly consumption growth to an annual frequency and calculate the skewness over the sample.

Table 1: Parameter values

Panel A: Basic parameters	
Average growth in consumption (normal times) $\bar{\mu}_C$ (%)	1.96
Average growth in dividend (normal times) $\bar{\mu}_D$ (%)	3.03
Volatility of consumption growth (normal times) σ (%)	1.45
Leverage ϕ	3.5
Rate of time preference β	0.003
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}	0.081
Mean reversion in expected consumption growth κ_{μ_1}	1.00
Minimum consumption disaster (%)	10
Power law parameter for consumption disaster	6.27
Panel C: Boom parameters	
Average probability of boom $\bar{\lambda}_2$ (%)	2.86
Mean reversion in boom probability κ_{λ_2}	0.11
Volatility parameter for booms σ_{λ_2}	0.081
Mean reversion in expected consumption growth κ_{μ_2}	1.00
Minimum consumption boom (%)	5
Power law parameter for consumption booms	15.00

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

Table 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
Smooth consumption	2.70	0.79	0.28	0.10	0.06
Jumps in consumption	0.81	0.78	0.46	0.28	0.44
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
Smooth consumption	0.83	0.52	0.31	0.20	0.33
Jumps in consumption	0.96	0.59	0.40	0.26	0.48

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. Smooth consumption refers to the model presented in the text, with jumps in expected consumption growth. Jumps in consumption refers to a model with jumps 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 3: Log consumption and dividend growth moments

Panel A: Consumption growth								
		No-Jump Simulations			All Simulations			
	Data	0.05	0.50	0.95	0.05	0.50	0.95	Population
mean	1.95	1.65	1.95	2.26	-0.31	1.65	2.70	1.50
standard deviation	1.45	1.22	1.44	1.66	1.47	3.16	7.52	4.24
skewness	-0.37	-0.50	0.00	0.48	-4.56	-1.63	2.21	-4.80
kurtosis	3.22	2.20	2.80	3.87	2.85	10.33	28.09	55.34

Panel B: Dividend growth								
		No-Jump Simulations			All Simulations			
	Data	0.05	0.50	0.95	0.05	0.50	0.95	Population
mean	1.67	1.84	2.91	3.98	-5.01	1.86	5.52	1.31
standard deviation	6.46	4.28	5.04	5.82	5.15	11.05	26.33	14.84
skewness	0.10	-0.50	0.00	0.48	-4.56	-1.63	2.21	-4.80
kurtosis	4.66	2.20	2.80	3.87	2.85	10.33	28.09	55.34

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 100,000 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 4: Value spread moments

	No-Jump Simulations				All Simulations			Population
	Data	0.05	0.50	0.95	0.05	0.50	0.95	
$\exp(E[\log(\text{value spread})])$	1.23	1.18	1.21	1.27	1.18	1.23	1.34	1.24
$\sigma(\log(\text{value spread}))$	0.08	0.02	0.04	0.07	0.03	0.07	0.16	0.10
Value spread autocorrelation	0.79	0.57	0.78	0.90	0.39	0.69	0.89	0.71

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 100,000 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 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	1.65	1.95	2.10	-0.12	1.62	2.57	1.49
$\sigma(R^b)$	2.75	0.13	0.25	0.50	0.38	2.43	5.76	3.24
$E[R^m - R^b]$	7.25	3.66	5.44	7.94	2.88	5.97	10.28	6.27
$\sigma(R^m)$	17.8	10.4	14.5	20.1	13.1	20.5	31.9	22.3
Sharpe ratio	0.41	0.27	0.38	0.51	0.14	0.30	0.46	0.28
$\exp(E[p - d])$	32.5	25.0	30.7	34.4	20.0	28.6	34.3	27.8
$\sigma(p - d)$	0.43	0.10	0.19	0.34	0.14	0.27	0.50	0.35
$\text{AR1}(p - d)$	0.92	0.57	0.78	0.91	0.53	0.78	0.91	0.85

