The Role of Volatility Shocks and Rare Events in Long-Run Risk Models

Nicole Branger* Paulo Rodrigues‡ Christian Schlag§

First version: November 19, 2010

This version: November 18, 2011

* Finance Center Münster, University of Münster, Universitätsstr. 14-16, 48143 Münster, Germany. E-mail: nicole.branger@wiwi.uni-muenster.de

‡ Maastricht University, Department of Finance, Tongersestraat 53, 6211 LM Maastricht, The Netherlands. E-mail: p.rodrigues@maastrichtuniversity.nl

§ Goethe University, House of Finance, Grueneburgplatz 1, 60323 Frankfurt am Main. E-mail: schlag@finance.uni-frankfurt.de

The Role of Volatility Shocks and Rare Events in Long-Run Risk Models

First version: November 19, 2010

This version: November 18, 2011

Abstract

We study a long-run risk model with a stochastic consumption growth rate, a stochastic volatility, a stochastic jump intensity, and a stochastic mean reversion level for the latter two processes.

First, using a square-root specification instead of the Ornstein-Uhlenbeck process suggested by Drechsler and Yaron (2010) for the long-run mean reversion level of uncertainty has far-reaching economic consequences: the equity risk premium is increasing not only with short-run but also with long-run uncertainty, and the predictive power of the current price-dividend ratio for future excess returns increases and comes closer to empirically observed values.

Second, we distinguish between two sources of a time-varying uncertainty of cash-flows, stochastic diffusive variance and stochastic jump intensity. We find that for most effects caused by time-varying uncertainty, time-variation in the jump intensity is much more important than time-variation in diffusive volatility risk.

Third, the empirically observed low correlation between changes in the level and changes in the slope of the implied volatility smile for the S&P 500 index can only be matched with a model where jump intensity and conditional variance are locally uncorrelated.

Keywords: Asset pricing, Epstein-Zin preferences, variance risk premium, jump risk, stochastic volatility, level and slope of implied volatility smile

JEL: G12

1 Introduction and Motivation

Long-run risk (LRR) models introduced by Bansal and Yaron (2004) represent an important class of approaches to explain asset pricing puzzles. The general idea behind them is that consumption growth is affected by a slowly moving persistent, but small source of risk which over the long run becomes relevant and makes consumption growth very risky. Together with preferences of the Epstein-Zin type this generates an equity premium which is significantly higher than that in standard models with CRRA preferences and simpler dynamics for consumption growth.

Stochastic volatility of realized and expected consumption growth, also introduced first by Bansal and Yaron (2004), generates an additional equity risk premium and furthermore leads to stochastic volatility of equity returns. Since this seminal paper much work has been done to extend the basic LRR model.[1] Some papers have added a second volatility factor or a stochastic mean reversion level for volatility. Other papers also consider jumps in consumption or in the state variables with an intensity which is usually assumed to be proportional to diffusive variance. The inclusion of a stochastic volatility channel implies that expected excess returns are no longer constant. The richer dynamics lead to predictability of future excess returns and induce a variance risk premium.

In this paper we study an LRR model with stochastic volatility, a stochastic jump intensity, and a stochastic long-run level of diffusive and jump risk. While

[1]For extensions of the basic LRR model see, e.g., Eraker and Shaliastovich (2008), Bansal and Shaliastovich (2011), Drechsler (2009), Zhou and Zhu (2009), Bansal and Shaliastovich (2010), Benzoni, Collin-Dufresne, and Goldstein (2010), Drechsler and Yaron (2010), Wachter (2010), and Wang and Bidarkota (2010).

all three factors are related to the total level of uncertainty in the economy they nevertheless capture different aspects of it. We find that they have significantly different implications for equity returns, the variance risk premium, and option prices. The fine structure of time-variation in these uncertainty factors themselves matters for explaining the dynamics of risk premia and the predictability of future excess returns as well as the dynamics of the implied volatility smile.

The paper closest to ours is Drechsler and Yaron (2010) (DY hereafter). They extend the LRR model by adding a stochastic process for the formerly constant mean reversion level in the variance process and by including jumps in the state variables. The resulting model is able to explain the observed large and positive variance risk premium as well as the performance of this variance risk premium as a predictor for future excess returns on the dividend claim. It also matches the patterns of time variation both in the level and in the variance of these excess returns.

The most important general insight from the DY model is that to explain the stylized facts about different asset classes, like equity and its variance, both a stochastic diffusive variance and a stochastic jump intensity should be included as state variables. However, in DY both of these sources of risk are driven by the same state variable (and are thus locally perfectly correlated), since intensity is specified as an affine function of conditional variance.

This choice of the specification for the dynamics of stochastic intensities is indeed very common in the option pricing literature. For example, Bates (2000) proposes a two-factor stochastic volatility model and makes the jump intensity an affine function of the two volatility factors. Similarly, in her study on jump risk premia, Pan (2002) models the Poisson intensity as a linear function of the conditional variance of stock returns. While such a specification is both convenient from a modeling perspective and of course more general than assuming a constant intensity, its empirical validity has been questioned, e.g. in a recent empirical study by Santa-Clara and Yan (2010). They state that the "... estimated correlation between the increments of the diffusive volatility and jump intensity is quite low, at 0.17." (p. 444). This suggests that a model where the stochastic jump intensity is locally *perfectly* correlated with the conditional variance is potentially misspecified.

Figure 1 shows the time-series plots of monthly observations for the level and the slope of the implied volatility smile for options on the S&P 500 index for the period from 1996 to 2006. Already a first inspection of the graphs shows that the two quantities tend to move rather independently and are by no means perfectly correlated. The empirical correlation between changes in level and slope measured over the given period is 0.07. Indeed we find that a model where variances and intensities are tied together firmly is unable to reproduce this low correlation. We therefore introduce an additional stochastic factor independent of the conditional variance, and make the jump intensity proportional to a weighted average of these two factors.[2]

Relaxing the assumption of a perfect correlation between two random variables obviously reduces the overall amount of long-run uncertainty in the economy. To make the independent intensity case comparable to the old specification with a proportional intensity we recalibrate the model such that it again reproduces the asset pricing moments observed in the data. This recalibration yields a parametriza-

[2] In our numerical analysis of this more flexible model we only consider the case where intensity is locally uncorrelated with volatility.

tion of the model with a lower probability for intensity jumps than for variance jumps and higher average jump sizes for both. While only one source of uncertainty (diffusive variance or intensity) can jump at any given point in time, it will jump by more every time it does.

A second issue which we analyze in our paper is the choice of the type of stochastic process for the long-run mean of stochastic volatility. Here we modify the DY setup by introducing a square-root diffusion for the long-run mean of volatility instead of the original Ornstein-Uhlenbeck (OU) dynamics. Besides the more technical aspect that the square-root process guarantees non-negativity of the underlying variable, we prefer this choice for its superior economic implications. With an OU-specification expected returns do not vary with the level of long-run uncertainty in the economy, i.e., the market price of risk for this factor is constant due to its constant conditional volatility. However, an important economic implication of this fact, namely that the expected excess return increases if the current (short-run) level of uncertainty increases, but does not react to an increase in the (long-run) mean reversion level, seems rather counterintuitive. With our square-root specification the level of long-run uncertainty not only adds to the *level* of the equity risk premium, but also to its *variation over time*.

