

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 Jonathan Berk, Adlai Fisher, Joao Gomes, Leonid Kogan, Nikolai Roussanov, Harald Uhlig and seminar participants at Boston University, McGill University, Stanford 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 They have not been peer-reviewed or been subject to the review by the NBER Board of Directors that accompanies official NBER publications.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, Revised April 2015
JEL No. G12

ABSTRACT

In this paper we exploit a fundamental difference between positive and negative rare events to explain the value premium. We show that if booms are expected but do not occur, average in-sample returns will be lower for assets that are exposed to booms than for those that are not. We build a general equilibrium endowment economy in which growth stocks are endogenously more exposed to booms than other assets. We show that this effect can account for most of the observed value premium, the low abnormal performance of growth stocks, and the high volatility and betas of growth stocks in the data.

Jerry Tsai
Department of Economics
University of Oxford
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 is subject to positive and negative events that are much larger than what would be expected under 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 have lower observed excess returns than samples in which disasters are relatively absent. However, these two types of samples are qualitatively similar in that they both exhibit a disaster premium. 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 have higher observed excess returns than samples in which booms

are relatively absent. However, these two types of samples are qualitatively different in that in one, there is a positive boom premium and in the other, the observed boom premium is negative. We use this simple theoretical observation to account for the value premium. In our model, the growth sector consists 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.

What is the source of the asymmetry between disasters and booms? Why, in other words, is the risk premium for bearing boom risk positive in population but negative in samples without rare events? Consider first the very simple problem faced by a risk-neutral investor, and assume (reasonably) that asset prices rise in booms and fall in disasters (we will show this in our model below). The risk neutral investor simply needs to be compensated for disaster risk in periods where disasters do not occur by higher returns. Similarly, he needs to be “compensated” for boom risk in periods where booms do not occur by *lower* returns. If returns were also higher when booms did not occur, then no-arbitrage would be violated.

Now consider the more realistic case of a risk averse investor. This investor requires a higher return to hold assets exposed to disaster and boom risk as compared with the risk-neutral investor. Because the disaster itself lowers prices, the positive returns due to disaster premia must come about at least in part during no-disaster periods. For disasters, this implies that both the population risk premium must be positive (because of risk aversion) and the observed expected return associated with disasters must be positive (both because of no-arbitrage, and because of the extra effect of risk aversion). Under a reasonable calibration, in fact, the population risk premium is an order of magnitude larger than the expected effect of the disaster on prices. This is why samples with disasters and those without look similar, as far as disaster risk premia go.

For booms, on the other hand, *the higher returns due to the risk premium come about when the boom itself is realized*. Thus the returns over periods without booms can be lower than for the asset without boom exposure. In fact, no-arbitrage implies that they must be

lower. Putting these facts together, we see that the population risk premium for booms is positive, while the observed expected return associated with booms is negative. Samples with and without booms look very different when it comes to boom premia: assets exposed to booms have higher true expected returns, but lower observed expected returns when booms do not occur.

This reasoning explains why one should expect a negative observed premium for boom risk in samples without rare events. However, it says nothing about the magnitude of the premium. To obtain quantitatively relevant results, a second mechanism is important. In our model, we make the standard assumption of constant relative risk aversion (CRRA), which leads to stationary rates of return. This standard assumption also implies that boom risk has a lower price than disaster risk. Because any risk premium for booms works against us, this second source of asymmetry between disasters and booms combines with the first to produce an economically significant value premium.

Specifically, when we calibrate our model, the observed equity premium is 5%. However, the observed return to growth stocks is only 3%, while the observed return to value stocks is 6%. Moreover, our model also explains why value stocks will have strong abnormal performance, and growth poor abnormal performance relative to the CAPM. The relative abnormal performance of value stocks implied by our model is 5%, as it is in the data. Indeed, our model naturally explains a number of aspects of the value and growth returns in the data that have posed a challenge to previous general equilibrium models. Namely:

1. Growth stocks have *higher* variance and covariance with the market despite having lower observed returns.
2. Growth stocks have yet higher covariances, and returns, during periods of high overall market returns, such as occurred in the 1990s.
3. The value-minus-growth return, unlike the market excess return, cannot be predicted by the price-dividend ratio. It can however be predicted by the value spread.

Moreover, while a full explanation of skewness puzzles is outside the scope of this study, our model does imply that high valuation stocks have high skewness and low expected returns, as in the data. Our model also implies that assets with high “upside” betas and low “downside” betas also have low excess returns, as in the data.

In explaining these facts, we tie our hands in that, despite the presence of disasters, value stocks are not more exposed to disaster risk (while plausible and in the spirit of our model, we assume this way to focus on the main mechanism of booms). Moreover, we assume that booms affect consumption as well as dividends; this implies boom risk is priced, and this works against us in finding a value premium.

Finally, because of the presence of disasters in the model, we can explain a high equity premium and equity volatility, along with low volatility of consumption growth. The model achieves this with a risk aversion coefficient as low as three. Note that low risk aversion is crucial to explaining the value puzzle in our setting; if risk aversion were too high, growth would carry a higher premium in population, and we would not be able to match low observed returns over the sample. The model generates realistic volatility through the mechanism of time-varying disaster and boom risk, combined with recursive preferences. Without this mechanism, equity claims would have counterfactually low volatility during normal times.

To summarize: thus far the literature has focused on one-sided rare events, namely disasters, to explain the equity premium. We show however, that the presence of booms has a large affect on the cross-section if some assets are exposed to them and some are not. That is, by introducing booms as well as disasters, one can explain not only the equity premium puzzle but the value puzzle as well.

Relation to the prior literature

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), Ai, Croce, and Li

(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 these models do not explain the observed pattern in postwar variances and covariances. 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

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. McQuade (2013) shows that stochastic volatility in production can generate a value premium, depending on how the risk of volatility is priced.

(2013) endogenously generate a cross-section of firms through differences in investment opportunities, but, like Lettau and Wachter, they assume an exogenous stochastic discount 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 (endogenously) small. Nonetheless, our model has a reasonable equity premium.

Our model features rare disasters, as do models of 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. 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 $Z_{1t} < 0$ and the second type ($j = 2$) to be booms, so that $Z_{2t} > 0$. When a disaster occurs, the process μ_{1t} jumps downward. It then reverts back to zero (absent any other bad shocks). Likewise, when a boom occurs, the process μ_{2t} jumps upward. It too reverts back. Absent any disaster or boom shocks, the drift rate of consumption is $\bar{\mu}_C$. This model allows for consumption to adjust smoothly as in the data (see Nakamura, Steinsson, Barro, and Ursúa (2013)), but yet to undergo periods of extreme growth rates in either direction.

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 given by λ_{jt} , which is stochastic, and follows the process

$$d\lambda_{jt} = \kappa_{\lambda_j} (\bar{\lambda}_j - \lambda_{jt}) dt + \sigma_{\lambda_j} \sqrt{\lambda_{jt}} dB_{\lambda_{jt}}, \quad j = 1, 2, \quad (4)$$

where the $B_{\lambda_{jt}}$ 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. Given our parameter choices, λ_{jt} is an excellent approximation of a probability of a disaster ($j = 1$) or of a boom ($j = 2$) over, say, a year, and we use the terms “probability” and “intensity” interchangeably. 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 $\beta > 0$. These assumptions lead to exact expressions that are available in closed-form up to ordinary differential equations. A detailed derivation is in Appendix B.⁴

2.2 The state-price density

We start by establishing how the various sources of risk are priced in the economy. We will use the notation $\mathcal{J}_j(\cdot)$ to denote how a process changes in response to a rare event of type

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

⁴Using 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.

j . For example, for the state-price density π_t , $\mathcal{J}_j(\pi_t) = \pi_t - \pi_{t-}$ if a type- j jump occurs at time t . In our complete-markets endowment economy, the state-price density represents the marginal utility of the representative agent.

Theorem 1. *The state-price density π_t follows the process*

$$\frac{d\pi_t}{\pi_{t-}} = \mu_{\pi_t} dt + \sigma_{\pi_t} dB_t + \sum_{j=1,2} \frac{\mathcal{J}_j(\pi_t)}{\pi_{t-}} dN_{jt}, \quad (7)$$

where

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

and

$$\frac{\mathcal{J}_j(\pi_t)}{\pi_t} = e^{b_{\mu_j} Z_{jt}} - 1, \quad \text{for } j = 1, 2, \quad (9)$$

and

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

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

Moreover, for $\gamma > 1$, $b_{\lambda_1} > 0$, $b_{\lambda_2} < 0$, and $b_{\mu_j} < 0$ for $j = 1, 2$.

Theorem 1 shows that, in the event of a disaster, marginal utility jumps upward, and in the event of a boom, marginal utility jumps downward. The exponential form of state prices derives from CRRA utility. Properties of the exponential function imply that the upward jump in response to a disaster is greater than the downward jump in response to a boom, even when the distributions for Z_{1t} and Z_{2t} are symmetric. Changes in λ_{jt} are also priced as reflected by the second and third elements of σ_{π_t} . Marginal utility rises when the risk of a disaster rises, and falls when the risk of a boom rises. Again, there is an asymmetry. The price for time-varying disaster risk, b_{λ_1} , is much larger in magnitude than the price of time-varying boom risk, b_{λ_2} , even when the distributions are symmetric.

Because the EIS is equal to 1, and because only expected consumption (not realized consumption) is subject to jumps, the riskfree rate in this economy is standard.

Corollary 2. *Let r_t denote the instantaneous risk-free rate in this economy. Then*

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

Table 1 summarizes the effect of shocks on prices and returns in the model. Note that only μ_{1t} and μ_{2t} effect the riskfree rate, not the rare event probabilities.

2.3 The aggregate market

Here, we derive formulas for the price-dividend ratio and the equity premium on the aggregate market. These formulas explicitly show how booms and disasters and their probabilities impact these quantities, and thus are important for understanding the qualitative and quantitative results for value and growth claims.

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

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

where

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

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

Valuation

Our first result gives the formula for the price of the aggregate market. By no-arbitrage,

$$F(D_t, \mu_t, \lambda_t) = E_t \int_t^\infty \frac{\pi_s}{\pi_t} D_s ds,$$

where π_s is the state-price density. Valuing the market comes down to calculating the expectation on the left hand side.

Theorem 3. 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

$$F(D_t, \mu_t, \lambda_t) = \int_0^\infty H(D_t, \mu_t, \lambda_t, \tau) d\tau. \quad (14)$$

where

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

$$b_{\phi\mu_j}(\tau) = \frac{\phi - 1}{\kappa_{\mu_j}} (1 - e^{-\kappa_{\mu_j}\tau}), \quad j = 1, 2 \quad (16)$$

and the remaining terms satisfy

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

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

with boundary conditions $b_{\phi\lambda_j}(0) = a_\phi(0) = 0$. Furthermore, the price-dividend ratio on the market portfolio is given by

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

Here and in what follows, $b_{\phi\mu}(\tau) = [b_{\phi\mu_1}(\tau), b_{\phi\mu_2}(\tau)]^\top$ and $b_{\phi\lambda}(\tau) = [b_{\phi\lambda_1}(\tau), b_{\phi\lambda_2}(\tau)]^\top$.

Equation 14 expresses the value of the aggregate market in terms of the sum (or, more precisely, the integral) of zero-coupon equity claims. H gives the values of these claims as functions of the disaster and boom terms μ_{1t}, μ_{2t} , the probabilities of a disaster and boom $\lambda_{1t}, \lambda_{2t}$ and the time τ until the dividend is paid.

These individual dividend prices (and, by extension the price of the market as a whole) have interpretations based on the primitive parameters. Because $\phi > 1$, (16) shows that prices are increasing in μ_{1t} and μ_{2t} . The presence of $\phi - 1$ shows that this is a trade-off between the effect of expected consumption growth on future cash flows and on the riskfree rate. In our model with $EIS = 1$ and $\phi > 1$, the cash flow effect dominates and the valuation of the market falls during disasters and rises during booms. Moreover, the more persistent

is the process (the lower is κ_{μ_j}), the greater is the effect of a change in μ_{jt} on prices.⁵ These formulas also show that the probability of rare events also affect prices, but the intuition is more subtle. We discuss it further below. First, we note the sign of the coefficients in the following corollary (see also the row corresponding to the market price-dividend ratio in Table 1):

Corollary 4. *The price-dividend ratio $G(\mu_t, \lambda_t)$ is increasing in μ_{jt} (for $j = 1, 2$), decreasing in λ_{1t} and increasing in λ_{2t} .*

The proof of this corollary is in Appendix B.4, but substantial intuition can be obtained from the formulas themselves. The functions $b_{\phi\mu_j}(\tau)$ are discussed above; now consider the functions $b_{\phi\lambda_j}(\tau)$. These would be identically zero without the last term in the ODE, $E_{\nu_j} \left[e^{b_{\mu_j} Z_{jt}} \left(e^{b_{\phi\mu_j}(\tau) Z_{jt}} - 1 \right) \right]$. 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.⁶ 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,⁷ representing the combined effect of changes in expected cash flows and in the riskfree rate. Thus the response of equity values to a change in the disaster

⁵The 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 $e^{\kappa_{\mu_j} \tau} > \kappa_{\mu_j} \tau + 1$.

⁶More precisely, this is the static disaster premium for zero-coupon equity with maturity τ .

⁷Again, more precisely, it is the price response of zero-coupon equity with maturity τ .

probability is determined by a risk premium effect, and a (joint) cash flow and riskfree rate effect. For disasters, both effects operate in the same direction, namely an increase in the probability lowers valuations. For booms, on the other hand, the risk premium effect lowers prices but the joint cash flow and riskfree rate effect raises prices. The left hand side of (20) is positive (because $e^{b_{\phi\mu_2}(\tau)Z_{2t}} > 1$), and thus the joint cash flow and riskfree rate effect dominates.

The left panel in Figure 1 shows the coefficients as functions of the maturity of zero-coupon equity τ .⁸ As discussed above, $b_{\phi\mu_1}(\tau)$ and $b_{\phi\mu_2}(\tau)$ are positive, reflecting the fact that the market is exposed to both disasters and booms. Both terms converge to their limits in a relatively short time because in our calibration (and in the data) rare events occur on the order of years rather than decades. Likewise, the figure shows that $b_{\phi\lambda_1}(\tau)$ is negative; an increase in the disaster probability lowers prices. Changes in the disaster probability are more persistent in our calibration than disasters themselves (this persistence is necessary to match the persistence of price-dividend ratios in the data), and this is reflected in the longer time to convergence. An increase in the boom probability raises prices, as shown by the fact that $b_{\phi\lambda_2}(\tau)$ is positive; like $b_{\phi\lambda_1}(\tau)$ it takes a long time to converge because of the persistence in the boom probability. Also interesting is the fact that $b_{\phi\lambda_2}(\tau)$ is so much smaller than $b_{\phi\lambda_1}(\tau)$.⁹ This occurs because the cash-flow/riskfree rate effect and the risk premium effect operate in the same direction for disasters but in opposite directions for booms.

The equity premium

We now turn to the equity premium. For our quantitative results, we will examine the average simulated excess return, calculated over a time horizon that matches the data.

⁸These figures are created based on the calibration discussed later in the paper. Here, we focus on qualitative properties of the figures, which are robust to many alternative calibrations.

⁹While we assume in our calibration that booms are somewhat smaller than disasters, this is a generic result that holds even when booms and disasters are the same size.

However, we gain intuition by examining the instantaneous equity premium which is an excellent approximation to its finite-horizon counterpart, and is available in closed form.

First note that by Ito's Lemmas, we can write the price process for the aggregate market as follows:

$$\frac{dF_t}{F_{t-}} = \mu_{F_t} dt + \sigma_{F_t} dB_t + \sum_{j=1,2} \frac{\mathcal{J}_j(F_t)}{F_{t-}} dN_{jt},$$

for some drift term μ_{F_t} , a (row) vector of diffusion terms σ_{F_t} , and terms $\frac{\mathcal{J}_j(F_t)}{F_{t-}}$ that denote percent change in price due to the rare event. Furthermore, note that $\frac{\mathcal{J}_j(F_t)}{F_{t-}} = \frac{\mathcal{J}_j(G_t)}{G_{t-}}$, because dividends themselves do not jump in this model; only the price-dividend ratio does.