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 100,000 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 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	4.34	6.06	8.52	2.59	5.26	8.28	5.34
$E[R^g - R^b]$	6.62	1.04	3.37	6.41	1.10	7.90	24.79	9.97
$E[R^v - R^g]$	1.34	1.26	2.74	4.03	-19.55	-2.42	3.46	-4.63
$\sigma(R^v)$	17.0	9.7	13.5	18.7	11.3	17.4	25.2	18.1
$\sigma(R^g)$	21.0	16.6	22.8	32.6	21.8	41.8	117.3	64.0
$\sigma(R^v - R^g)$	11.7	10.7	14.6	19.9	13.1	35.1	116.1	60.7
Sharpe ratio, value	0.48	0.34	0.45	0.60	0.13	0.31	0.51	0.29
Sharpe ratio, growth	0.32	0.05	0.15	0.24	0.04	0.18	0.29	0.16
Sharpe ratio, value-growth	0.11	0.07	0.19	0.33	-0.20	-0.07	0.25	-0.08
alpha, value	1.26	0.79	1.08	1.50	-0.22	0.86	2.82	1.23
alpha, growth	-1.26	-5.96	-4.32	-3.10	-9.62	-3.17	1.36	-4.28
alpha, value-growth	2.53	3.97	5.41	7.33	-1.53	4.07	12.29	5.51
beta, value	0.92	0.86	0.92	0.96	0.26	0.83	0.96	0.66
beta, growth	1.09	1.21	1.41	1.63	1.23	1.69	3.64	2.27
beta, value-growth	-0.16	-0.75	-0.49	-0.25	-3.32	-0.87	-0.28	-1.62

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 100,000 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 7: Long-horizon regressions of aggregate market and value-minus-growth returns on the price-dividend ratio and the value spread

	1-year Horizon				3-year Horizon				5-year Horizon			
	Data	NJ	All	Pop.	Data	NJ	All	Pop.	Data	NJ	All	Pop.
Panel A: Market returns on the price-dividend ratio												
Coef.	-0.12	-0.28	-0.17	-0.09	-0.29	-0.67	-0.42	-0.23	-0.41	-0.91	-0.60	-0.35
<i>t</i> -stat	[-2.41]				[-3.37]				[-3.37]			
R^2	0.09	0.15	0.07	0.02	0.22	0.33	0.16	0.05	0.27	0.43	0.22	0.07
Panel B: Market returns on the value spread												
Coef.	-0.50	-0.76	-0.33	-0.06	-1.18	-1.88	-0.86	-0.17	-1.28	-2.61	-1.26	-0.25
<i>t</i> -stat	[-1.86]				[-2.28]				[-3.13]			
R^2	0.05	0.05	0.02	0.001	0.12	0.12	0.06	0.002	0.09	0.16	0.08	0.003
Panel C: Value-minus-growth returns on the price-dividend ratio												
Coef.	0.01	0.11	0.06	0.01	0.05	0.27	0.16	0.02	0.09	0.38	0.24	0.02
<i>t</i> -stat	[0.37]				[0.51]				[0.76]			
R^2	0.00	0.03	0.02	1×10^{-4}	0.01	0.07	0.05	2×10^{-4}	0.02	0.10	0.07	2×10^{-4}
Panel D: Value-minus-growth returns on the value spread												
Coef.	0.46	1.18	0.47	0.01	1.13	2.87	1.23	0.03	1.48	3.95	1.81	0.05
<i>t</i> -stat	[2.52]				[2.44]				[2.37]			
R^2	0.10	0.10	0.03	1×10^{-5}	0.19	0.25	0.08	5×10^{-5}	0.21	0.34	0.12	7×10^{-5}

Notes: Coefficients and R^2 -statistics from predictive regressions in annual (overlapping) postwar data and in the model. In Panels A and B the excess market return is regressed against the price-dividend ratio and the value spread respectively. Panels C and D repeat this exercise with the value-minus-growth return. The value spread is defined as the log book-to-market ratio of the value sector minus log book-to-market ratio of the aggregate market. For the data coefficients, we report *t*-statistics constructed using Newey-West standard errors. Population moments (Pop.) are calculated from simulating monthly data from the model for 600,000 years and then aggregating to an annual frequency. We also simulate 100,000 60-year samples and report the average value for each statistic both from the full set of simulations (All) and for the subset of samples for which no jumps occur (NJ).