The main contributions of our paper are as follows. First, introducing the square-root specification for the long-run mean of conditional volatility yields the economically sensible result that expected excess returns not only depend on the current level of uncertainty but also on its long-run mean. Furthermore we achieve a significant improvement in the goodness of fit of the predictive regressions of future excess returns on the dividend claim on the current p-d ratio, taking us much closer to the values observed in the data than the DY model. For example, with a square-root process for the long-run mean of volatility the regression of five-year excess returns on the current p-d ratio exhibits an R^2 of roughly 18 percent, compared to 12 percent for the OU specification in DY, and to 23 percent in the data. The intuition behind this improvement in predictability is as follows: Future excess returns are predictable via the current p-d ratio, because both expected excess return and the p-d ratio are functions of the short-run uncertainty factors, 'stochastic volatility' and 'stochastic intensity', and these variables are themselves predictable. With a square-root specification for the long-run mean, excess returns also depend on this variable, while in the case of an OU process they do not. Putting it differently, in the square-root case the current value of long-run volatility has informational value for future excess returns via its impact on expected excess returns, while this is not true in the OU case considered in the DY paper.

Second, decoupling jump intensities and conditional variances allows us to separately analyze the role of diffusive and jump risks. With an intensity which is strictly proportional to the conditional variance, it is only the total amount of uncertainty in the economy which matters, but not its distribution across different types of risk factors. This would not be much of a problem, if separating the two risk sources just resulted in their contributions to the output of the model being roughly equal. We find, however, that rather the opposite is true. The variance risk premium is almost exclusively driven by jump risk (represented in our model by the independent factor driving the jump intensity). Variation in the expected excess returns on the dividend claim is driven by changes in all three uncertainty factors, where the jump intensity is most important, followed by current diffusive uncertainty, while the the long-run level of uncertainty comes in third.

Third, as indicated above, we are interested in matching the (potentially surprisingly) low correlation between changes in the level and in the slope of the implied volatility smile for indices like the S&P 500. When the intensity is proportional to the conditional variance, we never observed values for this correlation in our simulations below 0.5, in most cases it even exceeded 0.6. With locally uncorrelated processes for intensity and variance on the other hand, it goes down to about 0.07 and thus perfectly explains the value observed in the data.

The remainder of this paper is organized as follows. In Section 2, we introduce the model setup, and show how the equilibrium is derived for our model economy. In Section 3, we present the results of our analysis of asset pricing moments and predictive regressions. Section 4 concludes.

2 Model Setup

2.1 The Investor

We assume a continuous time endowment economy with a single perishable consumption good. The preferences of the representative agent are described by a recursive utility function introduced by Epstein and Zin (1989). In discrete time this utility function is given by:

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

We denote utility and consumption at time t by U_t and C_t, respectively. β is the investor's subjective discount factor. The parameter θ is defined as $\frac{1-\gamma}{1-\frac{1}{\psi}}$, where $\psi > 1$

and $\gamma > 1$ denote the elasticity of intertemporal substitution (EIS) and the relative risk aversion parameter, respectively. θ will be negative, given the restrictions on γ and ψ.

The recursive utility specification allows to choose the level of relative risk aversion and the EIS independently of each other. For the special case $\gamma = \frac{1}{\psi}$, the preferences in (1) collapse to the usual power utility specification where the EIS is just the inverse of relative risk aversion. In the case where $\gamma > \frac{1}{\psi}$ the investor has a preference for early resolution of uncertainty.

2.2 The Economy

Our economy is described by the processes for consumption and dividends, as well as for a set of state variables governing the future cash flow dynamics, all collected in the n-dimensional vector Y. The first and the last element of Y are the log of consumption (i.e. $c_t \equiv \ln C_t = e_c' Y_t$ with $e_c = (1, 0 \ldots 0)'$) and the log of dividends (i.e., $\delta_t \equiv \ln D_t = e_\delta' Y_t$ with $e_\delta = (0, \ldots 0, 1)'$), respectively. The remaining $n-2$ elements of Y are the state variables which, in our setup, are the long-run growth factor and a number of 'uncertainty factors' which will be described in detail below.

The dynamics of Y are given by the following system of stochastic differential equations:

$$dY_t = \mu(Y_t)dt + G(Y_t)dW_t + \xi_t dN_t \tag{2}$$

where W is an n-dimensional Brownian motion and N is an m-dimensional Poisson process with intensity $l(Y_t)$. The jump sizes are collected in the $n \times m$-matrix ξ_t with typical element $\xi_{t,ij}$ representing the change in variable i caused by a jump in Poisson process j at time t. The drift $\mu(Y_t)$, the variance-covariance matrix $G(Y_t)G(Y_t)'$ of the diffusion terms, and the vector of jump intensities $l(Y_t)$ is assumed to be an affine function of the state variables. This model setup closely follows the specification in Eraker and Shaliastovich (2008).

The reference model for our analysis is the continuous-time version of the specification in DY. There the authors assume that the dynamics of consumption and dividends are driven by a long-run growth risk factor and a stochastic volatility component, whose long-run mean is also time-varying. Both the long-run growth factor and the stochastic volatility are subject to jumps in two independent Poisson processes ($m = 2$). The jump intensities are driven exclusively by the stochastic volatility factor (of which they are an affine function), so that the two variables are locally perfectly correlated.

We extend this setup proposed by DY in two directions. First, we relax the assumption just described concerning jump intensities and conditional variance by including an additional stochastic factor governing the evolution of the jump intensity, possibly together with the conditional variance (in the sense of the intensity being a weighted average of the conditional variance and our new factor). This allows us to decompose the current amount of 'total risk' in the economy into the amount of jump risk (represented by the level of the jump intensity) and the quantity of diffusive uncertainty (represented by the current level of conditional diffusive variance). So in our model there might well be an increase in jump risk without a simultaneous increase in diffusion risk, and vice versa, while this would not be possible in the DY setup.

Second, we assume that the long-run mean of the stochastic volatility component follows a square-root instead of an OU-process, so that in our model this long-run mean stays positive with probability one. Besides representing a fundamentally important modeling improvement (guaranteeing that the long-run variance stays positive) this change has profound implications for theoretical asset pricing results as well as for the ability of the model to explain stylized facts in the data. Imposing a stochastic volatility for the long-run mean of economic uncertainty results in a direct dependence of expected excess returns on this factor, which is the key to the better performance of our model when it comes to predicting future excess returns by the current p-d ratio.

Put together the vector of cash flow and state variables is given by $Y = (c, x, \sigma^2, \alpha, \bar{\sigma}^2, \delta)'$. As introduced by Bansal and Yaron (2004) x is a small persistent component that captures long-run risks in the growth rate of consumption and dividends. Stochastic volatility is driven by two components. σ^2 controls the conditional diffusive variance of consumption, dividends, and the state variables, and $\bar{\sigma}^2$ is its long-run mean. The state variable α is the second component (besides σ^2) which drives the stochastic jump intensity.