The instantaneous expected return is defined as the expected percent price appreciation, plus the dividend yield. In the notation above,

$$r_t^m = \underbrace{\mu_{F_t} + \frac{1}{F_t} \sum_{j=1,2} \lambda_{jt} E_{\nu_j} [\mathcal{J}_t(F_t)]}_{\text{expected price appreciation}} + \underbrace{\frac{D_t}{F_t}}_{\text{dividend yield}}. \quad (21)$$

Theorem 5. *The instantaneous equity premium relative to the risk-free rate r_t is*

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

Theorem 5 divides the equity premium into three components. The first is the standard term arising from the consumption CAPM (Breenden (1979)). The second component is the sum of the premium directly attributable to disasters ($j = 1$) and to booms ($j = 2$). These are covariances between the state-price density and the aggregate market value during rare events, multiplied by probabilities that the rare events occur.¹⁰ We call this second term the *static rare event premium* because it is there regardless of whether the probabilities of rare events are constant or time-varying.¹¹

¹⁰These terms take the form of uncentered second moments, but they are indeed covariances; this is because the jump occurs instantaneously and so the conditional expected change in the variable is negligible.

¹¹However, the term "static premium" is somewhat of a misnomer here, since even the direct effect of

The third component in (22) represents the compensation the investor requires for bearing the risk of changes in the rare event probabilities. 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). The following corollary shows that all terms in (22) are positive. A closely related result is that the increases in both the disaster and the boom probability increase the equity premium, as indicated in Table 1 and discussed in what follows.¹²

Corollary 6. *1. The static disaster and boom premiums are positive.*

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

Proof For Result 1, recall that $b_{\mu_j} < 0$ for $j = 1, 2$ (Theorem 1). First consider disasters ($j = 1$). Note $Z_{1t} < 0$, so $e^{b_{\mu_1} Z_{1t}} - 1 > 0$. Furthermore, because G_t is increasing in μ_{1t} (Corollary 4), $\mathcal{J}_1(G_t) < 0$. It follows that the static disaster premium is positive. Now consider booms ($j = 2$). Because $Z_{2t} > 0$, $e^{b_{\mu_2} Z_{2t}} - 1 < 0$. Because G_t is increasing in μ_{2t} , $\mathcal{J}_2(G_t) > 0$. Therefore the static boom premium is also positive.

To show the second statement, first consider disasters ($j = 1$). Recall that $b_{\lambda_1} > 0$ (Theorem 1). Further, $\partial G / \partial \lambda_1 < 0$ (Corollary 4). For booms ($j = 2$), each of these quantities takes the opposite sign. The result follows. \square

In words, the static premiums are positive because state prices and aggregate market values move in opposite directions during rare events: during disasters, marginal utility is high,

rare events on the price-dividend ratio is a dynamic one. In a model with time-additive utility, only the instantaneous co-movement with consumption would matter for risk premia, not changes to the consumption distribution. Thus there would only be the CCAPM term under our assumptions.

¹²Table 1 indicates that there is no effect of μ_{1t} and μ_{2t} on risk premia. In fact, there is a second-order effect. The reason is that changes to μ_{1t} and μ_{2t} result in changes in duration of the equity claim, and this impacts sensitivity to λ_{1t} and λ_{2t} , in addition to disasters and booms themselves. This effect is negligible in our calibration.

but valuations are low while during booms the opposite is true. Thus exposure to disasters and booms has a direct positive impact on the equity premium.

Exposure to disasters and booms also increases the equity premium indirectly through the dynamic effect of time-varying probabilities (the resulting term is what we refer to as the λ -premium). An increase in disaster risk raises marginal utility and lowers valuations, likewise an increase in boom risk lowers marginal utility and raises market valuations. Thus exposure to time-varying probabilities of rare events further increases the equity premium.

Figure 2 (top right panel) shows these terms as a function of the disaster probability for the calibration discussed later in the paper. The dotted line that is essentially at zero shows the CCAPM. The dash-dotted line shows the static disaster premium; lying above it is the full static premium, that includes the premium due to booms. Finally, the solid line is the full equity premium, which includes the λ -premium. Under a calibration designed to match equity volatilities, the λ -premium due to disasters is substantial as in Wachter (2013). The λ -premium due to booms, however, is extremely small, as we discuss further in the next section.¹³

Observed returns in samples without rare events

We now 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

$$r_{nj,t}^m = \mu_{Ft} + \frac{D_t}{F_t}. \quad (23)$$

¹³This figure shows that the static boom premium is small. This is not a generic result; it arises in our calibration because booms are smaller than disasters. While booms have a smaller effect on marginal utility and thus on state prices, they have a larger effect on asset prices because of Jensen’s inequality. If the leverage parameter ϕ and risk aversion are equal, and booms and disasters are symmetric, then these terms would be of the same size. On the other hand, the λ -premium due to booms is smaller than for disasters, even under these conditions.

Based on this definition, on Theorem 5, and on $\mathcal{J}_j(F_t)/F_t = \mathcal{J}_j(G_t)/G_t$, these expected returns can be calculated as follows:

Corollary 7. *The observed expected excess return in a sample without jumps is*

$$r_{\text{nj},t}^m - r_t = \phi\gamma\sigma^2 - \underbrace{\sum_{j=1,2} \lambda_{jt} E_{\nu_j} \left[e^{b_{\mu_j} Z_{jt}} \frac{\mathcal{J}_j(G_t)}{G_t} \right]}_{\text{observed static rare-event premium}} - \underbrace{\sum_{j=1,2} \lambda_{jt} \frac{1}{G_t} \frac{\partial G}{\partial \lambda_j} b_{\lambda_j} \sigma_{\lambda_j}^2}_{\lambda\text{-premium}}. \quad (24)$$

As in the true risk premium, there are components of the observed premium attributable to disasters ($j = 1$) and to booms ($j = 2$). The premium for time-varying λ risk takes the same form for both the observed and true premium cases, because rare-event risk varies whether a rare event occurs or not. It is the static premium, or the premium due to the rare event itself, that differs.¹⁴ It follows immediately from (24) (and indeed, it can be inferred from the definition (23)), that the observed static premium for disasters is higher than the true static premium, while the observed static premium for booms is lower. In fact, for booms it will be sufficiently lower so that the observed static premium is negative, as stated in the corollary below.

Corollary 8. *The observed static disaster premium in a sample without jumps is positive. The observed static boom premium in a sample without jumps is negative.*

Proof The result follows from (24), from $e^{b_{\mu_j} Z_{jt}} > 0$ and from $\mathcal{J}_1(G_t) < 0$ and $\mathcal{J}_2(G_t) > 0$ (because prices are increasing in μ_{jt}). \square

We now return to a question raised the introduction: why does the premium due to booms switch signs? The intuition given there is reflected in the very simple proof of Corollary 8. First, the relevant component of state prices $e^{b_{\mu_j} Z_{jt}}$ is positive, regardless

¹⁴We 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.

of parameter values (this is a reflection of the fact that absence of arbitrage holds in the model). Second, during booms, asset prices rise. The observed (static) boom premium is equal to the negative of the percent change in asset prices multiplied by the relevant component of the state price. In other words, the observed premium due to booms must be negative to compensate for the positive returns when booms are realized; otherwise no-arbitrage would be violated. This effect is mitigated by risk aversion. The greater is γ , the closer to zero the observed premium due to booms can be.¹⁵ Of course, as shown in Corollary 6, the true premium due to booms arises from the covariance of state prices with asset prices and must be positive. This risk premium, and more, is realized by the investor when the boom actually occurs.

We can see the difference between booms and disasters by contrasting the left panels in Figure 2 with those of Figure 3. Figure 2 shows risk premia as a function of disaster probability; Figure 3 shows risk premia as a function of boom probability, and hence better highlights the role of booms. In Figure 2 there is very little difference between true risk premia and observed risk premia in samples without rare events. In Figure 3, true and observed risk premia are qualitatively different. The true boom premium is positive, and decreasing in the disaster probability, while the observed boom premium is negative, and decreasing in the probability.

Before leaving the section on risk premia, we note the importance of the asymmetry in the price of boom versus disaster risk. The discussion above pertains to just the static part of the premium, not the λ -premium.¹⁶ If the λ -premium for booms is large enough it could

¹⁵More precisely, what matters $\gamma - 1$, or in fact the difference between γ and the inverse of the EIS. The reason is that the rare events change the consumption distribution rather than consumption itself. The relevant notion of risk neutrality is thus time-additive utility, in which the agent is indifferent over the timing of the resolution of uncertainty. In this case, $\gamma = 1$ and indeed the static rare-events premia reduce to price changes.

¹⁶This raises the question of why we bother making the rare event probabilities vary at all, since the main effect in the model does not depend on these terms. The answer is given in the introduction: if the rare event probabilities were constant, than equity volatility would be essentially zero except when

eliminate the in-sample observed negative returns for booms. However, the λ -premium for booms, unlike that for disasters is extremely small. One important reason for this is that the price of risk for λ_{2t} is small in magnitude compared to the price of risk for λ_{1t} ; that is $b_{\lambda_1} > -b_{\lambda_2}$, on account of constant relative risk aversion. Second, changes in the probability of boom have a smaller affect on prices, as explained in the discussion following Theorem 3.

To summarize, this section shows that, while the true premia for both disaster and boom risk are positive, the observed premium for disaster risk is positive while the observed premium for boom risk is negative. These results are directly relevant for the cross-section because, as we will see, value and growth claims differ based on their exposure to the types of risks discussed here.

2.4 Growth and value sectors

In the previous section, we define the market portfolio and its properties. Based on these results, we can see the differential affects of booms and disasters on prices and on risk premia. These results will be useful as we now turn to assets that differ in their exposure to these sources of uncertainty.

The value sector

We first define a sector of the economy that is not exposed to the risks (and benefits) of booms. It makes intuitive sense that firms and industries will differ in their ability to directly profit from technological progress. For clarity, we define a stark dichotomy. We create a sector that does not directly benefit from a boom, but is otherwise identical to the market. A second sector, one that captures the benefits of the boom, is simply defined as everything else that is in the market portfolio.¹⁷

disasters were to take place.

¹⁷In reality, what we think of as value and growth stocks could be combinations of the value and growth claim, with growth stocks loading more on the boom sector. For reasons that we explain below, such

Consider an asset with cash flows following the process

$$\frac{dD_t^v}{D_t^v} = \mu_{Dt}^v dt + \phi \sigma dB_{Ct}, \quad (25)$$

where $\mu_{Dt}^v = \bar{\mu}_D + \phi \mu_{1t}$. This asset is the same as the aggregate market, except it is not exposed to positive jumps. We use the subscript v to denote “value”. As we will show, this asset will have a higher ratio of fundamentals-to-price than the market as a whole. This is the defining characteristic of value in the data.¹⁸

Corollary 9. *The price-dividend ratio for value is below that of the market.*

While a formal proof is in Appendix B.5, the reasoning is straightforward. The value claim is not exposed to booms, and these are unambiguously positive events. Thus dividend growth for the market is weakly greater than dividend growth for value in every state of the world, and strictly greater in some states of the world. It follows that the price-dividend ratio, which is the present discounted value of future dividends scaled by current dividends, is lower for value than for the market.

Pricing for the value claim is directly analogous to that of the market. Because this price takes a form so similar to that of the market (Theorem 3), that we briefly summarize the results and leave the details to Appendix B.5. The price of the claim to the dividend stocks would have higher stock market valuations relative to fundamentals, consistent with the definition of growth in the data. Writing down such share dynamics would greatly complicate the model without adding to the intuition. Note that this structure implies that observed dividend growth is only higher for the growth sector if a rare boom actually occurs. Thus our model is consistent with the results of Chen (2012), who finds relatively small differences in the measured growth rate on growth stocks as compared to value stocks.

¹⁸Note that in this model, dividends, earnings and cash flows are all the same. While it is most traditional to define value firms as those with a high ratio of book-to-market value, using other measures of fundamentals than book value, such as dividends, earnings, and cash flows work equally well. That is, the same “value puzzle” holds if firms are sorted on these alternate measures, and in fact all the measures are highly correlated (see, e.g. Lettau and Wachter (2007)).

stream (25) is given by

$$F^v(D_t^v, \mu_t, \lambda_t) = \int_0^\infty H^v(D_t^v, \mu_t, \lambda_t, \tau) d\tau, \quad (26)$$

where

$$H^v(D_t^v, \mu_t, \lambda_t, \tau) = D_t^v \exp \{ a_\phi^v(\tau) + b_{\phi\mu}^v(\tau)^\top \mu_t + b_{\phi\lambda}^v(\tau)^\top \lambda_t \}. \quad (27)$$

The price-dividend ratio for the value claim is therefore

$$G^v(\mu_t, \lambda_t) = \int_0^\infty \exp \{ a_\phi^v(\tau) + b_{\phi\mu}^v(\tau)^\top \mu_t + b_{\phi\lambda}^v(\tau)^\top \lambda_t \} d\tau, \quad (28)$$

with $a_\phi^v(\tau)$, $b_{\phi\mu}^v(\tau)$, and $b_{\phi\lambda}^v(\tau)$ given in Appendix B.5. We highlight one important difference between these terms and their counterparts for the market portfolio. Note that

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

The analogous term for the market is $b_{\phi\mu_2}(\tau) = \frac{\phi-1}{\kappa_{\mu_2}} (1 - e^{-\kappa_{\mu_2}\tau})$. From (29), we see that the price of the value claim fluctuates with booms, even though the cash flow process does not itself depend on booms. The reason is that, when a boom occurs, the riskfree rate rises because the representative agent has a greater desire to borrow against the future. This causes asset prices to fall. This effect is present for the aggregate market, but it is dominated by the expected cash flow effect. For the value claim, this is the only effect booms have on prices.

Naturally, the difference in $b_{\phi\mu_2}^v(\tau)$ carries over to $b_{\phi\lambda_2}^v(\tau)$, which reflects how the price responds to changes in the probability of a boom. An increase in the probability of a boom decreases the price of the value claim because the riskfree rate effect dominates everything else. We summarize these results as follows:

Corollary 10. *The price-dividend ratio for the value claim $G^v(\mu_t, \lambda_t)$ is increasing in μ_{1t} , decreasing in μ_{2t} , and decreasing in the probability of a rare event λ_{jt} , for $j = 1, 2$.*

Figure 1 compares the coefficients on value with those for the market. We see that the response of the value claim to disasters and to changes in the disaster probability are

almost indistinguishable. The response for booms is quite different. The function $b_{\phi\mu_2}^v(\tau)$ is negative and decreasing in τ , rather than positive and increasing as it is for the market. It is also about half the magnitude of the market coefficient, because the riskfree rate effect alone is small compared with the (combined) cash flow and riskfree effect for the market. We see this also when considering the response of the price of the value premium to changes in the boom probability. Again, $b_{\phi\lambda_2}^v(\tau)$ is negative and decreasing, and small in magnitude when compared with the corresponding function for the market.

These results lead directly to formulas for the risk premium on the value claim. See Appendix B.5 for details. The proof of Corollary 12 is similar to that for the market and so it is omitted.