Table 8: Data on portfolios formed on the book-to-market ratio

	All stocks					Top size quintile				
	G					V				
	1	2	3	4	5	1	2	3	4	5
Panel A: 1947 – 2010										
$E[R^i - R^f]$	6.08	6.71	8.04	8.61	10.36	6.22	6.46	7.26	7.32	8.36
<i>se</i>	(0.59)	(0.55)	(0.52)	(0.54)	(0.62)	(0.57)	(0.54)	(0.52)	(0.55)	(0.62)
$\sigma[R^i - R^f]$	16.26	15.12	14.47	14.96	17.20	15.85	14.94	14.48	15.16	17.32
β_i	1.06	1.02	0.95	0.96	1.05	1.02	0.99	0.91	0.93	0.99
<i>se</i>	(0.02)	(0.01)	(0.02)	(0.02)	(0.03)	(0.02)	(0.01)	(0.02)	(0.02)	(0.03)
% of total market	40.84	22.24	17.03	12.22	7.66	33.85	16.86	12.37	8.01	4.33
Panel B: 1990 – 2000										
$E[R^i - R^f]$	11.23	10.48	9.72	10.02	10.57	12.23	10.95	9.54	9.51	10.41
<i>se</i>	(1.42)	(1.27)	(1.18)	(1.14)	(1.16)	(1.39)	(1.31)	(1.25)	(1.28)	(1.38)
$\sigma[R^i - R^f]$	16.26	14.62	13.58	13.15	13.38	16.02	15.01	14.40	14.73	15.82
β_i	1.11	1.09	0.99	0.92	0.89	1.09	1.11	1.00	0.95	0.93
<i>se</i>	(0.05)	(0.03)	(0.04)	(0.04)	(0.05)	(0.05)	(0.03)	(0.05)	(0.06)	(0.06)
% of total market	41.83	22.00	17.26	11.58	7.34	33.98	16.56	12.34	7.43	4.07
Panel C: 1929 – 1932										
$E[R^i - R^f]$	-21.79	-20.99	-28.65	-21.52	-26.35	-21.48	-19.26	-28.72	-29.79	-16.94
<i>se</i>	(5.77)	(5.71)	(7.14)	(8.82)	(9.12)	(5.77)	(5.70)	(7.47)	(8.79)	(8.74)
$\sigma[R^i - R^f]$	39.98	39.58	49.44	61.08	63.21	39.99	39.48	51.79	60.89	60.58
β_i	0.95	0.94	1.15	1.38	1.44	0.95	0.93	1.19	1.36	1.23
<i>se</i>	(0.04)	(0.03)	(0.06)	(0.13)	(0.12)	(0.04)	(0.03)	(0.07)	(0.11)	(0.17)
% of total market	46.62	24.56	18.22	8.40	2.20	42.72	20.30	14.50	6.01	0.89
Panel D: 2008										
$E[R^i - R^f]$	-43.42	-40.68	-49.46	-51.68	-38.88	-42.64	-38.93	-52.93	-56.76	-38.56
<i>se</i>	(6.51)	(5.24)	(5.81)	(6.53)	(7.77)	(6.21)	(4.64)	(5.31)	(6.45)	(7.71)
$\sigma[R^i - R^f]$	22.55	18.14	20.11	22.63	26.90	21.51	16.07	18.38	22.33	26.70
β_i	1.01	0.81	0.90	1.00	1.21	0.96	0.69	0.78	0.96	1.18
<i>se</i>	(0.07)	(0.04)	(0.06)	(0.06)	(0.06)	(0.07)	(0.05)	(0.09)	(0.08)	(0.09)
% of total market	36.14	25.71	16.40	13.79	7.96	29.40	20.31	12.22	10.03	4.99

Notes: Statistics for portfolio excess returns formed by sorting stocks by the ratio of book equity to market equity during various subperiods of the data. The panel reports means and β s with respect to the value-weighted CRSP market portfolio. Data are at a monthly frequency. We multiply excess returns by 1200 to obtain annual percent returns. Excess returns are measured relative to the 30-day Tbill. The left panel reports results from the full set of equities, while the right panel looks only at the top size quintile.

Table 9: Upside and downside betas

	$E[R - R^b]$ (%)	relative β^+	relative β^-
Value	6.06	-0.0024	0.0062
Growth	3.37	0.0686	-0.0634

Notes: This table reports the relative upside and downside betas for the value and growth portfolios in the model, along with the average returns on these portfolios. The first column reports the average annual excess returns for value and growth, the second row reports the relative upside betas, $\beta_i^+ - \beta_i$, where $\beta_i^+ = \frac{\text{cov}(R^i - R^m | R^m > \mu_m)}{\text{var}(R^m | R^m > \mu_m)}$, μ_m is the average return on the aggregate market within each simulation, and β_i is the regular CAPM beta for value and growth portfolio. The third column reports the relative downside betas, $\beta_i^- - \beta$, where $\beta_i^- = \frac{\text{cov}(R^i - R^m | R^m < \mu_m)}{\text{var}(R^m | R^m < \mu_m)}$. We simulate 100,000 60-year samples and report the median values from the subset of samples for which no jumps occur.