The drift of Y at time t is given by

$$\mu(Y_t) = M + K\,Y_t = \begin{pmatrix} \mu_c \\ 0 \\ 0 \\ 0 \\ \kappa_{\bar{\sigma}}\bar{\bar{\sigma}}^2 \\ \mu_\delta \end{pmatrix} + \begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & -\kappa_x & 0 & 0 & 0 & 0 \\ 0 & 0 & -\kappa_\sigma & 0 & k_{\sigma,\bar{\sigma}} & 0 \\ 0 & 0 & 0 & -\kappa_\alpha & k_{\alpha,\bar{\sigma}} & 0 \\ 0 & 0 & 0 & 0 & -\kappa_{\bar{\sigma}} & 0 \\ 0 & \phi & 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} c_t \\ x_t \\ \sigma_t^2 \\ \alpha_t \\ \bar{\sigma}_t^2 \\ \delta_t \end{pmatrix}, \quad (3)$$

where, e.g., $k_{\sigma,\bar{\sigma}}$ denotes the loading of the drift of σ^2 on $\bar{\sigma}^2$.

The consumption and the dividend process have a drift of $\mu_c + x_t$ and $\mu_\delta + \phi x_t$, respectively. $\phi > 1$ plays the role of a 'leverage factor', since dividends, which are paid by (presumably levered) firms, have a larger exposure to changes in the economic environment than consumption. The long run risk factor x reverts to a long run mean of zero, so that we can interpret its current value as the deviation of expected consumption growth from its long run mean.

σ^2, α, and $\bar{\sigma}^2$ are all modeled as mean-reverting processes with the respective mean-reversion speeds κ_σ, κ_α, and $\kappa_{\bar{\sigma}}$. While $\bar{\sigma}^2$ mean-reverts to its long run mean $\bar{\bar{\sigma}}^2$, the other two state variables σ^2 and α are driven back towards the stochastic mean-reversion levels $\frac{k_{\sigma,\bar{\sigma}}}{\kappa_\sigma}\bar{\sigma}_t^2$ and $\frac{k_{\alpha,\bar{\sigma}}}{\kappa_\alpha}\bar{\sigma}_t^2$, respectively. In our numerical analysis below we choose $k_{\sigma,\bar{\sigma}}$ and $k_{\alpha,\bar{\sigma}}$ such that the long-run means of all three processes are equal to one.[3]

The conditional covariance matrix for the diffusion components, $G(Y_t)G(Y_t)'$, is given as

$$G(Y_t)G(Y_t)' = \begin{pmatrix} g_{cc} & 0 & 0 & 0 & 0 & g_{c\delta} \\ 0 & g_{xx} & 0 & 0 & 0 & 0 \\ 0 & 0 & g_{\sigma\sigma} & 0 & 0 & 0 \\ 0 & 0 & 0 & g_{\alpha\alpha} & 0 & 0 \\ 0 & 0 & 0 & 0 & g_{\bar{\sigma}\bar{\sigma}} & 0 \\ g_{c\delta} & 0 & 0 & 0 & 0 & g_{\delta\delta} \end{pmatrix} \tag{4}$$

[3]Note that, in contrast to standard specifications without jumps there is a difference here between the *mean reversion level* of a process and its *long-run mean.*

where

$$
\begin{aligned}
g_{jj} &= \sigma_j^2(w_j\sigma_t^2 + 1 - w_j) \qquad j \in \{c, x, \sigma, \delta\} \\
g_{\alpha\alpha} &= \sigma_\alpha^2(w_\alpha\alpha_t + 1 - w_\alpha) \\
g_{\bar{\sigma}\bar{\sigma}} &= \sigma_{\bar{\sigma}}^2(w_{\bar{\sigma}}\bar{\sigma}_t^2 + 1 - w_{\bar{\sigma}})
\end{aligned}
\tag{5}
$$

and

$$
g_{c\delta} = w_{c\delta}\sigma_c\sigma_\delta\left(\sqrt{w_c w_\delta}\sigma_t^2 + \sqrt{(1-w_c)(1-w_\delta)}\right).
$$

The entries on the main diagonal of the variance-covariance matrix in (4) represent the conditional variances of the elements of Y, and $g_{c\delta}$ is the conditional covariance between consumption growth and dividend growth. The weights w_j ($j \in \{c, x, \sigma, \delta\}$) determine to what degree the conditional variance of the respective variable is affected by the conditional variance process σ^2. w_α and $w_{\bar{\sigma}}$ have an analogous interpretation. The weights are restricted to be between zero and one, and in our parametrization we set $w_c = 0.5$, $w_\delta = 0.125$, and $w_x = 1$.[4] Setting $w_\sigma = w_\alpha = 1$, as we will do below, implies that both σ^2 and α follow classical square-root processes (albeit with jumps). The choice of $w_{\bar{\sigma}}$ determines a key feature of the model, namely if $\bar{\sigma}$ follows an OU process ($w_{\bar{\sigma}} = 0$), or if this variable is governed by a square-root process ($w_{\bar{\sigma}} = 1$). Since we choose the processes for σ^2, α, and $\bar{\sigma}^2$ such that their long-run means are equal to one, σ_c, σ_δ, σ_x, σ_σ, σ_α, and $\sigma_{\bar{\sigma}}$ are the average diffusion volatilities of consumption growth, dividend growth, and the state variables.

In order to stay within the affine model class we impose a zero correlation between any pair of square-root processes. For reasons of parsimony, we introduce a non-zero correlation only between the shocks to the dividend and the consump-

[4]The numerical values of all parameters can be found in Table 1.

tion process in our calibrations. The specification of the covariance term between consumption and dividends follows the one suggested in DY, and we set $w_{c\delta} = 0.2$.

For future reference, we will also write the variance covariance matrix as the sum of a matrix of constants and additional summands containing matrices representing the impact of state variables:

$$G(Y_t)G(Y_t)' = h + H_\sigma \sigma_t^2 + H_\alpha \alpha_t + H_{\bar{\sigma}} \bar{\sigma}_t^2 \tag{6}$$

where h, H_σ, H_α, and $H_{\bar{\sigma}}$ are assumed to be symmetric and positive definite.

The consequences of the choice of $w_{\bar{\sigma}}$ in (5) become immediately visible in (6). In the OU case the variance of $\bar{\sigma}^2$ is given by $h(5,5) = \sigma_{\bar{\sigma}}^2$, while $H_{\bar{\sigma}}$ is equal to the zero matrix so that the current level of $\bar{\sigma}^2$ does not have an impact on the conditional variance-covariance matrix. In the square-root case $H_{\bar{\sigma}}$ will have the non-zero element $H_{\bar{\sigma}}(5,5) = \sigma_{\bar{\sigma}}^2$, and $h(5,5) = 0$. This difference between the two specifications will become important below, when we analyze expected excess returns on risky assets.