Corollary 11. 1. *The value sector premium relative to the risk-free rate r_t is*

$$r_t^v - r_t = \phi\gamma\sigma^2 - \sum_{j=1,2} \lambda_{jt} E_{\nu_j} \left[(e^{b_{\mu_j} Z_{jt}} - 1) \frac{\mathcal{J}_j(G_t^v)}{G_t^v} \right] - \sum_{j=1,2} \lambda_{jt} \frac{1}{G_t^v} \frac{\partial G^v}{\partial \lambda_j} b_{\lambda_j} \sigma_{\lambda_j}^2. \quad (30)$$

2. *The observed expected excess return on the value sector in a sample without rare events is*

$$r_{nj,t}^v - r_t = \phi\gamma\sigma^2 - \sum_{j=1,2} \lambda_{jt} E_{\nu_1} \left[e^{b_{\mu_j} Z_{jt}} \frac{\mathcal{J}(G_t^v)}{G_t^v} \right] - \sum_{j=1,2} \lambda_{jt} \frac{1}{G_t^v} \frac{\partial G^v}{\partial \lambda_j} b_{\lambda_j} \sigma_{\lambda_j}^2. \quad (31)$$

Corollary 12. 1. *The static boom premium for the value sector is negative (it is positive for the market).*

2. *The λ_2 -premium is negative (it is positive for the market).*

3. *The observed static boom premium for the value sector is positive (it is negative for the market).*

Other components of the risk premium and observed risk premium on value take the same sign as the market.

Note that the price of the value sector responds in the opposite direction to the market in the case of booms; this is the reason for the change in signs.

We show the components of the value sector premium next to the market as a function of disaster probability (Figure 2) and as a function of boom probability (Figure 3). The difference is most apparent when we consider risk premiums as a function of the boom probability. For the market portfolio, the static observed boom premium is negative in samples without rare events. For the value sector, the static observed boom premium is slightly positive. This reflects the intuition in the introduction: when investors are expecting booms and they do not occur, the observed returns on assets exposed to booms will be lower than on assets not exposed to booms (or assets that fall in price when booms occur). We can see this directly by comparing the left and right figures in Panel B of Figure 3.

The growth sector

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,

$$F_t^g = F_t - F_t^v. \quad (32)$$

Figure 4 shows dividends (Panel A) and prices (Panel B) for a typical simulation that contains a boom. As Panel A shows, prior to a boom, the dividend on the value claim and the market are identical, and the dividend on the growth claim equals zero. When a boom occurs, a wedge opens up between the market dividend and the value dividend. This wedge is the dividend on the growth claim.

Figure 4, Panel B shows the price of a claim to the value sector, the market, and the growth sector. Note that even though the growth sector pays no dividends prior to the boom, it has a non-zero price because investors anticipate the possibility of future dividends. When a boom occurs, the price of the growth claim immediately rises, the value of the aggregate market also rises, but by less, and the price of the value claim falls slightly. After the boom, the price of the growth claim and the overall market remain high relative to value, reflecting permanently higher dividend growth. However, all claims do rise in

value over time because aggregate dividends are growing.

We can use the basic accounting identity (32), and the analogous identity for dividends, to derive an expression for the dividend-price ratio on the growth claim. We use this expression to show that the growth claim has a lower dividend-price ratio than the value claim.

Corollary 13. *The dividend-price ratio for the growth sector is below the dividend-price ratio for the value sector.*

Proof. It follows from the definition of the growth dividend and from the accounting identity (32) that

$$\begin{aligned} \frac{D_t^g}{F_t^g} &= \frac{D_t - D_t^v}{F_t - F_t^v} \\ &= \frac{D_t - D_t^v}{D_t G_t - D_t^v G_t^v} \\ &= \frac{1}{G_t^v} \frac{D_t - D_t^v}{D_t \frac{G_t}{G_t^v} - D_t^v}. \end{aligned} \tag{33}$$

Note that $1/G_t^v$ is the dividend-price ratio on the value claim. Because value has a lower price-dividend ratio than the market as a whole (Corollary 9), $\frac{G_t}{G_t^v} > 1$. Because the dividend on the market is at least as big as the dividend on value, the term multiplying $1/G_t^v$ is less than one. \square

Not surprisingly, the value of the growth sector is an increasing function of expected dividend growth during a boom as well as the probability of a boom. Less obvious is the fact that growth is also exposed to the risk of a disaster. Prior to a boom, growth has no cash flows to fall in the case of disaster. However, after a boom takes place, the cash flows that accrue to growth fall by the same percentage amount in the event of disaster as the rest of the dividends in the economy. Anticipating this, investors price the effect of a disaster into growth stocks before the boom occurs. A full proof is given in Appendix B.5.

Corollary 14. *The price of the growth sector is increasing in μ_1 , decreasing in λ_1 , and increasing in μ_2 and λ_2 .*

The most important aspect of this structure is that the growth sector is essentially a levered claim to the proceeds of the boom. Relative to the market as a whole, it has a low value. However, it bears the entire risk of the boom.

Is this model non-stationary?

A concern one might have with the model thus far is non-stationarity in the relative size of the value sector. It might seem that the value sector grows ever smaller as a proportion of the market as a whole and the growth claim eventually equals the entire market. However, in the data, the size of the value sector does not appear to be trending downward. The prior literature on the cross-section has taken different views toward the issue of stationarity. Some authors (who do not emphasize the value premium) view a non-stationary cross-section is acceptable (Cochrane, Longstaff, and Santa-Clara (2008), Martin (2013)), while others explicitly model share dynamics so a stationary cross-section is achieved (Lettau and Wachter (2007), Santos and Veronesi (2010)).

In fact, our model makes it possible to define the value sector such that its relative size is stationary. At each time t , we redefine the value sector as the asset that pays current dividend equal to the dividend on the market (namely D_t), and with future dividend stream evolving according to (25). That is,

$$\frac{dD_s^v}{D_s^v} = \mu_{D_s^v}^v dt + \phi \sigma dB_{C_s}, \quad s \geq t \tag{34}$$

with boundary condition $D_t^v = D_t$. The price of this value sector is still the no-arbitrage value of the dividend stream (34), and so is equal to (26) with $D_t^v = D_t$ in the first argument.¹⁹

This definition of the value sector is consistent over time, just as it is possible to have, say, a company change its mix of debt and equity in the capital structure without violating the promises made to either stake holder. Our model is one in which there are two claims,

¹⁹Note that Figure 4 shows the time path of prices without redefining value's dividends; it is therefore what cash flows and price appreciation look like from the point of view of the owner of each of the claims.

“value” and “growth” that add up to the market as a whole. In the event of a large boom, both claims are dissolved and two new claims are created, with the owners of the previous claims receiving the correct number shares of the new claims. After a boom, some of the capital from the previous growth sector goes to fund the new growth sector, and some goes into the value sector. The owners of the previous value sector own less of the new one, but the value of their holdings stays the same. For example a large technology company that is no longer innovative might use its current capital to fund riskier start-ups that will be the new “growth” firms; the remainder of the company could be thought of as a value firm. We are not claiming that what happens in our model is exactly what happens in the data, but rather pointing out that it is not necessary, either in the model or in the data, for the value of a stream of cash flows to be identified with a single firm or sector over time.

Note that price-dividend ratios and instantaneous expected returns are invariant to whichever definition we use, because they depend only on the stationary state variables μ_t and λ_t . Returns defined over a finite unit of time are also invariant, as we show in Appendix C. Thus the only difference between a model where value is “redefined” whenever a boom takes place and one where it is not is that, in the former case value’s share both in terms of market valuation and dividends remains stationary whereas in the latter it will eventually fall to zero.

3 Quantitative results

3.1 Data

Our data come from various sources. We will compare our rare events in the model to tail events in the data, using international consumption data described in detail in Barro and Ursúa (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 use the log book-to-market ratio on the market minus the log book-to-market ratio on value 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 2 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 Ursúa (2008) estimate that the probability of a rare disaster in OECD countries is 2.86%.²⁰ We use this number as our average disaster

²⁰We 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.

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 6.27. 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.²¹ Following Barro and Ursúa (2008), we assume a 10% minimum disaster size.

The power law distribution for booms is quite difficult to observe directly. We could use international data on consumption growth, and in fact such data provide plenty of evidence of extreme positive growth rates. However, one could reasonably ask whether these data are directly applicable to a developed country like the United States. We thus turn to asset markets, and in particular, to the size of the growth sector. The size of the growth sector in the model turns out to be sensitive to the thickness of the tail of the power law distribution, a thicker tail implying a larger growth sector. We can therefore infer tail thickness by matching the size of the growth sector in the model to the size of the growth sector in the data. In following this strategy, we follow David and Veronesi (2013), who also use asset markets to determine beliefs about states that rarely occur.

Specifically, our growth sector is in fact an extreme example of a growth stock, making the lowest book-to-market decile a reasonable target for the size of the growth sector. To identify the minimum jump size, we choose an annual growth rate of 5%, a value that would be an unusual normal-times observation under our parameter values, and yet is conservative. Given this minimum jump size, we require the model to match the relative

²¹One 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 extremely limited in this regard.

book-to-market ratio of value (deciles 2–9) as compared with the market as a whole.²² Given our other parameter choices, this implies a power law parameter of 15, corresponding to a thinner tail than for disasters. As we later discuss, it turns out that the value premium is quite insensitive to the choice of this parameter.

Despite the fact that we use asset market data to infer our distribution for booms, we still want to compare these booms to those we see in international data. The disaster distribution and the boom distribution in the model and in the data are reported in Figures 5 and 6. These figures show that, except for small disasters at the five-year horizon, our assumptions imply less extreme distributions than the data.²³ In particular, our model implies fewer, and smaller, booms than observed in the international data.

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 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. We discuss this in detail in Appendix A, but it is also apparent from the equation for b_{λ_j} in state prices (see Theorem 1), that, for a solution to exist, parameters must satisfy

$$\frac{1}{2\sigma_{\lambda_1}^2} (\kappa_{\lambda_1} + \beta)^2 \geq E_{\nu_1} [e^{b_{\mu_1} Z_1} - 1]. \quad (35)$$

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

²²In the model, the “book” values of the market and of value are the same. Requiring the model to match relative market valuations produces very similar answers.

²³It 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.

least, this is not the case.²⁴ 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. Such a model would not be able to account for the observed levels of stock market volatility.²⁵

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. This would give a high number of 4.7.²⁶ 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

²⁴Chen, Dou, and Kogan (2013) note this lack of identification as a disadvantage of rare events models.

²⁵Why, intuitively, is there such a constraint? Note that utility is a solution to a recursive equation; the above discussion reflects the fact that there is no guarantee in general that a solution to this recursion exists. In this particular case, it appears that the problematic region of the parameter space is one in which there is a lot of uncertainty that is resolved very slowly. A sufficiently slow resolution of uncertainty could lead to infinitely negative utility for our recursive utility agent.

²⁶However, dividend and consumption growth are not perfectly correlated in the data, suggesting that a lower number is more reasonable.

value of 5.1). Given that high values of ϕ are helpful for the moments of equity returns, our minimization chooses this upper bound.²⁷

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.²⁸ 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 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

²⁷We have simulated from a calibration 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, not surprisingly, because it is the risk of rare events, rather than the normal-times consumption risk, that drives our results. Lowering ϕ itself does lead to somewhat lower observed equity and value premia, but the difference is not large. A ϕ of 3 implies an equity premium of 5.1% (as compared with 5.4 in our main calibration) and an observed value premium of 2.6% (as compared with 2.7 in our main calibration).

²⁸Attempting to match the very high persistence of the price-dividend ratio leads to unstable results.

that the post-war period is exceptional in that a boom has not been observed. Moreover, while increasing the probability of a boom does increase the observed value premium, it also reduces the observed equity premium. Thus requiring the model to match the equity premium as well as the value premium is an additional source of discipline.²⁹

3.3 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.³⁰ Given a simulated series of the state variables, we obtain price-dividend ratios and one-period dividend growth rates on the market and on value. Using these quantities we simulate returns as described in Appendix C. 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 rare events. It is this subset of simulations that is the most interesting comparison for postwar data.

²⁹A 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.

³⁰The 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)).

3.3.1 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.³¹ 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.

While other papers show that disasters can explain the equity premium, our paper departs from most of the literature in that we put the disaster in expected rather than realized consumption. This implies that the disaster unfolds over a realistic time interval rather than nearly instantaneously. Because we still find an equity premium, our paper constitutes a response to Constantinides (2008) and Julliard and Ghosh (2012), as well as Mehra and Prescott (1988), who critique the near-instantaneous feature of traditional disaster models.

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). This effect would be straightforward to implement but would substantially

³¹In the definition of kurtosis that we use, three is the value for the normal distribution.

complicate the notation and exposition without changing any of the underlying economics. Moreover, 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.³² 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 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. Some behavior of asset 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.

3.3.2 Unconditional moments of value and growth portfolios

Table 6 reports cross-sectional moments in the model. As a “tight” data comparison, we take the growth portfolio as the bottom decile formed by sorting on book-to-market and the value portfolio as the remaining nine deciles. This comparison has the advantage that, in both the model and in the data, the two portfolios considered sum to the market. However, we also report excess returns for more traditional measures of value and growth in Table 8. The top panel shows excess returns, standard deviations, and betas for the full sample period. As this table shows, the value premium, as measured by comparing the top to the bottom quintile of stocks is 4.28%.

Table 6 shows that our model can account for a value spread of 2.74%, a substantial

³²The 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.

fraction of the data value, even though the definition of “value” in our model arguably accounts for more stocks. The 95% critical value from simulations of our model is 4.03%, not far from the data value of 4.28%. These values correspond to simulations without rare events. The population value premium is in fact negative, because growth stocks are in fact more risky than value stocks in the model. Yet even looking across the full set of simulations implies that it is not unlikely to observe a value premium in any particular sample.

Table 6 also shows that value stocks have lower standard deviations than growth stocks and higher Sharpe ratios. Both of these results hold across the full set of simulations, as well as in the samples without rare events. Both of these affects are strongly present in the data. Clearly the value premium in the model does not represent a return for risk; in the model it arises, as explained in Section 2, because investors are willing to accept a lower return on growth in most periods, in return for an occasional very high payout.³³

The fact that the value premium is not a return for bearing risk leads to a surprising comparative static: the value premium is quite insensitive to the size of booms in the economy. Figure 7 shows the average value spread across simulations, where we vary parameters of the boom distribution. In Panel A, we vary the probability of the boom and in Panel B we vary the size of the tail parameter for the power law distribution. The value premium is indeed increasing in the probability of a boom. If there are no rare booms in the model, then the value premium predicted by the model will be zero.³⁴ Conversely, high probabilities of a boom are associated with high value premiums, though for probabilities that are high enough, it becomes quite unlikely that we would have observed a sample containing no booms. However, once one has specified a heavy right tail for consumption growth, the precise form of this tail matters remarkably little. Specifically, Panel B shows

³³A 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.

³⁴We examine the sensitivity across a range from 0.6% probability to 5%. Below this 0.6%, the growth sector is extremely small and return moments are unstable.

that the value premium barely changes at all as a function of the power law tail parameter. Lower values of this parameter are associated with thicker tails. At extremely low values (equal to 1 or 2), dividend growth is high enough that the aggregate market price, and hence the growth sector price fail to converge (see Assumption 2 in Appendix A). Starting with a thickness parameter of 3 and going arbitrarily high, it appears that there is no change in the observed expected return on value stocks relative to growth stocks.

Why is it that the observed value premium is so insensitive to the shape of the boom distribution? The reason is that two opposing forces determine the observed value premium, as explained in the Section 2 and in the introduction. On the one hand, the greater the probability of large booms, the riskier growth stocks become, and the more negative is the true value premium in population. On the other hand, the greater the probability of large booms, the lower is the return on growth stocks a risk neutral investor is willing to accept in samples without booms. These two effects roughly cancel, leaving the actual observed value premium to be very robust to many different assumptions on the distribution for booms.

Table 6 also shows that our model can explain the relative alphas and betas for value and growth stocks. Growth stocks have a high covariance with the market, because they are a levered bet on the occurrence of booms. The market return also depends on innovations to the boom probability, and so changes to this probability cause these two assets to move together. Moreover, when booms actually occur, both the market as a whole and growth stocks especially rise in value. Given that growth stocks have higher betas and lower observed returns than value stocks, it is of course not surprising that they have negative alphas. In fact, they have negative alphas in population as well as in samples without rare events, because a large part of their risk comes from changes to the probability of a boom, and the premium associated with this risk is low. Thus, unlike previous models of the value premium, our model is able to explain the patterns in betas on growth and value in the data. It does so in a way that is consistent with the patterns in expected returns.