Finally, we turn to the jump components. In our model, the long-run growth factor x, the stochastic variance σ^2, and the intensity factor α are subject to jumps, so that the Poisson process N in Equation (2) is three-dimensional. As usual we assume that the there are no simultaneous jumps in the different Poisson processes, i.e. there are no joint jumps in the state variables. Concerning the jump size distributions we assume that jumps in x are normal. Jumps in σ^2 and α have to be positive, and we rely on exponential distributions here. The intensities of the jumps $l(Y_t) \in \mathbb{R}^3$ are modeled as an affine function of the state variables, i.e.

$$l(Y_t) = l_0 + l_1 Y_t,$$

with $l_0 \in \mathbb{R}^3$ and $l_1 \in \mathbb{R}^{3\times 6}$. To keep the model tractable we assume that the jump intensities only depend on σ_t^2 and α_t. We can thus rewrite the elements l_{it} of the intensity vector at time t as

$$l_{it} = \tilde{l}_{i,0} + \tilde{l}_{i,1}\left[(1-\varphi_i)\,\sigma_t^2 + \varphi_i \alpha_t\right] \qquad i = x, \sigma, \alpha. \tag{7}$$

Setting φ_i equal to zero or one in (7) results in a jump intensity which is either perfectly correlated with the stochastic diffusion variance σ^2, as in DY, or moves independently of it. Setting $0 < \varphi_i < 1$ generates an intermediate case, where the intensity is no longer perfectly correlated with either σ^2 or α. In our parametrization we set $\tilde{l}_{i,0} = 0$, $\tilde{l}_{i,1} > 0$, and $\varphi_i \equiv \varphi$ for all i.

2.3 The Equilibrium

In this section we summarize the equilibrium solution for our model. A detailed derivation can be found in Eraker and Shaliastovich (2008).

The Euler equation reads

$$E_t\left[\frac{M_{t+h}}{M_t} e^{\int_t^{t+h} d\ln R_{i,s}}\right] \;=\; 1 \tag{8}$$

where M and $d\ln R_i$ denote the stochastic discount factor and the log return on an asset or portfolio i, respectively. The dynamics of the pricing kernel in continuous time are given by

$$d\ln(M_t) = \theta\ln(\beta)dt - \frac{\theta}{\psi}\,dc_t - (1-\theta)d\ln R_{c,t}, \tag{9}$$

where $d\ln R_{c,t}$ is the log-return on the consumption claim at time t. To solve for this log-return, Eraker and Shaliastovich (2008) approximate it via log-linearization following Campbell and Shiller (1988):

$$d\ln(R_{c,t}) = k_0 dt + k_1 dv_t - (1-k_1)v_t dt + dc_t. \tag{10}$$

v_t is the log wealth-consumption ratio, and k_0 and $k_1 \in (0,1)$ are called 'linearization constants', which depend on the average log wealth-consumption ratio (details are given in Appendix A.1).

v_t is assumed to be an affine function of Y_t, i.e.

$$v_t = A_0 + A_1' Y_t, \tag{11}$$

where A_0 and A_1 together with k_0 and k_1 solve a non-linear system of equations given in Appendix A.1. With Epstein-Zin preferences and $\gamma > 1/\psi$ one obtains $A_{1x} < 0$, where A_{1x} is the component related to x in the vector A_1. Furthermore, $A_{1\sigma}$, $A_{1\alpha}$, and $A_{1\bar{\sigma}}$ are all negative (see again Appendix A.1). Plugging the above expression for v_t into (10) yields

$$d\ln R_{c,t} \;=\; [k_0 - (1-k_1)A_0 - (1-k_1)A_1' Y_t]\,dt + (k_1 A_1 + e_c)' dY_t.$$

Together with Equation (9), this yields the dynamics of the pricing kernel, from which we can then derive the risk-free rate and the market prices of risk. The risk-free rate is $r_t = r_0 + r_1' Y_t$, where the coefficients r_0 and r_1 are given in Appendix A.2. The market prices of risk are collected in the vector $\Lambda \in \mathbb{R}^6$ with

$$\Lambda \;=\; \gamma e_c + (1-\theta)k_1 A_1.$$

Jumps in x are assumed to be normally distributed with mean $\mu_{\xi^x} \leq 0$ and variance $\sigma^2_{\xi^x}$. Their intensity under the risk-neutral measure $\mathbb{Q}$ is given by the intensity under the physical measure $\mathbb{P}$, multiplied by the term $\exp\{-(1-\theta)k_1 A_{1x}\mu_{\xi^x} + 0.5(1-$

$\theta)^2 k_1^2 A_{1x}^2 \sigma_{\xi^x}^2\} > 1$. The mean jump size under $\mathbb{Q}$ is equal to $\mu_{\xi^x} - (1-\theta)k_1 A_{1x}\sigma_{\xi^x}^2$, the variance does not change.[5] With $\theta < 0$ and $A_{1x} > 0$ jumps in the long-run risk factor x are thus more negative and more frequent under $\mathbb{Q}$ than under $\mathbb{P}$.

Jumps in the conditional variance σ^2 and in the intensity factor α exhibit similar characteristics. Their $\mathbb{Q}$-intensities are equal to the $\mathbb{P}$-intensities multiplied by the terms $(1+(1-\theta)k_1 A_{1\sigma}\mu_V)^{-1} > 1$ and $(1+(1-\theta)k_1 A_{1\alpha}\mu_\alpha)^{-1} > 1$, respectively. The mean jump sizes under $\mathbb{Q}$ are $(1+(1-\theta)k_1 A_{1\sigma}\mu_V)^{-1}\mu_V$ and $(1+(1-\theta)k_1 A_{1\alpha}\mu_\alpha)^{-1}\mu_\alpha$, which are both greater than their counterparts under $\mathbb{P}$, μ_V and μ_α, given that $A_{1\sigma}$ and $A_{1\alpha}$ are negative.

Once we know the pricing kernel, we can price any other claim, given its future payoffs. The log return on the dividend claim can be approximated like the log return on the consumption claim:

$$d\ln R_{\delta,t} \quad = \quad k_{d0}dt + k_{d1}dv_{\delta,t} - (1-k_{d1})v_{\delta,t}dt + d\delta_t \tag{12}$$

where $v_{\delta,t}$ denotes the log p-d ratio, again assumed to be affine in Y_t:[6]

$$v_{\delta,t} \quad = \quad A_{d0} + A_{d1}'Y_t.$$

Plugging the expression for $v_{\delta,t}$ into the approximation in (12) yields

$$d\ln R_{d,t} = [k_{d0} - (1-k_{d1})A_{d0} - (1-k_{d1})A_{d1}'Y_t]\,dt + (k_{d1}A_{d1} + e_\delta)'dY_t. \tag{13}$$

The exposure of the return to the risk factors is thus given by $\omega \equiv k_{d1}A_{d1} + e_\delta$. The expected excess return then follows from these exposures and the market prices of

[5]For details of the computation of risk-neutral jump intensities and mean jump sizes we again refer to Eraker and Shaliastovich (2008).