3.3.3 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.³⁵ 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.

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 rare events than 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

³⁵Lettau and Wachter (2011) show that if a single factor drives risk premia, then population values of predictive coefficients should be proportional across asset classes.

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.³⁶ The present model does, however, predict a positive coefficient. A high value of the price-dividend ratio on the market indicates a relatively high probability of a boom. In samples without rare events, the return on growth will be lower than the return on value when the boom probability is high. In the population, the coefficient is negative (and quite small); times of high λ_2 precede periods of high returns on growth when jumps occur with their proper frequency.³⁷

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

³⁶Roussanov (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.

³⁷ 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.

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.³⁸

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

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 will be periods when growth will outperform value, namely when 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 quintiles on the basis of the book-to-market ratio. We examine results for the full CRSP universe, as well as for the top size quintile, 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

³⁸In 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.

the lowest book-to-market quintile. Not only does growth 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 a 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 and in 2008, there appears to be little relation between performance and book-to-market quintile. 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 is not reflected in the performance of value stocks during these periods remains an open question.

4.2 Downside and upside risk

Ang, Chen, and Xing (2006) define downside β to be the covariance divided 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 stocks with higher downside β s have higher returns. Lettau, Maggiori, and Weber (2014) show that this finding also holds 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 that is rarely observed. In contrast, the studies mentioned above focus 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 and normal-times variation. An asset that is more exposed to disaster risk will also tend to fall during market downturns, because a downturn 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 link between the volatility of the rare event probabilities and their magnitudes. Moreover, while changes in boom probabilities in general have a smaller effect on the market return than disaster probabilities, upturns will be disproportionately caused by changes in the boom probability.

Table 9 shows the relative upside and downside betas calculated for the value and growth sectors of the economy, along with the observed risk premia. Indeed, the table shows that low relative upside betas and high relative downside betas are associated with high observed returns, just as in the data.

4.3 Skewness in the time series and cross-section

Our model predicts that positive skewness should be associated with low future 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 (2014), 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 rare events) go in the opposite 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 small samples without rare events. 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 requires little 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. Finally, we abstract from differential exposure to disaster risk. 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.

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

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. 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.³⁹ Assumption 3 states that the asymptotic slope of $a_{\phi}(\tau)$ is negative. This is the dynamic analogue of the condition that the growth rate be less than the discount rate in the static Gordon growth model. This

³⁹Note 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.

condition pertains to the market portfolio. If it is satisfied, the analogous condition for the value claim is satisfied automatically.⁴⁰

B Proofs of Theorems

This Appendix contains a detailed solution of the model. Section B.1 describes notation. Section B.2 describes the derivation of the value function and thus the state-price density under Assumption 1. Section B.3 contains general results on pricing equities in a model with jumps. This section does not require parametric assumptions on the processes for dividends or state prices. Sections B.4 and B.5 give results under the parametric assumptions in the main text; for these two sections, we make Assumptions 1–3.

B.1 Notation

- Definition of jump notation $\mathcal{J}(\cdot)$:

Let X_t be a pure diffusion process, and let μ_{jt} , $j = 1, 2$ be defined as in (3). Consider a scalar 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

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

for $j = 1, 2$, and

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

⁴⁰Specifically, 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 $\zeta_{\phi_2}^v > \zeta_{\phi_2}$. It is also the case that $\zeta_{\phi_1}^v = \zeta_{\phi_1}$.

- We use the notation κ_μ to denote the column vector $[\kappa_{\mu_1}, \kappa_{\mu_2}]^\top$, and similarly for κ_λ , σ_λ and $\bar{\lambda}$. Recall that we have already defined $\lambda_t = [\lambda_{1t}, \lambda_{2t}]^\top$, $\mu_t = [\mu_{1t}, \mu_{2t}]^\top$, $B_{\lambda t} = [B_{\lambda_{1t}}, B_{\lambda_{2t}}]^\top$ and $B_t = [B_{Ct}, B_{\lambda t}^\top]^\top$.
- We 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 $[\sigma_{\lambda_1}^2, \sigma_{\lambda_2}^2]^\top$.
- We use the notation $*$ to denote element-by-element multiplication for vectors of the same dimensionality.
- Partial derivatives with respect to a vector will be assumed to be row vectors. That is, given a function $h(\mu)$, $\partial h / \partial \mu = [\partial h / \partial \mu_1, \partial h / \partial \mu_2]$, and similarly for λ .
- Because the processes λ are independent, a consequence of Ito's Lemma is that cross-partial derivatives do not enter into the pricing equations. Thus, given a function $h(\lambda)$, we use the notation $\partial^2 h / \partial \lambda^2$ to denote the row vector $[\partial^2 h / \partial \lambda_1^2, \partial^2 h / \partial \lambda_2^2]$.

B.2 The state-price density

As a first step to computing the state-price density, we compute continuation value V_t as a function of aggregate wealth W_t and the state variables. As is often the case in models involving square root processes, there are technically two possible solutions corresponding to the choice of a solution to a quadratic equation. We select the solution such that, when there is no rare event risk, utility is equivalent to the case when there are in fact no rare events. Further discussion of this selection procedure is contained in Wachter (2013).

Lemma B.1. *In equilibrium, continuation value $V_t = J(W_t, \mu_t, \lambda_t)$, where W_t is the wealth of the representative agent, and J is given as follows:*

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

where

$$I(\mu_t, \lambda_t) = \exp \{ a + b_\mu^\top \mu_t + b_\lambda^\top \lambda_t \}, \quad (\text{B.2})$$

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:

$$\begin{aligned} a &= \frac{1-\gamma}{\beta} \left(\bar{\mu}_C - \frac{1}{2}\gamma\sigma^2 \right) + (1-\gamma) \log \beta + \frac{1}{\beta} b_\lambda^\top (\kappa_\lambda * \bar{\lambda}) \\ b_{\mu_j} &= \frac{1-\gamma}{\kappa_{\mu_j} + \beta}, \\ b_{\lambda_j} &= \frac{1}{\sigma_{\lambda_j}^2} \left(\beta + \kappa_{\lambda_j} - \sqrt{(\beta + \kappa_{\lambda_j})^2 - 2E_{\nu_j} [e^{b_{\mu_j} Z_{jt}} - 1] \sigma_{\lambda_j}^2} \right). \end{aligned}$$

Furthermore,

$$\frac{W_t}{C_t} = \beta^{-1}. \quad (\text{B.3})$$

where C_t is aggregate consumption.

Proof Let S_t 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 l . This relation implies that S_t satisfies

$$\frac{dS_t}{S_t} = \frac{dC_t}{C_t} = \mu_{C_t} dt + \sigma dB_{C_t}. \quad (\text{B.4})$$

Consider an agent who allocates wealth between the claim to aggregate consumption and the risk-free asset. Let α_t be the fraction of wealth in the consumption claim, 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 denotes the instantaneous risk-free rate. Optimal consumption and portfolio choice must satisfy the following Hamilton-Jacobi-Bellman (HJB) equation:

$$\begin{aligned} \sup_{\alpha_t, c_t} \left\{ \frac{\partial J}{\partial W} (W_t \alpha_t (\mu_{C_t} - r_t + l^{-1}) + W_t r_t - c_t) + \frac{\partial J}{\partial \lambda} (\kappa_\lambda * (\bar{\lambda} - \lambda_t)) - \frac{\partial J}{\partial \mu} (\kappa_\mu * \mu_t) \right. \\ \left. + \frac{1}{2} \frac{\partial^2 J}{\partial W^2} W_t^2 \alpha_t^2 \sigma^2 + \frac{1}{2} \left(\frac{\partial^2 J}{\partial \lambda^2} \right) (\sigma_\lambda^2 * \lambda_t) + \lambda_t^\top \bar{J} (J(W_t, \mu_t, \lambda_t)) + f(c_t, J_t) \right\} = 0. \quad (\text{B.5}) \end{aligned}$$

In equilibrium, $\alpha_t = 1$ and $c_t = C_t = W_t l^{-1}$. Substituting these policy functions into (B.5) implies

$$\begin{aligned} \frac{\partial J}{\partial W} W_t \mu_{C_t} + \frac{\partial J}{\partial \lambda} (\kappa_\lambda * (\bar{\lambda} - \lambda_t)) - \frac{\partial J}{\partial \mu} (\kappa_\mu * \mu_t) + \frac{1}{2} \frac{\partial^2 J}{\partial W^2} W_t^2 \sigma^2 \\ + \frac{1}{2} \left(\frac{\partial^2 J}{\partial \lambda^2} \right) (\sigma_\lambda^2 * \lambda_t) + \lambda_t^\top \bar{\mathcal{J}}(J(W_t, \mu_t, \lambda_t)) + f(C_t, J_t) = 0. \end{aligned} \quad (\text{B.6})$$

From the envelope condition $\partial f / \partial C = \partial J / \partial W$, we obtain $\beta = l^{-1}$, and prove (B.3). Given that the consumption-wealth ratio equals β^{-1} , it follows that

$$\begin{aligned} f(C_t, V_t) &= f(W_t l^{-1}, J(W_t, \mu_t, \lambda_t)) \\ &= \beta W_t^{1-\gamma} I(\mu_t, \lambda_t) \left(\log \beta - \frac{\log I(\mu_t, \lambda_t)}{1-\gamma} \right), \end{aligned} \quad (\text{B.7})$$

where, to derive (B.7), we conjecture (B.2). Substituting (B.7) and (B.1) into (B.6), we find

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

Note that $\mu_{C_t} = \bar{\mu}_C + \mu_{1t} + \mu_{2t}$. Collecting coefficients on μ_{jt} results in linear equations for b_{μ_j} for $j = 1, 2$. Solving these equations yields

$$b_{\mu_j} = \frac{1-\gamma}{\kappa_{\mu_j} + \beta}, \quad j = 1, 2.$$

Collecting coefficients on λ_{jt} yields a quadratic equation for b_{λ_j} for $j = 1, 2$. Given the root selection procedure described at the start of this Appendix, the solutions are

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

Collecting the constant terms implies

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

This verifies our conjecture on the form of I . □

Lemma B.2. *State prices can be characterized as follows:*

$$\pi_t = \exp \left\{ \int_0^t \frac{\partial}{\partial V} f(C_s, V_s) ds \right\} \beta^\gamma C_t^{-\gamma} I(\mu_t, \lambda_t) \quad (\text{B.8})$$

for $I(\mu_t, \lambda_t)$ given in Lemma B.1.

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

$$\pi_t = \exp \left\{ \int_0^t \frac{\partial}{\partial V} f(C_s, V_s) ds \right\} \frac{\partial}{\partial C} f(C_t, V_t). \quad (\text{B.9})$$

From (6), we obtain

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

Lemma B.1 shows that, in equilibrium, $V_t = J(\beta^{-1}C_t, \mu_t, \lambda_t)$, where the form of J is given in (B.1). It follows that

$$\frac{\partial}{\partial C} f(C_t, V_t) = \beta^\gamma C_t^{-\gamma} I(\mu_t, \lambda_t). \quad (\text{B.10})$$

Equation B.8 then follows from (B.9). \square

Proof of Theorem 1 We apply Ito's Lemma to (B.8) to find

$$\frac{d\pi_t}{\pi_{t-}} = \mu_{\pi t} dt + \sigma_{\pi t} dB_t + \sum_{j=1,2} \frac{\mathcal{J}_j(\pi_t)}{\pi_{t-}} dN_{jt},$$

where

$$\sigma_{\pi t} = \left[-\gamma\sigma, \quad b_{\lambda_1}\sigma_{\lambda_1}\sqrt{\lambda_{1t}}, \quad b_{\lambda_2}\sigma_{\lambda_2}\sqrt{\lambda_{2t}} \right], \quad (\text{B.11})$$

and

$$\frac{\mathcal{J}_j(\pi_t)}{\pi_t} = e^{b_{\mu_j} Z_{jt}} - 1, \quad \text{for } j = 1, 2. \quad (\text{B.12})$$

Moreover, from Lemma B.1, it follows that a , b_{μ_j} and b_{λ_j} take the form described. It follows immediately 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 $b_{\lambda_2} < 0$. \square

Proof of Corollary 2 The risk-free rate is obtained by taking the derivative of the HJB (B.5) with respect to α_t , evaluating at $\alpha_t = 1$ and setting it equal to 0. \square

Lemma B.3. *The drift of the state-price density is given by*

$$\mu_{\pi t} = -r_t - \lambda_t^\top \frac{\bar{\mathcal{J}}(\pi_t)}{\pi_t} \quad (\text{B.13})$$

$$= -\beta - \mu_{Ct} + \gamma\sigma^2 - \sum_{j=1,2} \lambda_{jt} E_{\nu_j} [e^{b_{\mu_j} Z_{jt}} - 1]. \quad (\text{B.14})$$

Proof Statement (B.13) follows from the absence of arbitrage. Statement (B.14) follows from Theorem 1, specifically,

$$\lambda_t^\top \frac{\bar{\mathcal{J}}(\pi_t)}{\pi_t} = \sum_{j=1,2} \lambda_{jt} E_{\nu_j} [e^{b_{\mu_j} Z_{jt}} - 1],$$

and from the expression for r_t in Corollary 2. \square

B.3 General equity pricing results

We establish some lemmas useful for pricing equities that require absence of arbitrage, but do not require parametric assumption on state prices or dividends.

We first establish a no-arbitrage restriction on the price today of a single dividend payment in τ periods.

Lemma B.4. *Let $H_t = H(D_t, \mu_t, \lambda_t, \tau)$ denote the time t price of a single future dividend payment at time $t + \tau$:*

$$H(D_t, \mu_t, \lambda_t, \tau) = E_t \left[\frac{\pi_{t+\tau}}{\pi_t} D_{t+\tau} \right]. \quad (\text{B.15})$$

Then,

$$\frac{dH_t}{H_t} = \mu_{H(\tau),t} dt + \sigma_{H(\tau),t} dB_t + \sum_{j=1,2} \frac{\mathcal{J}_j(H_t)}{H_t} dN_{jt} \quad (\text{B.16})$$

for scalar processes $\mu_{H(\tau),t}$ and (row) vector processes $\sigma_{H(\tau),t}$. Moreover,

$$\mu_{\pi t} + \mu_{H(\tau),t} + \sigma_{\pi t} \sigma_{H(\tau),t}^\top + \frac{1}{\pi_t H_t} \lambda_t^\top \bar{\mathcal{J}}(\pi_t H_t) = 0. \quad (\text{B.17})$$

Proof Equation B.16 is a result of Ito's Lemma and the Markov property of the state variables and dividends (which implies that H is indeed a function as stated). No-arbitrage implies

$$\pi_t H(D_t, \lambda_t, \mu_t, T-t) = E_t [\pi_s H(D_s, \lambda_s, \mu_s, T-s)] \quad \text{for } s > t. \quad (\text{B.18})$$

For the remainder of the argument, we simplify notation by writing $\mu_{Ht} = \mu_{H(\tau),t}$ and $\sigma_{Ht} = \sigma_{H(\tau),t}$. Ito's Lemma applied to $\pi_t H_t$ implies

$$\begin{aligned} \pi_t H_t = \pi_0 H_0 + \int_0^t \pi_s H_s (\mu_{Hs} + \mu_{\pi s} + \sigma_{\pi s} \sigma_{Hs}^\top) ds + \int_0^t \pi_s H_s (\sigma_{Hs} + \sigma_{\pi s}) dB_s \\ + \sum_{j=1,2} \sum_{0 < s_{ij} \leq t} \left(\pi_{s_{ij}} H_{s_{ij}} - \pi_{s_{ij}^-} H_{s_{ij}^-} \right), \end{aligned} \quad (\text{B.19})$$

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.19) yields:

$$\begin{aligned} \pi_t H_t = \pi_0 H_0 + \underbrace{\int_0^t \pi_s H_s \left(\mu_{Hs} + \mu_{\pi s} + \sigma_{\pi s} \sigma_{Hs}^\top + \sum_{j=1,2} \lambda_j \frac{\bar{\mathcal{J}}_j(\pi_s H_s)}{\pi_s H_s} \right) ds}_{(1)} \\ + \underbrace{\int_0^t \pi_s H_s (\sigma_{Hs} + \sigma_{\pi s}) dB_s}_{(2)} \\ + \underbrace{\sum_{j=1,2} \left(\sum_{0 < s_{ij} \leq t} \left(\pi_{s_{ij}} H_{s_{ij}} - \pi_{s_{ij}^-} H_{s_{ij}^-} \right) - \int_0^t \pi_s H_s \lambda_j \bar{\mathcal{J}}_j(\pi_s H_s) ds \right)}_{(3)}. \end{aligned} \quad (\text{B.20})$$

Equation B.18 implies that $\pi_t H_t$ is a martingale. Moreover, the terms labeled (2) and (3) on the right hand side of (B.20) have zero expectation. Therefore the term labeled (1) must also equal zero in expectation. The process $\pi_t H_t$ is strictly positive, and given that the equation must hold for all t , the integrand must equal zero. \square

We now prove an extension of Lemma B.4 that holds for the integral of functions $H(\cdot, \tau)$ over τ .