[6]Details on the computation of the coefficients A_{d0} and A_{d1} and the linearization constants k_{d0} and k_{d1} are given in Appendix A.3.

risk. For the dividend claim, the expected excess return (i.e, the equity premium) is

$$
\begin{aligned}
&\frac{E[d\ln R_{d,t}]}{dt} - (r_0 + r_1' Y_t) \qquad (14)\\
&= \omega' \left(h + H_\sigma \sigma_t^2 + H_\alpha \alpha_t + H_{\bar\sigma} \bar\sigma_t^2\right) \Lambda \\
&\quad - 0.5\, \omega' \left(h + H_\sigma \sigma_t^2 + H_\alpha \alpha_t + H_{\bar\sigma} \bar\sigma_t^2\right) \omega \\
&\quad + \sum_{i=x,\sigma^2,\alpha} l_{i,t} \left(k_{d1} A_{d1,i} E\left[\xi^i\right] + E\left[e^{-(1-\theta)k_1 A_{1i}\xi^i}\right] - E\left[e^{(k_{d1}A_{d1,i}-(1-\theta)k_1 A_{1i})\xi^i}\right]\right)
\end{aligned}
$$

The first term on the right-hand side gives the premia for diffusive risk, while the second represents a Jensen correction term. The three terms of the sum are the premia for jump risk in x, σ^2, and α.

As with CRRA, there will be a premium for consumption diffusion risk equal to the product of relative risk aversion γ and the covariance of dividend risk with consumption risk, $g_{c\delta}$ (see Equation (4)). With Epstein-Zin preferences, however, state variables are also priced. Given $\gamma > \frac{1}{\psi}$ the dependence of the p-d ratio on the state variables is in line with intuition, since it increases with x and decreases with the uncertainty factors. The premium on the long-run growth rate x is then positive, since both the exposure of the stock price and the market price of risk are positive. Furthermore the premia on the uncertainty factors σ^2, α, and $\bar\sigma^2$ are also very likely to be positive, since the negative exposures are multiplied by negative market prices of risk. Put together, the premia on all diffusive risk factors are positive.

To assess whether these premia are constant or vary over time, note that they are proportional to the local diffusion (co)variances of the dividend and of the state variables. As pointed out above it is especially the type of stochastic process for $\bar\sigma$ which is important here. The expected excess return thus increases in σ^2 and — when the jump intensities and the conditional variance are not perfectly correlated

— also in α. When $\bar{\sigma}^2$ is modeled as an OU process like in the DY model, the expected excess return on the dividend claim will not depend on the *level* of this variable (its contribution will only enter the matrix of constants h), whereas with our square-root specification changes in $\bar{\sigma}$ will have an immediate effect on the equity risk premium.

The premia for jumps in x, σ and α depend on the exposures $k_{d1}A_{d1}$ of the p-d ratio to the state variables, on the market prices of risk, and on the jump intensity. As discussed above jumps in x are more frequent and more negative under the risk-neutral than under the true measure. Together with the fact that the stock price is increasing in x, this will result in a positive contribution of the premium for x to the total risk premium on the dividend claim. A similar argument shows that the premia for jumps in σ^2 and α have to be positive, since the p-d ratio loads negatively on these two variables, their jump intensities are higher under $\mathbb{Q}$ than under $\mathbb{P}$, and the mean jump size is greater under the risk-neutral than under the physical measure. Furthermore, the jump part of the equity risk premium is proportional to the jump intensity, which is an affine function of α and σ^2.

2.4 Variance Risk and Variance Risk Premium

In our analysis of the model we will also run predictive regressions of future excess returns on the dividend claim on the current variance risk premium. The variance risk premium is defined as the difference of the expected quadratic variation of the return over a time interval τ under the risk-neutral measure $\mathbb{Q}$ and the physical measure $\mathbb{P}$, i.e. it is given by the expression

$$E_t^{\mathbb{Q}}\left[\int_t^{t+\tau}(d\ln R_{d,s})^2\right]-E_t^{\mathbb{P}}\left[\int_t^{t+\tau}(d\ln R_{d,s})^2\right]. \tag{15}$$

Note that a positive variance risk premium implies that the investor is actually willing to accept a negative average return on an asset with positive exposure to variance risk.[7]

To compute the variance risk premium, we start with the squared local return on the dividend claim. From (13) one obtains

$$\begin{aligned}(d\ln R_{d,t})^2 &= (k_{d1}A_{d1}+e_\delta)'\,(dY_t dY_t')\,(k_{d1}A_{d1}+e_\delta)\\ &= (k_{d1}A_{d1}+e_\delta)'\,[G(Y_t)G(Y_t)'dt+\xi_t\text{diag}(dN_t)\xi_t']\,(k_{d1}A_{d1}+e_\delta),\end{aligned}\tag{16}$$

which comprises the variance due to diffusion risk and the contribution of jumps to the squared return. Taking expectations under the physical measure $\mathbb{P}$ yields

$$\begin{aligned}&E_t^{\mathbb{P}}\left[(d\ln R_{d,t})^2\right]\\ &= (k_{d1}A_{d1}+e_\delta)'\left(h+H_\sigma\sigma_t^2+H_\alpha\alpha_t+H_{\bar\sigma}\bar\sigma_t^2\right)(k_{d1}A_{d1}+e_\delta)\,dt\\ &\quad+\sum_{i=x,\sigma^2,\alpha} l_{i,t}(k_{d1}A_{d1,i})^2E_t^{\mathbb{P}}\left[(\xi_t^i)^2\right]dt.\end{aligned}\tag{17}$$

An analogous formula gives the expectation under the risk-neutral measure $\mathbb{Q}$.

[7]Examples for papers dealing with the variance risk premium are Bollerslev, Sizova, and Tauchen (2011), Bollerslev, Tauchen, and Zhou (2010), Carr and Wu (2009), Egloff, Leippold, and Wu (2010), and Todorov (2010).

By integrating (17) from t to $t+\tau$ one obtains the $\mathbb{P}$-expectation in (15):

$$
\begin{aligned}
& E_t^{\mathbb{P}}\left[\int_t^{t+\tau} (d\ln R_{d,s})^2\right] \\
&= (k_{d1}A_{d1}+e_\delta)'\left(h\,\tau + H_\sigma E_t^{\mathbb{P}}\left[\int_t^{t+\tau}\sigma_s^2 ds\right] + H_\alpha E_t^{\mathbb{P}}\left[\int_t^{t+\tau}\alpha_s ds\right] + H_{\bar\sigma} E_t^{\mathbb{P}}\left[\int_t^{t+\tau}\bar\sigma_s^2 ds\right]\right) \\
&\qquad \cdot (k_{d1}A_{d1}+e_\delta) \\
&\quad + \sum_{i=x,\sigma^2,\alpha} (k_{d1}A_{d1,i})^2 E_t^{\mathbb{P}}\left[(\xi_t^i)^2\right] E_t^{\mathbb{P}}\left[\int_t^{t+\tau} l_{i,s} ds\right], \qquad (18)
\end{aligned}
$$

where

$$
E_t^{\mathbb{P}}\left[\int_t^{t+\tau} l_{i,s} ds\right] = \tilde{l}_{i,0}\,\tau + \tilde{l}_{i,1}\left(\varphi_i E_t^{\mathbb{P}}\left[\int_t^{t+\tau}\alpha_s ds\right] + (1-\varphi_i)E_t^{\mathbb{P}}\left[\int_t^{t+\tau}\sigma_s^2 ds\right]\right).
$$

Again, the corresponding quantity under $\mathbb{Q}$ can be computed in just the same fashion. Closed form solutions to the integrals in (18) can be found in Appendix A.4. Taking the difference of the two expectations finally yields an expression for the variance risk premium as defined in (15). Since the variation of the stock return arises because the dividend and the p-d ratio are exposed to diffusive and jump risk, the total variance risk premium can consequently be decomposed into the premia for these two types of risk. In our model the local diffusion variance is stochastic (represented by the first term in Equation (17)) and depends on the state variables σ^2, α, and $\bar\sigma^2$ which all have different dynamics under $\mathbb{P}$ and $\mathbb{Q}$, so that that there will be a non-zero premium for diffusive variance risk. Since all of these 'uncertainty state variables' have a mean which is greater under $\mathbb{Q}$ than under $\mathbb{P}$, their contribution to the variance risk premium is positive.