Lemma B.5. *Define*

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

assuming this indefinite integral exists. Then

$$\frac{dF_t}{F_t} = \mu_{F_t} dt + \sigma_{F_t} dB_t + \sum_{j=1,2} \frac{\mathcal{J}_j(F_t)}{F_t} dN_{jt}, \quad (\text{B.21})$$

for a scalar process μ_{F_t} and a 1×3 vector process σ_{F_t} , satisfying

$$\mu_{\pi t} + \mu_{F_t} + \frac{D_t}{F_t} + \sigma_{\pi t} \sigma_{F_t}^\top + \lambda_t^\top \frac{\bar{\mathcal{J}}(\pi_t F_t)}{\pi_t F_t} = 0. \quad (\text{B.22})$$

Proof First note that (B.21) follows from Ito's Lemma. We will show (B.22) using the corresponding result for H , Lemma B.4.⁴¹ By applying Ito's Lemma applied to both F and H , we find

$$F(D_t, \mu_t, \lambda_t) \sigma_{F_t} = \int_0^\infty H(D_t, \mu_t, \lambda_t, \tau) \sigma_{H(\tau),t} d\tau$$

and

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

Moreover, because H is a function of τ but F is not,

$$F(D_t, \mu_t, \lambda_t) \mu_{F_t} = \int_0^\infty \left(H(D_t, \lambda_t, \mu_t, \tau) \mu_{H(\tau),t} - \frac{\partial}{\partial \tau} H(D_t, \mu_t, \lambda_t, \tau) \right) d\tau. \quad (\text{B.23})$$

Equation B.23 can be rigorously derived by applying Ito's Lemma to F , and differentiating under the integral sign. Namely,

$$F(D_t, \mu_t, \lambda_t) \mu_{F_t} = \int_0^\infty \left(H_t(\tau) \mu_{D_t} + \sum_{j=1,2} \frac{\partial}{\partial \lambda_j} H_t(\tau) (\bar{\lambda}_j - \lambda_j) + \sum_{j=1,2} \frac{\partial}{\partial \mu_j} H_t(\tau) \mu_j + \frac{1}{2} \sum_{j=1,2} \frac{\partial^2}{\partial \lambda_j^2} H_t(\tau) \right) d\tau,$$

where $H_t(\tau) = H(D_t, \mu_t, \lambda_t, \tau)$. We then observe that $\mu_{H(\tau),t}$ is equal to the integrand, plus $\frac{\partial}{\partial \tau} H(D_t, \mu_t, \lambda_t, \tau)$.

Finally,

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

⁴¹It is also possible to show (B.22) directly along the same lines as Lemma B.4.

The first equality holds because $H_t(\tau)$ must equal zero in the limit as τ approaches infinity; otherwise the indefinite integral would not exist as assumed. The second equality follows from no-arbitrage. Equation B.22 then follows Lemma B.4. \square

Consider the instantaneous expected return described in Section 2.3. Because we only need to assume a Markov structure, we will use superscript D to denote the fact that we are describing returns on a general dividend claim, as opposed to, say, the market portfolio. The following lemma contains a convenient and general characterization of the risk premium in economies with jumps.

Lemma B.6. *Let r_t^D denote the instantaneous expected return*

$$r_t^D = \mu_{F_t} + \frac{D_t}{F_t} + \lambda_t^\top \frac{\bar{\mathcal{J}}(F_t)}{F_t}. \quad (\text{B.24})$$

Then

$$r_t^D - r_t = -\sigma_{\pi t} \sigma_{F_t}^\top - \sum_{j=1,2} \lambda_{jt} E_{\nu_j} \left[\frac{\mathcal{J}_j(F_t)}{F_t} \frac{\mathcal{J}_j(\pi_t)}{\pi_t} \right]. \quad (\text{B.25})$$

Proof We start with the result in Lemma B.5, substituting in for $\mu_{\pi t}$ using (B.13):

$$\underbrace{-r_t - \lambda_t^\top \frac{\bar{\mathcal{J}}(\pi_t)}{\pi_t}}_{\mu_{\pi t}} + \mu_{F_t} + \frac{D_t}{F_t} + \sigma_{\pi t} \sigma_{F_t}^\top + \lambda_t^\top \frac{\bar{\mathcal{J}}(\pi_t F_t)}{\pi_t F_t} = 0.$$

We then substitute in for $\mu_{F_t} + \frac{D_t}{F_t}$ using the definition of r_t^D , (B.24), and add and subtract $\lambda_t^\top \frac{\bar{\mathcal{J}}(F_t)}{F_t}$ to find

$$r_t^D - r_t + \sigma_{\pi t} \sigma_{F_t}^\top - \lambda_t^\top \left(\frac{\bar{\mathcal{J}}(\pi_t)}{\pi_t} + \frac{\bar{\mathcal{J}}(F_t)}{F_t} - \frac{\bar{\mathcal{J}}(\pi_t F_t)}{\pi_t F_t} \right) = 0 \quad (\text{B.26})$$

Finally note that by definition of \mathcal{J} ,

$$E_{\nu_j} \left[\frac{\mathcal{J}_j(F_t)}{F_t} \frac{\mathcal{J}_j(\pi_t)}{\pi_t} \right] = \frac{\bar{\mathcal{J}}_j(F_t \pi_t)}{F_t \pi_t} - \frac{\bar{\mathcal{J}}_j(F_t)}{F_t} - \frac{\bar{\mathcal{J}}_j(\pi_t)}{\pi_t}, \text{ for } j = 1, 2.$$

The result follows. \square

B.4 Pricing the aggregate market

We now specialize to the case of the market portfolio as defined in the main text.

Lemma B.7. *Assume dividends follow the process (13) with the state-price density described in Appendix B.2. The function H , defined in (B.15), takes an exponential form:*

$$H(D_t, \mu_t, \lambda_t, \tau) = D_t \exp \left\{ a_\phi(\tau) + b_{\phi\mu}(\tau)^\top \mu_t + b_{\phi\lambda}(\tau)^\top \lambda_t \right\}, \quad (\text{B.27})$$

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

$$b_{\phi\mu_j}(\tau) = \frac{\phi - 1}{\kappa_{\mu_j}} (1 - e^{-\kappa_{\mu_j}\tau}), \quad (\text{B.28})$$

$$\frac{db_{\phi\lambda_j}}{d\tau} = \frac{1}{2} \sigma_{\lambda_j}^2 b_{\phi\lambda_j}(\tau)^2 + \left(b_{\lambda_j} \sigma_{\lambda_j}^2 - \kappa_{\lambda_j} \right) b_{\phi\lambda_j}(\tau) + E_{\nu_j} \left[e^{b_{\mu_j} Z_{jt}} \left(e^{b_{\phi\mu_j}(\tau) Z_{jt}} - 1 \right) \right], \quad (\text{B.29})$$

$$\frac{da_\phi}{d\tau} = \bar{\mu}_D - \bar{\mu}_C - \beta + \gamma \sigma^2 (1 - \phi) + b_{\phi\lambda}(\tau)^\top (\kappa_\lambda * \bar{\lambda}). \quad (\text{B.30})$$

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

Proof First note that the boundary conditions follow from $H(D_t, \mu_t, \lambda_t, 0) = D_t$. As earlier, we write $H_t = H(D_t, \mu_t, \lambda_t, \tau)$ to simplify notation. To prove the remaining statements, we conjecture the exponential form (B.27). Recall that

$$\frac{dH_t}{H_t} = \mu_{H(\tau),t} dt + \sigma_{H(\tau),t} dB_t + \sum_{j=1,2} \frac{\mathcal{J}_j(H_t)}{H_t} dN_{jt}$$

as in Lemma B.4. Ito's Lemma applied to (B.27) implies

$$\frac{\bar{\mathcal{J}}_j(\pi_t H_t)}{\pi_t H_t} = E_{\nu_j} \left[e^{(b_{\mu_j} + b_{\mu_j} \phi(\tau)) Z_{jt}} - 1 \right], \quad (\text{B.31})$$

$$\begin{aligned} \mu_{H(\tau),t} &= \frac{1}{H} \left(\frac{\partial H}{\partial D} \mu_{Dt} + \frac{\partial H}{\partial \lambda} (\kappa_\lambda * (\bar{\lambda} - \lambda_t)) - \frac{\partial H}{\partial \mu} (\kappa_\mu * \mu_t) \right. \\ &\quad \left. - \frac{\partial H}{\partial \tau} + \frac{1}{2} \left(\frac{\partial^2 H}{\partial \lambda^2} \right) (\sigma_\lambda^2 * \lambda_t) \right) \\ &= \mu_{Dt} + b_{\phi\lambda}(\tau)^\top (\kappa_\lambda * (\bar{\lambda} - \lambda_t)) - b_{\phi\mu}(\tau)^\top (\kappa_\mu * \mu_t) \\ &\quad - \left(\frac{da_\phi}{d\tau} + \lambda_t^\top \frac{db_{\phi\lambda}}{d\tau} + \mu_t^\top \frac{db_{\phi\mu}}{d\tau} \right) + \frac{1}{2} (b_{\phi\lambda}(\tau)^2)^\top (\sigma_\lambda^2 * \lambda_t), \end{aligned} \quad (\text{B.32})$$

and

$$\begin{aligned}\sigma_{H(\tau),t} &= \frac{1}{H} \left(\frac{\partial H}{\partial D} \mu_D [\sigma_D, 0, 0] + \frac{\partial H}{\partial \lambda_1} [0, \sigma_{\lambda_1} \sqrt{\lambda_{1t}}, 0] + \frac{\partial H}{\partial \lambda_2} [0, 0, \sigma_{\lambda_2} \sqrt{\lambda_{2t}}] \right) \\ &= \left[\phi \sigma, b_{\phi \lambda_1}(\tau) \sigma_{\lambda_1} \sqrt{\lambda_{1t}}, b_{\phi \lambda_2}(\tau) \sigma_{\lambda_2} \sqrt{\lambda_{2t}} \right].\end{aligned}\quad (\text{B.33})$$

We now apply Lemma B.4, substituting (B.31–B.33), along with state-price density expressions (8) and (B.14) into the no-arbitrage condition (B.17) to find

$$\begin{aligned}\mu_D t + b_{\phi \lambda}(\tau)^\top (\kappa_\lambda * (\bar{\lambda} - \lambda_t)) - b_{\phi \mu}(\tau)^\top (\kappa_\mu * \mu_t) + \frac{1}{2} (b_{\phi \lambda}(\tau)^2)^\top (\sigma_\lambda^2 * \lambda_t) \\ - \gamma \phi \sigma^2 + b_{\phi \lambda}(\tau)^\top (b_\lambda * \sigma_\lambda^2 * \lambda_t) - \beta - \mu_C t + \gamma \sigma^2 (1 - \phi) + \sum_{j=1,2} \lambda_{jt} E_{\nu_j} \left[e^{(b_{\mu_j} + b_{\phi \mu_j}(\tau)) Z_{jt}} - e^{b_{\mu_j} Z_{jt}} \right] \\ - \left(\frac{da_\phi}{d\tau} + \lambda_t^\top \frac{db_{\phi \lambda}}{d\tau} + \mu_t^\top \frac{db_{\phi \mu}}{d\tau} \right) = 0.\end{aligned}$$

Matching the terms multiplying μ_j implies

$$\frac{db_{\phi \mu_j}}{d\tau} = -\kappa_{\mu_j} b_{\phi_j \mu} + (\phi - 1);$$

Equation B.28 then follows from the boundary condition. Matching the terms multiplying λ_j implies (B.29) and matching the constant terms implies (B.30). This also verifies the conjectured form for H . \square

Under our assumptions, the solutions $b_{\phi \lambda_j}(\tau)$ have finite limits.

Lemma B.8.

$$\lim_{\tau \rightarrow \infty} b_{\phi \lambda_j}(\tau) = -\frac{1}{\sigma_{\lambda_j}^2} \left(\zeta_{\phi_j} - \kappa_{\lambda_j} + b_{\lambda_j} \sigma_{\lambda_j}^2 \right), \quad (\text{B.34})$$

where

$$\zeta_{\phi_j} = \sqrt{(b_{\lambda_j} \sigma_{\lambda_j}^2 - \kappa_{\lambda_j})^2 - 2 E_{\nu_j} \left[e^{(b_{\mu_j} + \frac{\phi-1}{\kappa_{\mu_j}}) Z_j} - e^{b_{\mu_j} Z_j} \right] \sigma_{\lambda_j}^2}. \quad (\text{B.35})$$

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

$$0 = \frac{1}{2}\sigma_{\lambda_j}^2 \bar{b}_{\phi\lambda_j}^2 + (b_{\lambda_j}\sigma_{\lambda_j}^2 - \kappa_{\lambda_j})\bar{b}_{\phi\lambda_j} + E_{\nu_j} \left[e^{\left(b_{\mu_j} + \frac{\phi-1}{\kappa_{\mu_j}}\right)Z_{jt}} - e^{b_{\mu_j}Z_{jt}} \right].$$

Equation B.34 gives a solution to this equation.⁴²

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

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

Therefore,

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

Therefore, by (B.34) for $j = 1$, $\lim_{\tau \rightarrow \infty} b_{\phi\lambda_1}(\tau) < 0$

Now, note that $Z_2 > 0$ implies that

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

Therefore,

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

(Assumption 2 implies that ζ_{ϕ_2} is real-valued). Moreover, $b_{\lambda_2} < 0$ (Theorem 1), so

$$0 < \zeta_{\phi_2} < \kappa_{\lambda_2} - b_{\lambda_2}\sigma_{\lambda_2}^2.$$

Therefore, by (B.34) for $j = 2$, $\lim_{\tau \rightarrow \infty} b_{\phi\lambda_2}(\tau) > 0$. □

Lemma B.9. *The functions $b_{\phi\lambda_j}(\tau)$ are monotonic.*

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

$$b_{\phi\lambda_1}(\tau^*) = \frac{1}{\sigma_{\lambda_1}^2} \left(\sqrt{(b_{\lambda_1}\sigma_{\lambda_1}^2 - \kappa_{\lambda_1})^2 - 2E_{\nu_1} \left[e^{\left(b_{\mu_1} + \frac{\phi-1}{\kappa_{\mu_1}}\right)Z_1} - e^{b_{\mu_1}Z_1} \right] \sigma_{\lambda_1}^2} - \kappa_{\lambda_1} + b_{\lambda_1}\sigma_{\lambda_1}^2 \right). \quad (\text{B.36})$$

⁴²We use the same procedure to select which of the two solutions of the quadratic to use as we used in Appendix B.2. That is, we chose the one in which rare events do not affect prices if there are in fact no rare events. We have verified that (B.34) does indeed correspond to the limit when the ordinary differential equation (17) is solved numerically.

However, differentiating (B.36) with respect to τ implies $db_{\phi\lambda_1}(\tau^*)/d\tau \neq 0$. 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. The same reasoning works for $b_{\phi\lambda_2}(\tau)$ \square

Corollary B.10. 1. $b_{\phi\mu_1}(\tau) > 0$,

2. $b_{\phi\mu_2}(\tau) > 0$,

3. $b_{\phi\lambda_1}(\tau) < 0$,

4. $b_{\phi\lambda_2}(\tau) > 0$.

Proof Results 1 and 2 follow immediately from the form of these functions in Lemma B.7 and our assumption that $\phi > 1$. Result 3 follows from $b_{\phi\lambda_1}(0) = 0$ and $\lim_{\tau \rightarrow \infty} b_{\phi\lambda_1}(\tau) < 0$ by Lemma B.8. By Lemma B.9, $b_{\phi\lambda_1}(\tau) < 0$. Similar reasoning shows $b_{\phi\lambda_2}(\tau) > 0$. \square

The following lemma shows that the integral determining the price of the market converges. The logic is the similar to that driving the convergence of a geometric sum.

Lemma B.11.

$$\lim_{\bar{\tau} \rightarrow \infty} \int_0^{\bar{\tau}} H(D, \mu, \lambda, \tau) d\tau < \infty$$

for all $D > 0$, vectors λ with both elements ≥ 0 and all μ_{jt} .

Proof Define $\bar{b}_{\phi\lambda_2}$ to be the limit defined in Lemma B.8, and $\bar{b}_{\phi\mu_2}$ to be the limit of $b_{\phi\mu_2}(\tau)$ (the existence of this limit follows immediately from the form of the function). By Corollary B.10,

$$H(D, \mu, \lambda, \tau) \leq D e^{\bar{b}_{\phi\lambda_2} \lambda_{2t} + \bar{b}_{\phi\mu_2} \mu_{2t}} e^{a_\phi(\tau)},$$

and thus it suffices to show that

$$\lim_{\bar{\tau} \rightarrow \infty} \int_0^{\bar{\tau}} e^{a_\phi(\tau)} d\tau < \infty.$$

Note that by (18),

$$a_\phi(\tau) = (\bar{\mu}_D - \bar{\mu}_C - \beta + \gamma\sigma^2(1 - \phi))\tau + \sum_{j=1,2} \kappa_{\lambda_j} \bar{\lambda}_j \int_0^\tau b_{\phi\lambda_j}(u) du.$$

Now define

$$\bar{a}_\phi = \bar{\mu}_D - \bar{\mu}_C - \beta + \gamma\sigma^2(1 - \phi) + \bar{b}_{\phi\lambda_2}\kappa_{\lambda_2}\bar{\lambda}_2.$$

It follows from the fact that $b_{\phi\lambda_2}(\tau) > 0$ and is increasing (Lemma B.9, Corollary B.10) that

$$a_\phi(\tau) \leq \bar{a}_\phi\tau + \kappa_{\lambda_1}\bar{\lambda}_1 \int_0^\tau b_{\phi\lambda_1}(u) du$$

Because $b_{\phi\lambda_1}(\tau)$ is decreasing, by Assumption 3 and by the equations for the limits in Lemma B.8, there exists a τ_0 such that

$$\bar{a}_\phi + b_{\phi\lambda_1}(\tau_0)\kappa_{\lambda_1}\bar{\lambda}_1 < 0.$$

Write

$$\begin{aligned} \int_0^{\bar{\tau}} e^{a_\phi(\tau)} d\tau &\leq \int_0^{\bar{\tau}} e^{\bar{a}_\phi\tau + \int_0^\tau b_{\phi\lambda_1}(u)\kappa_{\lambda_1}\bar{\lambda}_1 du} d\tau \\ &= e^{\bar{a}_\phi\tau_0 + \int_0^{\tau_0} b_{\phi\lambda_1}(u)\kappa_{\lambda_1}\bar{\lambda}_1 du} \int_0^{\bar{\tau}} e^{\bar{a}_\phi(\tau-\tau_0) + \kappa_{\lambda_1}\bar{\lambda}_1 \int_{\tau_0}^\tau b_{\phi\lambda_1}(u) du} d\tau \\ &\leq e^{\bar{a}_\phi\tau_0 + \int_0^{\tau_0} b_{\phi\lambda_1}(u)\kappa_{\lambda_1}\bar{\lambda}_1 du} \int_0^{\bar{\tau}} e^{\bar{a}_\phi(\tau-\tau_0) + \kappa_{\lambda_1}\bar{\lambda}_1 b_{\phi\lambda_1}(\tau_0)(\tau-\tau_0)} d\tau \end{aligned}$$

where the last line follows from the monotonicity of $b_{\phi\lambda_1}(\tau)$. Convergence as $\bar{\tau} \rightarrow \infty$ follows from the properties of the exponential function. \square

Proof of Theorem 3 Note that, by no-arbitrage,

$$F(D_t, \mu_t, \lambda_t) = E_t \int_t^\infty \frac{\pi_s}{\pi_t} D_s ds,$$

assuming the right-hand side is well-defined. It follows from Lemma B.7 that $H(D_t, \mu_t, \lambda_t, \tau) = E_t \left[\frac{\pi_{t+\tau}}{\pi_t} D_\tau \right]$ takes the required form. Lemma B.11 shows that the indefinite integral of these terms exists, and therefore we can write

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

The equation for the price-dividend ratio follows immediately from dividing this equation by D_t . \square

Proof of Corollary 4 Because

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

the dynamics of G follow immediately from Corollary B.10

Proof of Theorem 5 We use the general formula for the risk premium, given in Lemma B.6:

$$r_t^m - r_t = -\sigma_{\pi t} \sigma_{F_t}^\top - \sum_{j=1,2} \lambda_{jt} E_{\nu_j} \left[\frac{\mathcal{J}_j(F_t)}{F_t} \frac{\mathcal{J}_j(\pi_t)}{\pi_t} \right].$$

It follows from

$$F(D_t, \mu_t, \lambda_t) = D_t G(\mu_t, \lambda_t),$$

and Theorem 3 that

$$\sigma_{F_t} = \left[\phi\sigma, \frac{1}{G} \frac{\partial G}{\partial \lambda_1} \sigma_{\lambda_1} \sqrt{\lambda_{1t}}, \frac{1}{G} \frac{\partial G}{\partial \lambda_2} \sigma_{\lambda_2} \sqrt{\lambda_{2t}} \right].$$

Because dividends do not jump,

$$\frac{\mathcal{J}_j(F_t)}{F_{t-}} = \frac{\mathcal{J}_j(G_t)}{G_{t-}}$$

We can find the terms corresponding to the state-price density in Theorem 1 (see (B.11) and (B.12)). Substituting in, we find the formula for the equity premium:

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

□

B.5 Pricing results for value and growth

We first show that the price-dividend ratio for value lies below that of the market. This result does not depend on our specific assumptions about the dividend process.

Proof of Corollary 9 Consider an asset with dividend stream given by

$$\frac{dD_s^v}{D_s^v} = \mu_{D_s}^v dt + \phi\sigma dB_{C_s},$$

and $\mu_{D_s}^v = \bar{\mu}_D + \phi\mu_{1s}$, for $s \geq t$, and normalize D_t^v to D_t , the dividend on the market. Comparing the evolution of D_s^v with D_s (given in (13)), it follows that

$$D_s^v \leq D_s \quad \text{for } s \geq t. \quad (\text{B.37})$$

There is a positive probability that, for some s , the inequality in (B.37) is strict in some states of the world. Moreover, it follows from (B.9) that $\pi_s > 0$ for all s . Therefore,

$$E_t \int_t^\infty \frac{\pi_s}{\pi_t} D_s^v ds < E_t \int_t^\infty \frac{\pi_s}{\pi_t} D_s ds.$$

Dividing by D_t implies

$$E_t \int_t^\infty \frac{\pi_s}{\pi_t} \frac{D_s^v}{D_t} ds < E_t \int_t^\infty \frac{\pi_s}{\pi_t} \frac{D_s}{D_t} ds.$$

Because $D_t^v = D_t$, the left-hand side is the price-dividend ratio on the value claim. By the same reason, the right-hand side is the price-dividend ratio on the market. \square

A subtlety in the result above is that it does not require that we define the value claim at time t as an asset having the same dividend as the market. Normalizing the dividends to be equal is a technical step within the proof, not an assumption. The price-dividend ratio on the value claim is determined only by future growth in dividends and in the state-price density, not the level of dividends.

It is convenient to state the pricing equations for the value claim. This result follows from an argument directly analogous to that of Theorem 3.⁴³

Corollary B.12. *Let $F^v(D_t, \mu_t, \lambda_t)$ denote the price of the value sector, namely the claim to cash flows satisfying (25), where D_t is the current dividend.⁴⁴*

$$F^v(D_t, \mu_t, \lambda_t) = \int_0^\infty H^v(D_t, \mu_t, \lambda_t, \tau) d\tau. \quad (\text{B.38})$$

⁴³Note that if the integral defining the price of the aggregate dividend claim converges, then the integral for the value claim must also converge because dividend growth for value is lower than for the market. See Appendix A for more discussion.

⁴⁴In the statement of this theorem, what we call the current dividend is irrelevant. We could as easily call the current dividend D_t^v .

where

$$H^v(D_t, \mu_t, \lambda_t, \tau) = D_t \exp \{ a_\phi^v(\tau) + b_{\phi\mu}^v(\tau)^\top \mu_t + b_{\phi\lambda}^v(\tau)^\top \lambda_t \}, \quad (\text{B.39})$$

$$b_{\phi\mu_1}^v(\tau) = \frac{\phi - 1}{\kappa_{\mu_1}} (1 - e^{-\kappa_{\mu_1}\tau}) \quad (\text{B.40})$$

$$b_{\phi\mu_2}^v(\tau) = -\frac{1}{\kappa_{\mu_2}} (1 - e^{-\kappa_{\mu_2}\tau}), \quad (\text{B.41})$$

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

$$\frac{db_{\phi\lambda_j}^v}{d\tau} = \frac{1}{2}\sigma_{\lambda_j}^2 b_{\phi\lambda_j}^v(\tau)^2 + (b_{\lambda_j}\sigma_{\lambda_j}^2 - \kappa_{\lambda_j}) b_{\phi\lambda_j}^v(\tau) + E_{\nu_j} \left[e^{b_{\mu_j}^v Z_{jt}} \left(e^{b_{\phi\mu_j}^v(\tau) Z_{jt}} - 1 \right) \right], \quad (\text{B.42})$$

$$\frac{da_\phi^v}{d\tau} = \bar{\mu}_D - \bar{\mu}_C - \beta + \gamma\sigma^2(1 - \phi) + b_{\phi\lambda_j}^v(\tau)^\top (\kappa_\lambda * \bar{\lambda}) \quad (\text{B.43})$$

with boundary conditions $b_{\phi\lambda_j}^v(0) = a_\phi^v(0) = 0$. Furthermore, the price-dividend ratio on the value sector is given by

$$G^v(\mu_t, \lambda_t) = \int_0^\infty \exp(a_\phi^v(\tau) + b_{\phi\mu}^v(\tau)^\top \mu_t + b_{\phi\lambda}^v(\tau)^\top \lambda_t) d\tau. \quad (\text{B.44})$$

Corollary B.13. *The functions $b_{\phi\lambda_j}^v(\tau)$ are monotonic. Moreover,*

1. $b_{\phi\mu_1}^v(\tau) > 0$,
2. $b_{\phi\mu_2}^v(\tau) < 0$,
3. $b_{\phi\lambda_1}^v(\tau) < 0$,
4. $b_{\phi\lambda_2}^v(\tau) < 0$.

Proof First note that $b_{\phi\mu_1}^v(\tau) = b_{\phi\mu_1}(\tau)$, and therefore, because the ordinary differential equations are the same, $b_{\phi\lambda_1}^v(\tau) = b_{\phi\lambda_1}(\tau)$. It follows from Lemma B.9 and Corollary B.10 that monotonicity holds for $j = 1$, and also that Results 1 and 3 must hold.

To establish monotonicity for $j = 2$, we follow the same strategy as in Lemmas B.8 and B.9. As in Lemma B.8, we have

$$\lim_{\tau \rightarrow \infty} b_{\phi\lambda_2}^v(\tau) = -\frac{1}{\sigma_{\lambda_2}^2} (\zeta_{\phi_2}^v - \kappa_{\lambda_2} + b_{\lambda_2}\sigma_{\lambda_2}^2),$$

where

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

Because $Z_2 > 0$, $e^{-\frac{1}{\kappa_{\mu_2}}Z_2} < 1$ and

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

which shows that

$$\lim_{\tau \rightarrow \infty} b_{\phi_{\lambda_2}}^v(\tau) < 0.$$

The same reasoning used in Lemma B.9 then implies monotonicity.

Finally, Result 2 follows directly from the form of $b_{\phi_{\mu_2}}^v(\tau)$. Result 4 follows from monotonicity and the same reasoning as in Corollary B.10. \square

Proof of Corollary 10 The fact that the price-dividend ratio on the value claim is increasing in μ_1 and μ_2 , and decreasing in λ_1 and λ_2 follows directly from Corollary B.13 and the form of the price-dividend ratio (B.44).

Proof of Corollary 11 The first result follows from combining the general risk premium result in Lemma B.6 with the equation for the price of the value sector (26). See the proof of Theorem 5 for more detail. The second result follows from the first, and from the definition of the expected return in samples without jumps. \square

We now show that price of the growth claim is increasing in μ_1 and μ_2 , decreasing in λ_1 and increasing in λ_2 (Corollary 14 in the main text), using the corresponding results for the market and the value sector.

Proof of Corollary 14 It follows from the definition of the growth sector price,

$$F_t^g = F_t - F_t^v,$$

that

$$\frac{\partial F_t^g}{\partial \mu_{jt}} = \frac{\partial F_t}{\partial \mu_{jt}} - \frac{\partial F_t^v}{\partial \mu_{jt}} \tag{B.45}$$

and

$$\frac{\partial F_t^g}{\partial \lambda_{jt}} = \frac{\partial F_t}{\partial \lambda_{jt}} - \frac{\partial F_t^v}{\partial \lambda_{jt}}.$$

The straightforward cases correspond to $j = 2$ (booms). It follows from Corollary 4 that $\frac{\partial F_t}{\partial \mu_{2t}} > 0$ and from Corollary 10 that $\frac{\partial F_t^v}{\partial \mu_{2t}} < 0$, and similarly for derivatives with respect to λ_{2t} . Therefore, the price of the growth sector is increasing in μ_{2t} and λ_{2t} .

Now consider $j = 1$. Theorem 3 and Corollary B.12 imply that we can rewrite (B.45) as

$$\frac{\partial F_t^g}{\partial \mu_{jt}} = \int_0^\infty b_{\phi\mu_1}(\tau) H(D_t, \mu_t, \lambda_t, \tau) d\tau - \int_0^\infty b_{\phi\mu_1}^v(\tau) H^v(D_t^v, \mu_t, \lambda_t, \tau) d\tau.$$

Moreover, $b_{\phi\mu_1}(\tau) = b_{\phi\mu_1}^v(\tau)$. Because the dividend at time $t+\tau$ on value is less than or equal to the dividend on the market, and strictly lower with positive probability, $H(D_t, \mu_t, \lambda_t, \tau) > H^v(D_t^v, \mu_t, \lambda_t, \tau)$. The result follows. The same argument works for the derivative with respect to λ_{1t} . \square

C Return simulation

This section describes how we simulate returns on the market and the value and growth sectors, given a simulated series of state variables and one-period observations on dividend growth. We simulate at a monthly frequency, which is sufficiently fine to capture the joint distribution of consumption, dividends, and the state variables.