The jump-related part of the variance risk premium is driven by the differences in the average jump sizes and intensities under $\mathbb{P}$ and $\mathbb{Q}$. Since jumps are more extreme and more frequent under the risk-neutral than under the physical measure, this part of the premium is also positive.

3 Numerical Results

3.1 General Approach

In our analysis we compare three models. The first is the DY model with an OU process for the long-run mean of variance $\bar{\sigma}^2$ and a jump intensity which is perfectly correlated with the stochastic variance σ^2. In the model which we call 'Extension 1' we replace the OU process for $\bar{\sigma}^2$ by a square-root process, while 'Extension 2' refers to the model where we additionally introduce the autonomous intensity process α and then recalibrate the model.

Since our model extends the DY setup, we use their parameters as a starting point for our calibration. Note that DY specify their model in discrete time and report the parameters for a monthly frequency, while we use a continuous time model and work with annual numbers. We thus first rewrite our model in discrete time and match the parameters with the values given in DY. Transforming the model back to continuous time then gives what we call the 'base case', i.e., the continuous-time formulation of the DY model. As in DY, the preference parameters are set to $\gamma = 10$, $\psi = 2$, and $\beta = 0.9881$. All the parameters for the three alternative models are shown in Table 1.

We rely on Monte Carlo simulation to compute cash flow and asset pricing moments as well as the statistics for coefficients and R^2-values in the predictive regressions. Our approach follows DY in that the simulation results are based on 1,000 runs over 77 years each, where the dynamics of all variables are discretized with a time step of one day. Nevertheless, due to the Euler discretization which we apply to the system of stochastic differential equations, negative values can occur for σ, α, and $\bar{\sigma}$. If this happens these negative values are replaced by a small positive number.[8]

The model output has to be compared to data at some point. The respective first columns in Tables 2, 6, 7, and 8 show the empirical values for cash flow and asset pricing moments and for the predictive regressions using the p-d ratio and the variance risk premium as right-hand side variables.

We use the S&P 500 index as the empirical proxy for the dividend claim.[9] Consumption data are taken from the National Income and Product Accounts (NIPA) tables of the Bureau of Economic Analysis (BEA).[10] Consumption is defined as per-capita consumption of non-durable goods and services. These figures have to be converted into real terms, where we use the personal consumption expenditure (PCE) deflator from the BEA. Dividends are computed using the difference between returns on a stock market index with and without dividends, as in Bansal, Dittmar, and Lundblad (2005). The cash flow and asset pricing moments (except the variance premium) are computed over the period from 1930 to 2006. The statistics for the variance risk premium (computed as in DY as the difference between the expectations of the integrated variance of the dividend claim over the next 30 days under the risk-neutral and the physical measure, respectively) are based on the period from January 1990 to March 2007.

[8]We replace a value below $\epsilon = 2^{-52}$ (which is the floating point precision in Matlab) by 2ϵ minus this value, i.e. we reflect along the value of ϵ.

[9]The empirical characteristics of the CRSP index are very similar to those of the S&P 500, so they are not shown.

[10]See http://www.bea.gov/national/nipaweb/Index.asp and http://www.bea.gov/national/nipaweb/DownSS2.asp)

In our analysis of the model we will also consider the level and the slope of the implied volatility smile of options on the S&P 500 index. The statistics for these characteristics are based on a sample of volatility surfaces provided by OptionMetrics for the same period as the one for the variance risk premium. We define the level of the implied volatility as the implied Black-Scholes volatility of a 1-month at-the-money option, and the slope as the difference between the implied volatilities of an option with a moneyness (defined as the option's strike price divided by the current index level) of 0.975 and the at-the-money option.

When comparing different model specifications it must be made sure that the models do not already differ with respect to the fundamental cash flow dynamics. Table 2 shows the moments of consumption and dividend growth for the different versions of our model. It is obvious that these moments are very similar across all three alternative models, so that they could not be readily distinguished based just on the time-series characteristics of the cash flow variables. Furthermore, the descriptive statistics from the model are also generally close to the data. The only exception is the correlation between log dividend growth and log consumption growth, which is higher in the data than in the model. Other than that basically all of the empirical moments are located between the 5th and the 95th percentile of the simulated model data.

3.2 Asset Pricing

3.2.1 DY

The model with an OU process for $\bar{\sigma}^2$ and an intensity proportional to σ^2 is the continuous-time version of DY. We will now provide a detailed discussion of the results for this specification and focus on the respective roles of jump and diffusion risk. This gives us a benchmark against which we will later assess the economic implications of the proposed modifications in the setup.

The first thing we are interested in is how the state variables impact quantities like the wealth-consumption ratio, the p-d ratio, and expected excess returns. The wealth-consumption ratio at time t, v_t, is given as $v_t = A_0 + A_1' Y_t$ with $A_{1c} = A_{1\delta} = 0$. The coefficients for the four state variables x, σ^2, α, and $\bar{\sigma}^2$ are shown in Panel A of Table 3. The numbers shown for the DY model here seem to suggest that v is mainly driven by the long-run growth factor x with a coefficient of $A_{1x} \approx 1.7$, while the uncertainty variables σ^2 and $\bar{\sigma}^2$ are less important with coefficients equal to or smaller than 0.03 in absolute value. When taking the typical volatility levels of the different variables into account (see Table 1) we find, however, that a one standard deviation change in any of them would have roughly the same effect on the wealth-consumption ratio.[11]

For the p-d ratio (Panel B) we obtain a similar result, with the numerical

[11]For the long-run growth rate x a one standard deviation change due to diffusion risk causes a change in the wealth-consumption ratio of $A_{1x}\sigma_x = 0.01485$. For a jump in x this number is $A_{1x}\sigma(\xi^x) = 0.01560$, while for σ^2 the corresponding numbers are -0.00504 for a diffusive and -0.01061 for a jump-driven change. For $\bar{\sigma}^2$ the impact of a typical change in the diffusive risk factor is -0.01179.

values of the effects of changes in the state variables being much larger than in the case of the wealth-consumption ratio due to the fact that the dividend dynamics are levered relative to consumption.[12]

It is interesting to compare the impact of σ^2 and $\bar{\sigma}^2$, the two factors determining the evolution of diffusive uncertainty in the economy in the short and in the long run, on the various quantities shown in Table 3. We find that a given change in $\bar{\sigma}^2$ has a much larger impact on the wealth-consumption ratio (by a factor of around 8) than an equally large change in σ^2. The main reason for this more pronounced sensitivity is that, with a low speed of mean reversion of $\kappa_{\bar{\sigma}} = 0.018$, $\bar{\sigma}^2$ is much more persistent than σ^2 ($\kappa_\sigma = 3.6$), and an investor with Epstein-Zin utility is particularly concerned about long-run risk factors. Equation (A.2) in Appendix A.1 shows that for $k_1 = 1$ the sensitivities of the wealth-consumption ratio to σ^2 and to $\bar{\sigma}^2$ would be linked via $A_{1,\bar{\sigma}} = \frac{k_{\sigma,\bar{\sigma}}}{\kappa_{\bar{\sigma}}} A_{1,\sigma}$. The importance of $\bar{\sigma}^2$ is thus increasing in its persistence (i.e., in $\kappa_{\bar{\sigma}}^{-1}$) and in its impact on σ^2 (i.e., in $k_{\sigma,\bar{\sigma}}$).[13]

The impact of σ^2 on the wealth-consumption ratio is the greater the more

[12] The numbers are 0.0543 for the change in the p-d ratio due to a one standard deviation shock in the diffusive part of x and 0.0571 for a jump. For σ^2 the effect of a one standard deviation change due to diffusion risk is -0.0343, for a jump it is -0.0722, while for $\bar{\sigma}^2$ a one standard deviation shock of the diffusive factor moves the p-d ratio by $= -0.0659$.

[13] Note that without jumps in σ^2, $k_{\sigma,\bar{\sigma}} = \kappa_\sigma$, so that the relative importance of the two 'volatility factors' would be determined solely by the ratio of their respective persistence levels, and the more persistent factor would be more important. With jumps in σ^2, $k_{\sigma,\bar{\sigma}}$ is decreasing in the intensity and average size of these jumps (jumps have a positive contribution to the long-run mean, and if we want to keep the long-run mean of σ^2 at a certain value, we have to reduce the mean-reversion level). This smaller impact of $\bar{\sigma}^2$ on σ^2 consequently decreases the overall impact of this variable on the wealth-consumption ratio.

pronounced its persistence, as we show in Appendix A.1. Note here that the current uncertainty σ^2 influences the wealth-consumption ratio via several channels simultaneously. First there is the direct impact on the volatility of consumption growth via the diffusion component in the dynamics for c. Second, σ^2 affects the dynamics of x via the diffusion component, and as the driving force behind the jump intensity dynamics. Third, it drives its own volatility and jump intensity.

The relevance of the different channels can be assessed by comparing the value of $A_{1\sigma}$ in the full model to the situation when the respective channel is closed. It turns out that a time-varying intensity for jumps in the state variables is very important in general, whereas time-varying diffusion risk mainly matters for x.[14] This difference between jump and diffusion risk is the main reason why we introduce the more general specification where the intensity and the diffusion channel are separated. The results for the p-d ratio (Panel B) are qualitatively the same as those for the wealth-consumption ratio. Numerically, the effects are larger, of course, due to the inherent leverage of dividends relative to consumption.

Next we turn to the expected excess (log) returns on the dividend claim. The coefficients of the state variables are shown in Panel C of Table 3). Table 4 then presents a detailed decomposition of the expected excess return on the dividend claim into its components. For each of the models considered in our analysis the rows of this table show the variables delivering contributions to the expected excess return, either through diffusive or through jump components. For example, the total

[14] Additional calculations show that $A_{1\sigma}$ drops from its value of -0.00416 in Panel A of Table 3 to -0.00319 if we set $w_c = 0$, to -0.001751 for $w_x = 0$, to -0.00192 if we close down its impact onto the intensity of jumps in x and assume a constant intensity instead, and finally to -0.003794 for $w_\sigma = 0$ and to -0.001224 if we close down its impact onto its own intensity.

equity risk premium in the DY model is 6.05%. As shown in the upper panel of Table 4 the share of diffusive risk in the total equity premium is 2.45%, for jump risk this number is 3.60%, and one can see from the table how the two summands are in turn generated by the different risk factors in the model. The exposure to x earns on average 1.31% for diffusion risk and 1.20% for jump risk. For σ^2 the corresponding numbers are 0.26% and 2.40%, which reflects the fact that most of the variation in σ^2 is due to jumps. The long-run mean $\bar{\sigma}^2$ does not jump, so it can only contribute through diffusive risks (on average 1.19%). The part of the risk premium which is due to the correlation of dividends with consumption can be neglected, due to a Jensen correction term it even turns out to be negative (-0.31).

Finally, Tables 3 and 4 show that the expected excess return is increasing in σ^2. Looking at the column for σ^2, one can see that its total contribution is 5.208% times its stationary mean of 1, yielding the corresponding number 0.05208 in Table 3 as the coefficient for σ^2 in the equity premium in the DY model. To get the intuition note that the contribution of the risk factors to the equity risk premium is proportional to their current exposure to diffusion and jump risk (see equation (14)). The premium for diffusion risk, e.g., is proportional to σ^2 for x and for σ^2 (since $w_x = w_\sigma = 1$) while it is constant for $\bar{\sigma}^2$ due to the OU-specification. The premium for jumps in x and σ^2 is proportional to the jump intensity and thus again to σ^2.

As can be seen by comparing the values from the data (first column in Table 6) to those generated by the DY model (next four columns in Table 6), the model is close to the data with respect to most quantities of interest in terms of empirical fit. Exceptions are the level of the log p-d ratio, which is on the low side relative to the data, and its volatility, which is roughly twice as high in the data as in the model.

The volatility of the risk-free rate seems way too low compared to the data. This issue has already been discussed in Beeler and Campbell (2009) who argue that the low model volatility is simply the consequence of not incorporating inflation into the model. So when the (real) risk-free rate in the data is computed as the difference between a nominal rate and realized inflation, it automatically becomes more volatile than in a model which does not include inflation from the start.

We are also interested in the properties of the variance risk premium. The premium is calculated for a horizon of 30 days and then annualized. Analogous to Table 4, Table 5 shows the contributions of the cash flow and the state variables, again divided into a diffusive and a jump part.

The variance risk premium represents a compensation to investors for bearing uncertainty about future return variance. As becomes immediately obvious from the numbers in the table the jump components are much more important than diffusive risks. Of the total variance risk premium of 96 basis points almost 93 come from jump components, so that the diffusive contribution is basically negligible. As can be seen from Table 5 the premium for jumps in σ^2 is much larger (87.7 bps) than the one for jumps in x (5 bps). The intuition here is that jumps in x have zero mean, while jumps in σ^2 are positive. Jumps in σ^2 are thus *always* bad news for the investor, while jumps in x are symmetric around zero and will thus represent a 'good' innovation in 50 percent of the cases. Risk aversion still introduces an asymmetry between positive and negative shocks of the same size, but nevertheless the effect for x is much weaker, and so is the associated premium. A comparison with the empirical properties of the variance risk premium shows that the DY model tends to generate premia which are slightly too low, while the volatility of the premium and its higher moments are matched quite well.