Consider an asset with price process F_t and dividend D_t . The realized return between time t and $t + \Delta t$ is given by

$$R_{t,t+\Delta t} = \frac{F_{t+\Delta t} + \int_t^{t+\Delta t} D_s ds}{F_t}$$

(see Duffie (2001, Chapter 6.L)). We specialize to the case of an asset that pays a positive dividend at each point in time (namely, the market or value claim). Using the approxima-

tion $D_{t+\Delta t}\Delta t \approx \int_t^{t+\Delta t} D_s ds$, we can compute returns as

$$\begin{aligned}
R_{t,t+\Delta t} &\approx \frac{F_{t+\Delta t} + D_{t+\Delta t}\Delta t}{F_t} \\
&= \frac{\frac{F_{t+\Delta t}}{D_{t+\Delta t}} + \Delta t}{\frac{F_t}{D_t}} \frac{D_{t+\Delta t}}{D_t} \\
&= \frac{G(\mu_{t+\Delta t}, \lambda_{t+\Delta t}) + \Delta t}{G(\mu_t, \lambda_t)} \frac{D_{t+\Delta t}}{D_t},
\end{aligned} \tag{C.1}$$

where G is the price-dividend ratio on the asset. The price-dividend ratio series is obtained from the simulated series of state variables using (19) for the market and (28) for value. Note that this definition of the return makes it unnecessary to have information on either the level of dividends or the level of prices. Thus this definition is invariant to whether value is redefined to have the same dividend at time t as the market.

Because the growth sector does not pay a dividend every period, we cannot use (C.1) to compute its return. However, its return is implied by the fact that the market is a portfolio of value and growth. Let $R_{t,t+\Delta t}^m$ denote the market return and $R_{t,t+\Delta t}^v$ denote the return on the value sector, each computed using the relevant quantities in (C.1). Furthermore, recall that F_t^v is the price of the value sector at time t and F_t is the price of the market. Then,

$$R_{t,t+\Delta t}^m \approx \frac{F_t^v}{F_t} R_{t,t+\Delta t}^v + \left(1 - \frac{F_t^v}{F_t}\right) R_{t,t+\Delta t}^g. \tag{C.2}$$

This equation is approximate rather than exact because the model is formulated in continuous time. Solving for the growth return in (C.2) yields

$$R_{t,t+\Delta t}^g = \frac{1}{1 - \frac{F_t^v}{F_t}} \left(R_{t,t+\Delta t}^m - \frac{F_t^v}{F_t} R_{t,t+\Delta t}^v \right). \tag{C.3}$$

We now assume that the time- t dividend on value and on the market are the same. That is, we define the value sector as in Section 2.4. It follows that

$$\frac{F_t^v}{F_t} = \frac{G^v(\mu_t, \lambda_t)}{G(\mu_t, \lambda_t)}. \tag{C.4}$$

Growth returns can then be computed using the state variable series and returns on value and the market using (C.3) and (C.4).

References

- Abel, Andrew, 1999, Risk premia and term premia in general equilibrium, *Journal of Monetary Economics* 43, 3–33.
- Ai, Hengjie, Mariano Massimiliano Croce, and Kai Li, 2013, Toward a Quantitative General Equilibrium Asset Pricing Model with Intangible Capital, *Review of Financial Studies* 26, 491–530.
- 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, 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 José F. Ursúa, 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 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.
- David, Alexander, and Pietro Veronesi, 2013, What Ties Return Volatilities to Price Valuations and Fundamentals?, *Journal of Political Economy* 121, 682–746.
- 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, 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.

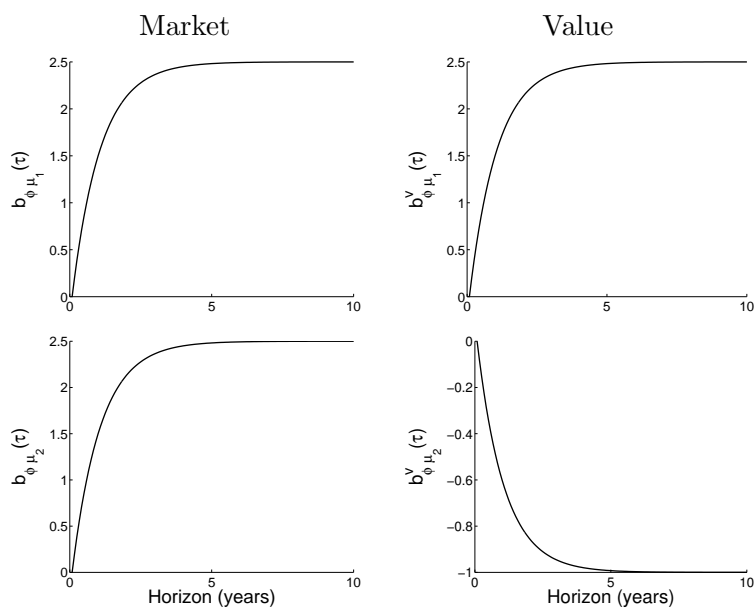
- Gourio, François, 2012, Disaster risk and business cycles, *American Economic Review* 102, 2734–2766.
- Hansen, Lars Peter, John C. Heaton, and Nan Li, 2008, Consumption strikes back? Measuring long run risk, *Journal of Political Economy* 116, 260–302.
- Jovanovic, Boyan, and Peter L. Rousseau, 2003, 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, 2014, Tail risk and asset prices, *Review of Financial Studies* 27, 2841–2871.
- 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, 2014, Conditional risk premia in currency markets and other asset classes, *Journal of Financial Economics* 114, 197–225.
- 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.
- McQuade, Timothy J., 2013, Stochastic Volatility and Asset Pricing Puzzles, Working paper, Stanford University.
- 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.

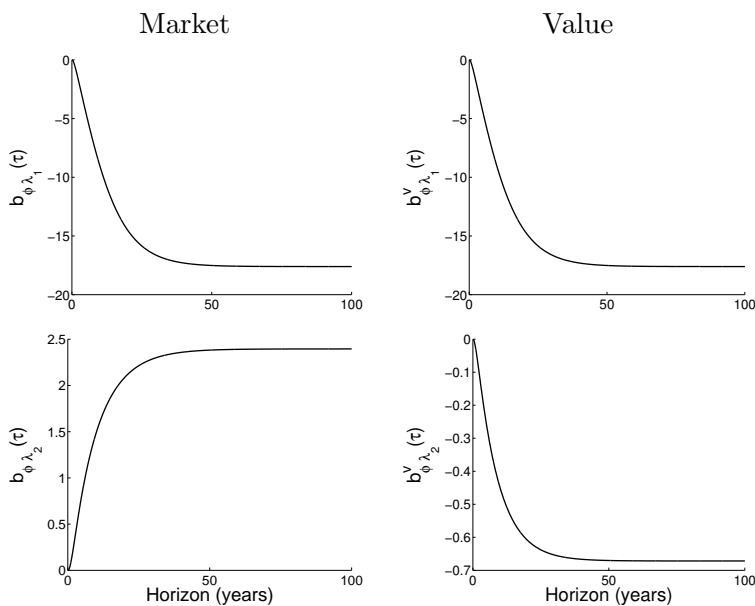
- Stambaugh, Robert F., 1999, Predictive regressions, *Journal of Financial Economics* 54, 375–421.
- Veronesi, Pietro, 2004, The Peso problem hypothesis and stock market returns, *Journal of Economic Dynamics and Control* 28, 707–725.
- Wachter, Jessica A., 2013, Can time-varying risk of rare disasters 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: Solution for the price-dividend ratio

Panel A: Expected growth rates

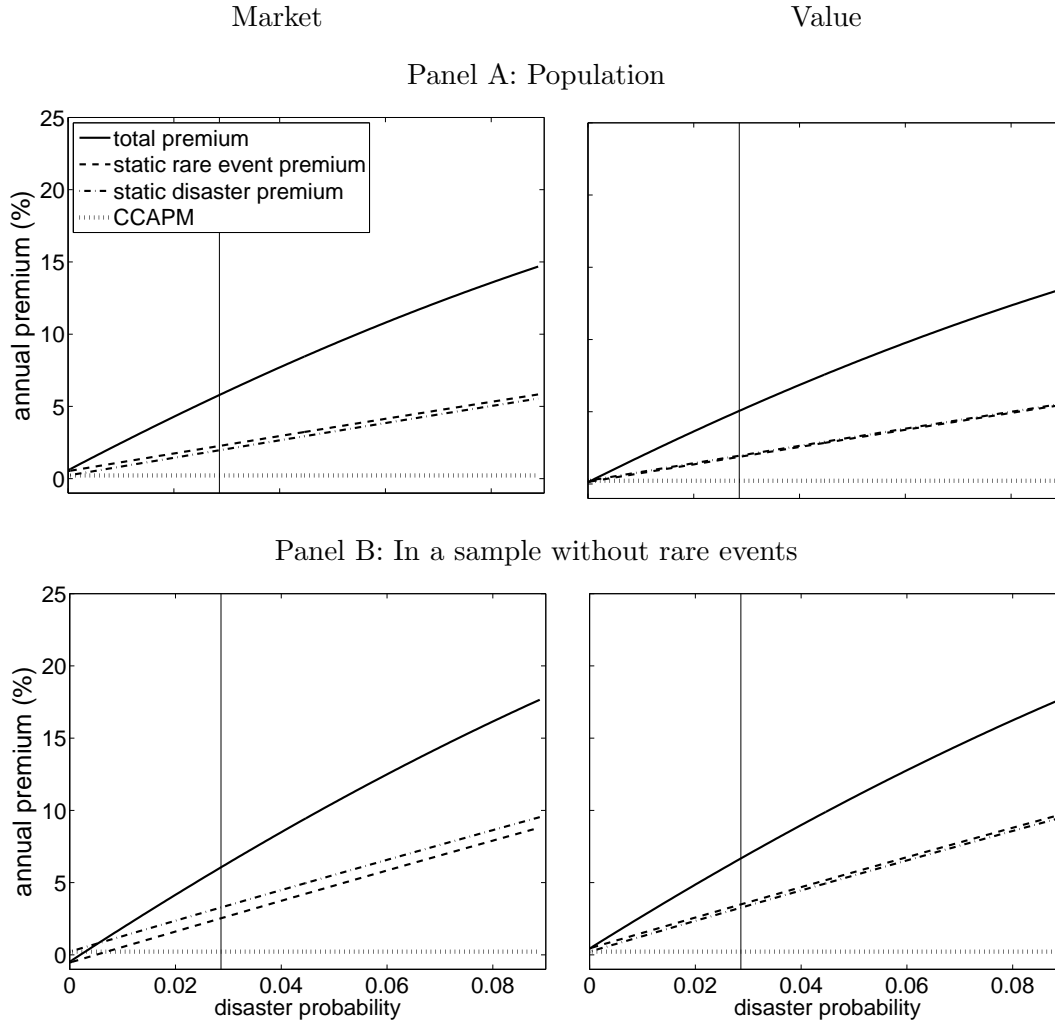


Panel B: Rare event probabilities



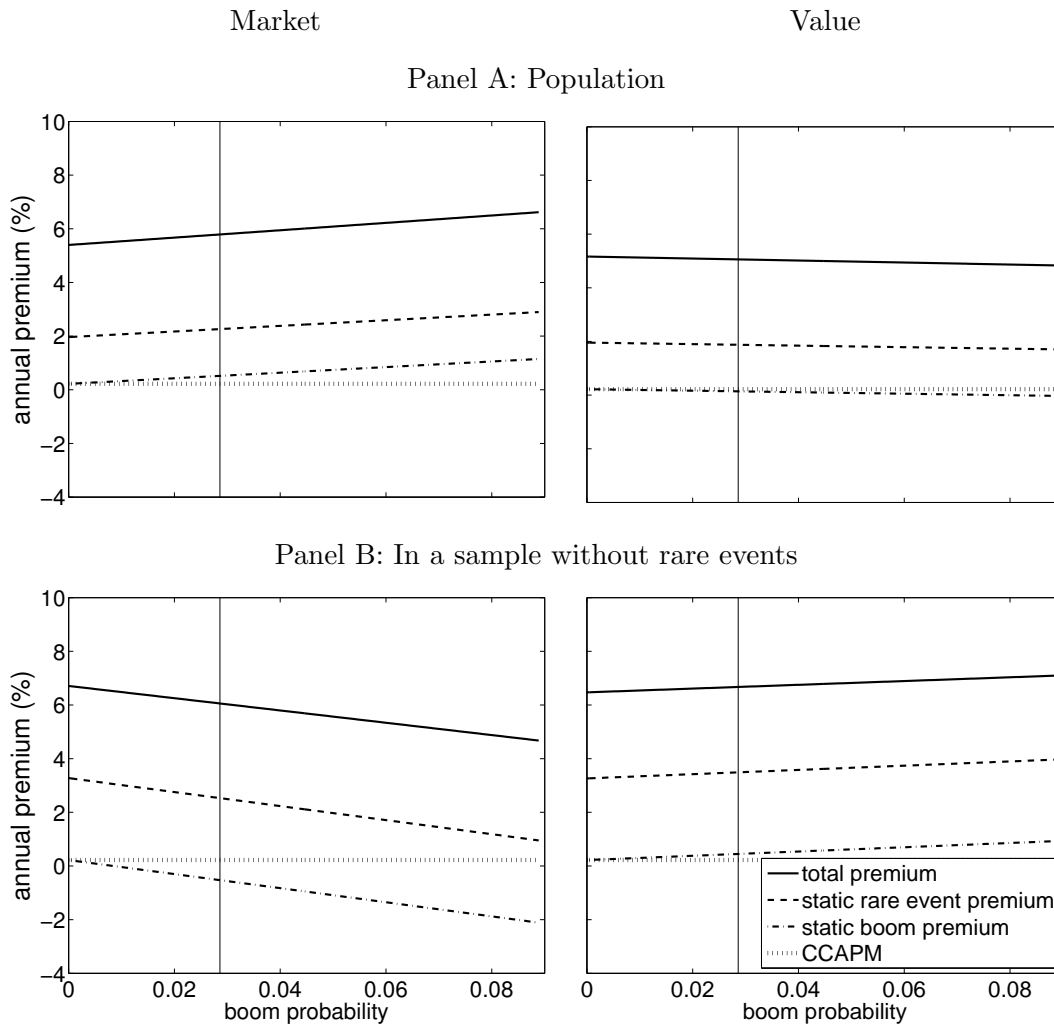
Notes: Panel A shows the coefficients multiplying μ_1 and μ_2 (the disaster and boom components in expected consumption growth, respectively) in the price-dividend ratio. Panel B shows the coefficients multiplying λ_{1t} (the probability of a disaster) and λ_{2t} (the probability of a boom). The left panel shows results for the market; the right shows results for the value premium. The scales on the left and the right may differ.

Figure 2: Risk premiums as a function of disaster probability



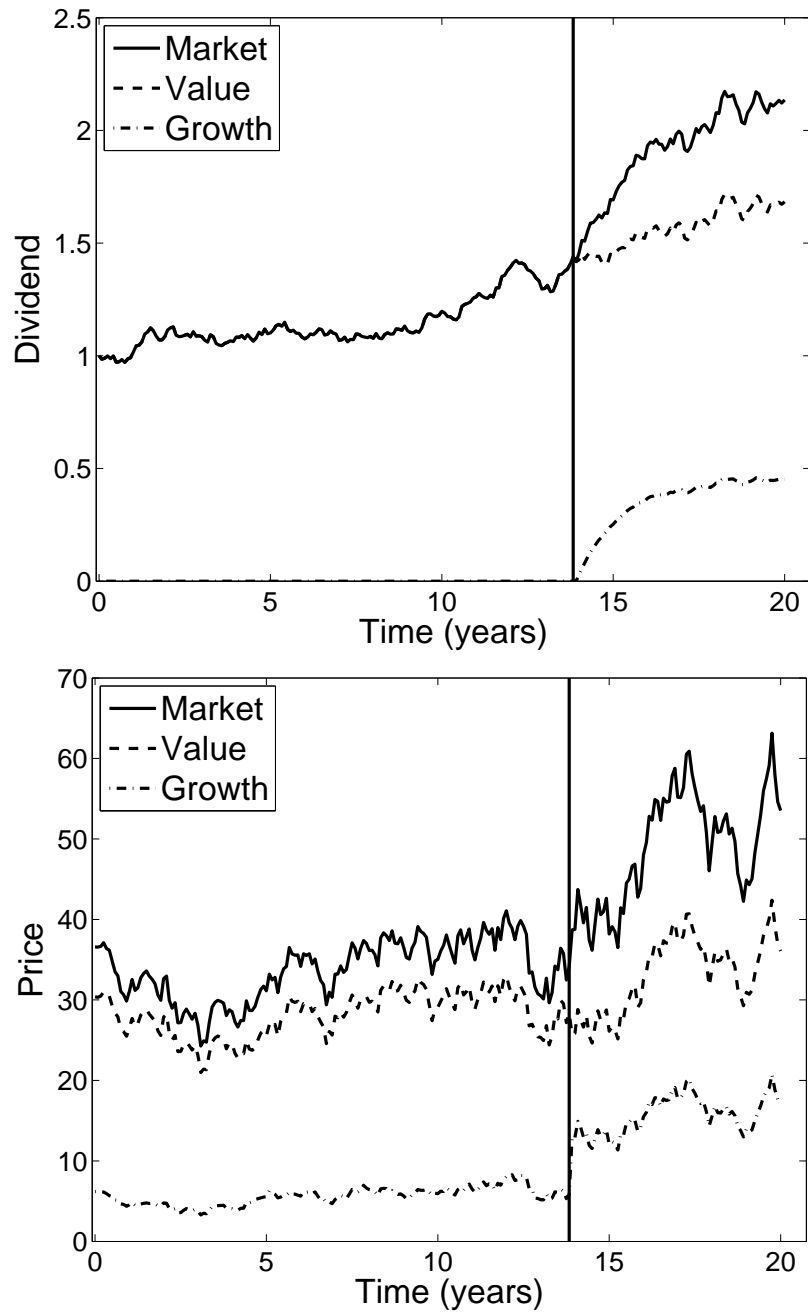
Notes: Panel A shows population risk premia as functions of the disaster probability with the boom probability fixed at its mean. The vertical line represents the mean of the disaster probability. Panel B shows what the observed expected excess return would be in a sample without rare events.

Figure 3: Risk premiums as a function of boom probability



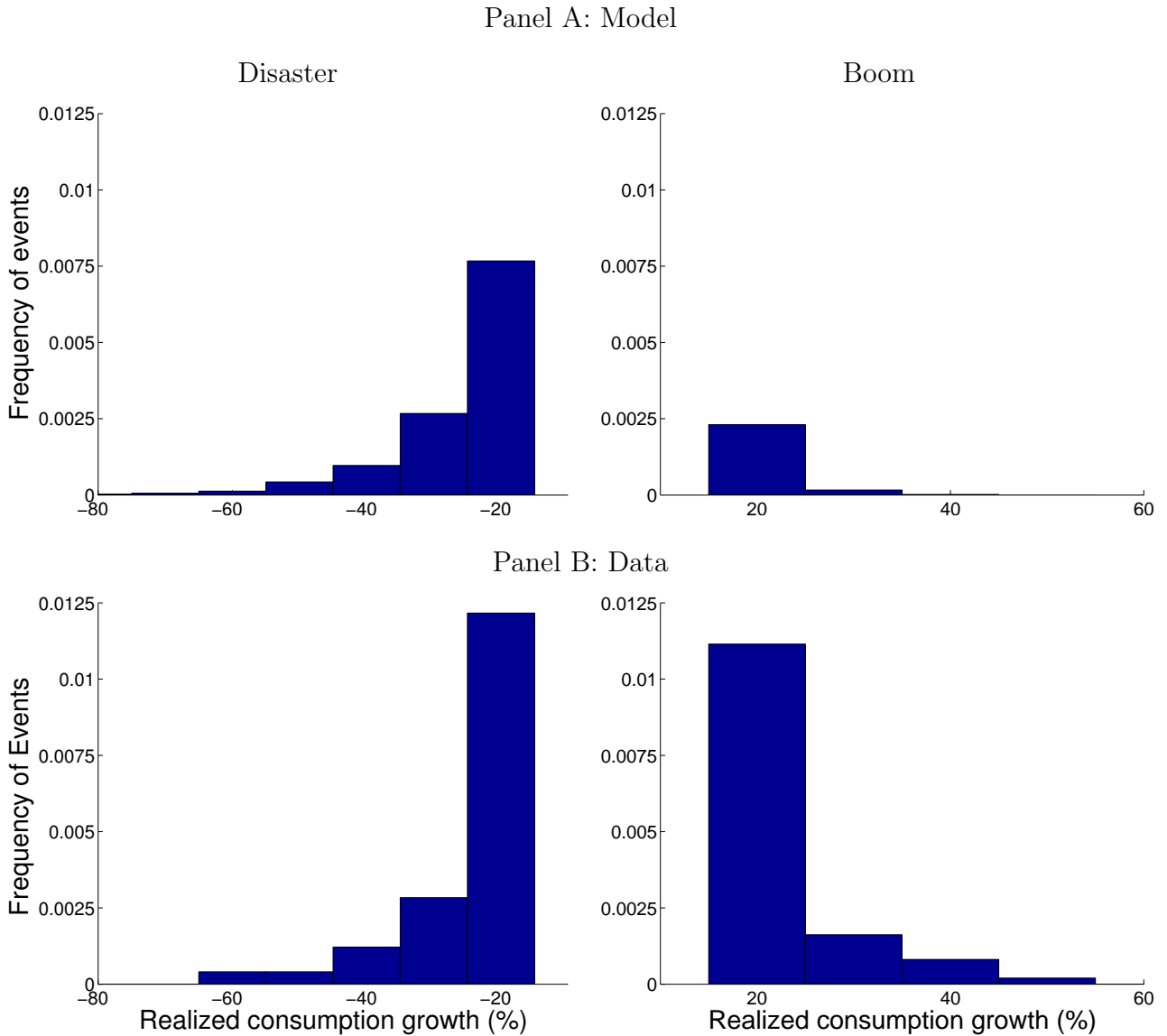
Notes: Panel A shows population risk premia as functions of the boom probability with the disaster probability fixed at its mean. The vertical line represents the mean of the boom probability. Panel B shows what the observed expected excess return would be in a sample without rare events.

Figure 4: Sample simulation path of asset dividends and prices



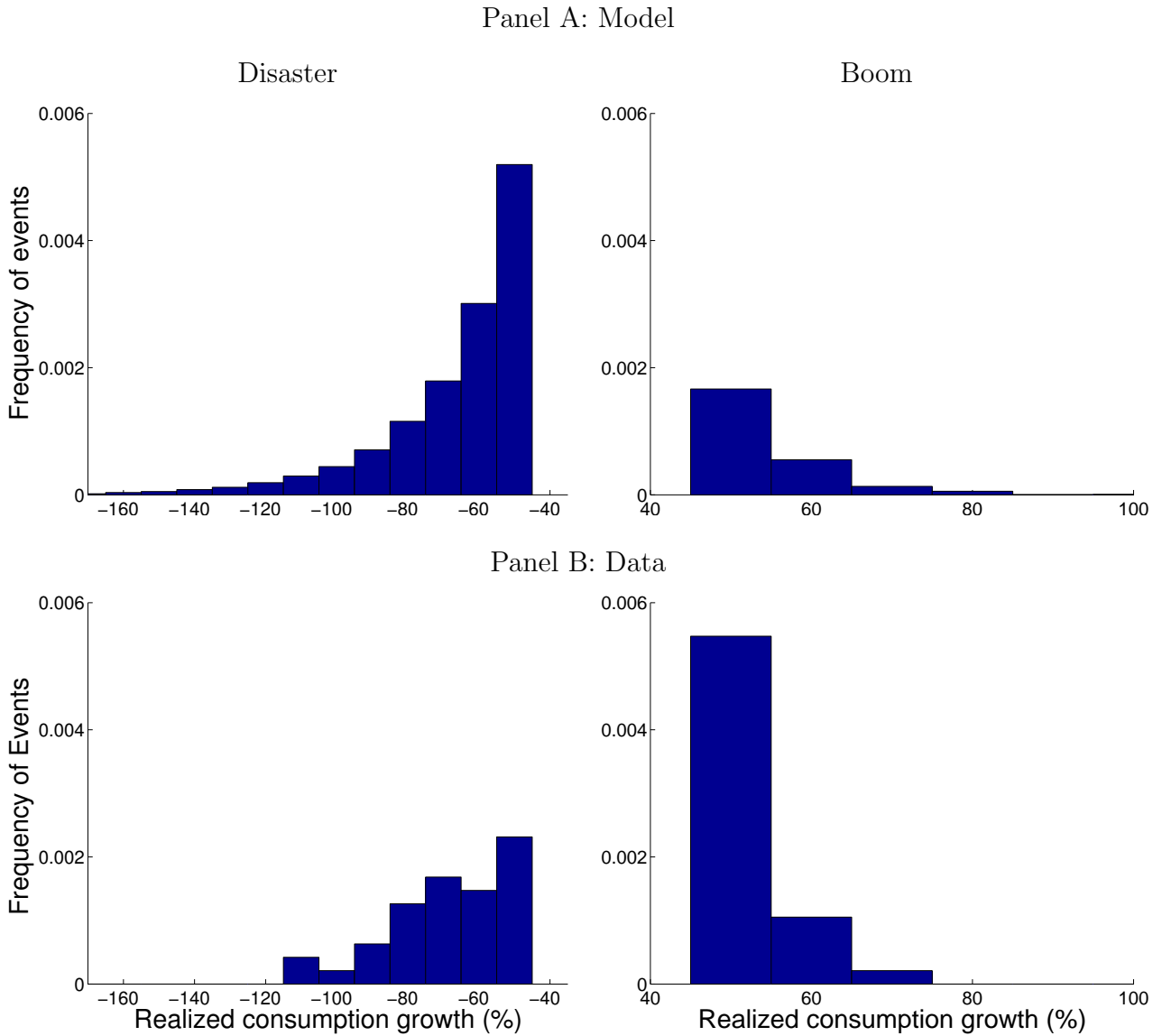
Notes: This figure shows results from a time series simulated from the model that includes a boom. The top figure shows dividends (initialized at one), while the bottom panel shows prices. The solid line marks the onset of the boom.

Figure 5: 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 Ursúa (2008). For the consumption booms, we exclude observations between 1944 and 1953.

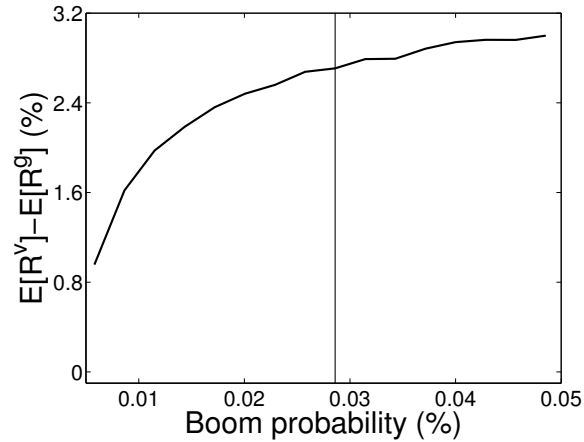
Figure 6: Tails of the five-year consumption growth rate distribution



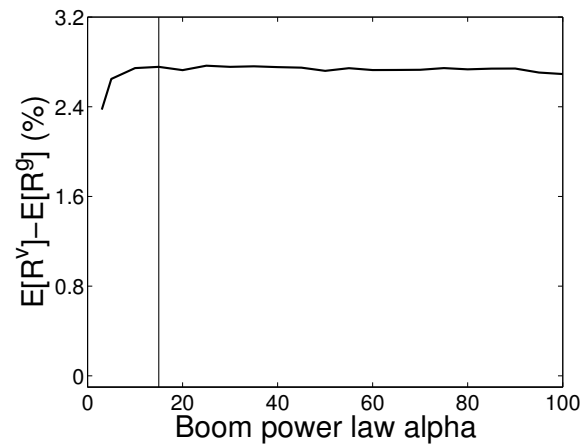
Notes: This figure shows histograms of five-year consumption growth rates. The right panel considers growth rates above 45%. The left panel considers growth rates below -45%. Panel A shows results from simulated data from the model. Panel B shows results from the data. Data are from Barro and Ursúa (2008).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

Figure 7: Sensitivity of the value premium to parameter choices

Panel A: Value premium as a function of average boom probability ($\bar{\lambda}_2$)

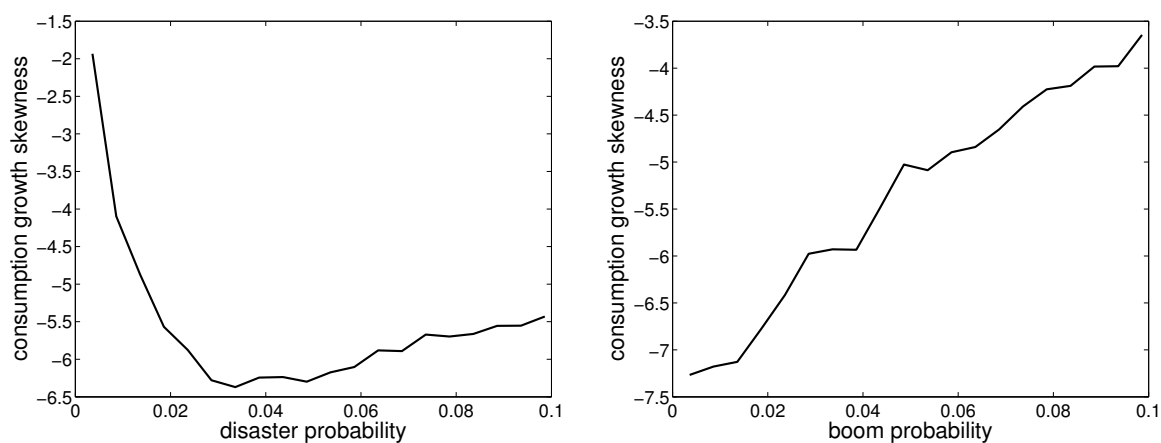


Panel B: Value premium as a function of the power law tail parameter (α)



Notes: Panel A reports the average value premium (defined as the difference between the return on the value sector and the return on the growth sector) across simulations with no rare events as a function of the average probability of a boom. All other parameters are unchanged. Panel B reports the average value premium across simulations with no rare events as a function of the power law parameter α . Lower values of α correspond to thicker tails.

Figure 8: Consumption growth skewness



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_2 , with the disaster probability equals to its mean. We calculate skewness by simulating 500,000 years of consumption growth at a monthly frequency for each value of the rare event probabilities. We aggregate the monthly consumption growth to an annual frequency and calculate the skewness over the sample.

Table 1: Effects of shocks on prices and returns

	High λ_1	High λ_2	Low μ_1	High μ_2
Riskfree rate	no change	no change	-	+
Market price-dividend ratio	-	+	-	+
Value price-dividend ratio	-	-	-	-
True equity premium	+	+	no change	no change
True risk premium on value	+	-	no change	no change
Observed equity premium	+	-	no change	no change
Observed risk premium on value	+	+	no change	no change

Notes: Signs of the effect of shocks to each state variable on the riskfree rate, price dividend ratios, and risk premia. “High λ_1 ” refers to an increase in the disaster probability. “High λ_2 ” refers to an increase in the boom probability. “Low μ_1 ” refers to a decrease in the component of expected consumption growth due to disasters. “High μ_2 ” refers to an increase in the component of expected consumption growth due to booms. The true and observed risk premia are all relative to the riskfree rate. The observed premia refers to the expected excess return that would be observed in a sample without rare events.

Table 2: 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 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 rare events 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 rare events 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 rare events 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 rare events 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
t -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
t -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
t -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
t -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 rare events 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 rare events occur.