Finally, we take a look at the performance of the DY model with respect to the characteristics of the implied volatility smile. While the level generated by the model is close to the empirically observed value, this is not the case for the slope. However, as we will see below, it is in general difficult for models of the type discussed in this paper to match the data here. The most striking discrepancy between the data and the model appears with respect to the correlation between the changes in the level and the slope of the smile. The value of less than 0.07 in the data shows that the two characteristics of the smile move basically independently, while the DY model indicates that they are strongly positively correlated. Of course, this is a direct consequence of the fact that the jump intensity is locally perfectly correlated with the conditional variance. As discussed in Yan (2011) the slope of the smile is a good proxy for jump risk, while the level of the smile for short-term options is mostly driven by diffusive volatility, so that with the jump intensity being proportional to the amount of diffusive risk, level and slope simply have to co-move strongly, which is not in accordance with the data.

3.2.2 Extension 1

This first extension modifies the DY setup by replacing the OU-specification for $\bar{\sigma}^2$ by a square-root process. The overall picture is that the total amount of risk in the economy increases, thus prices decrease and risk premia increase. For example, Table 3 shows that the log p-d ratio decreases from an average value of 2.95 for the DY case to on average 2.83 for the new specification. Similarly, the average expected excess return on the dividend claim goes up from 6.05% to 6.77%. The intuition behind this increase in overall risk is that the volatility of $\bar{\sigma}^2$ is no longer deterministic, but now varies with its own level, and this additional channel makes this factor more important. Its impact on the wealth-consumption and the p-d ratio increases in absolute values from -0.034 to -0.044 and from -0.19 to -0.26, respectively (Table 3), while the coefficients for the other variables remain more or less unchanged. In sum this causes the valuation ratios to decrease. In line with this intuition Table 4 shows that the risk premium due to the exposure to $\bar{\sigma}^2$ significantly increases from around 1.2% to (on average) 2.1% for the dividend claim.

Compared to the benchmark setup of DY the expected excess return on the dividend claim now depends on $\bar{\sigma}^2$, whereas its loading was zero in the OU case. As pointed out above the level of the equity risk premium depends on the variance due to diffusion risk and on the jump intensities (see Equation (14)). Switching from an OU to a square-root specification for $\bar{\sigma}^2$ implies that the formerly constant variance of $\bar{\sigma}^2$ is now proportional to it. Therefore both the risk-free rate and the expected excess returns on the dividend claim explicitly depend on this variable. In numbers, the contribution of $\bar{\sigma}^2$ to the equity risk premium increases from 1.2% to on average 2.1%. The other variables contribute about as much to the equity risk premium as they did before.

The economic consequences of the specification change are that a positive shock to $\bar{\sigma}^2$ now has three effects. The first two, namely that this immediately increases the overall level of long-run uncertainty in the economy and that it subsequently leads to an increase in the level of diffusion and jump risk via a higher mean-reversion level for σ^2, were already present in the DY case. The third effect is new: conditional expected excess returns are now also a function of the level of this long-run uncertainty, so that the investor instantaneously requires a higher expected return when the expected future level of uncertainty increases, i.e. when the investor forecasts worse developments of the economy in the future. In the benchmark setup, on the other hand, the expected excess return only increases when the short-run uncertainty factor increases. Comparing the model to the data (see the columns labeled 'Extension 1' in Table 6) shows that the volatility of the p-d ratio, although still somewhat low, is now higher and thus closer to the data.

Finally, we take a look at the variance risk premium. Table 5 shows that it is lower for the square-root specification than for the DY model (87 basis points vs. 96 basis points). The diffusive part of the premium increases marginally, while the jump part decreases substantially, at least in relative terms, from 92 basis points to 83 basis points. Due to the relatively short horizon of 30 days the variance risk premium is mainly paid for jump risk, i.e. for jumps in the return of the dividend claim due to jumps in x and σ^2, and the premium for jumps in σ^2 is still by far the largest component of the variance risk premium. With the additional channel for $\bar{\sigma}^2$ this factor has become more important than in DY, while the impact of short-run uncertainty σ^2 has decreased slightly. Consequently, the variance risk premium, which is mainly a premium for jumps in σ^2, also decreases.

As before, the last set of stylized facts we look at are those concerning the smile. The main point one has to note here is that the correlation between the changes in the level and the slope of the smile is still far too high compared to the data. The numerical values for level, slope, and this correlation are ver similar to those for the DY model. This shows that the specification of the dynamics for $\bar{\sigma}^2$ is not the key to matching option-related moments in the data.

3.2.3 Extension 2

Our second extension relaxes the assumption that the jump intensity l is perfectly correlated with the stochastic variance σ^2. We introduce an autonomous intensity process α which is independent of σ^2.

Since the separation of the intensity from the variance process reduces the overall amount of uncertainty in the economy, so that we have to recalibrate the model. The parameters are shown in Table 1. The most important characteristic of this recalibration is that now the average intensity for jumps in the jump intensity itself (represented by α) is 0.1, compared to an average intensity of 0.8 for jumps in σ^2. Furthermore, the average jump size in α and σ^2 is now 5.8, compared to 2.55 before. More than doubling the average jump size is motivated by the fact that jumps in α and σ^2 occur independently and that now jumps in α are much less frequent than were intensity jumps in the DY model and our Extension 1. Finally, σ^2 is now more volatile with a higher mean reversion speed, while exactly the opposite is true for α.

In terms of cash flow moments this new variant of the model cannot be distinguished from any of the previous specifications. Concerning asset pricing moments shown in Table 6 the expected excess return on the dividend claim is now 5.7% and thus lower than in DY and in Extension 1, but very close to the data. The risk-free rate is slightly below 1 percent and thus also very close to its empirical value. Concerning the variance risk premium we obtain an improvment compared to the other models, representing again a step closer to the data.

In terms of the quantitative output of the model the overall picture is not surprising. There is now less overall risk in the model due to the independence of the variance and the intensity processes, which implies higher valuation ratios as shown in Table 3. The new factor α of course influences the wealth-consumption as well as the p-d ratio, but the sum of the coefficients for σ^2 and α in the new model is (in absolute terms) less than the previous coefficient of σ^2. The coefficient for $\bar{\sigma}^2$ is also lower than before: $\bar{\sigma}^2$ drives the long-run level of σ^2 and α, and when the latter two are less important together than the 'old' σ^2 alone, their long-run level will also be less important as a determinant of both valuation ratios.

The main point of our new specification is that we can now distinguish between the impact of the amount of diffusion risk (represented by σ^2) and the amount of jump risk (represented by α). Table 3 shows that the coefficient for α in the p-d ratio is much larger in absolute terms than the one for σ^2. Intensity risk is thus much more relevant than the risk of adverse changes in volatility. Now that σ^2 has lost its double role as diffusive and jump-related risk factor, it is only responsible for the less important diffusion.

Looking at the coefficients of the state variables in the expected excess return on the dividend claim in Table 3 we find again that α (with a coefficient of 0.0293) is more important than σ^2 (with a coefficient of 0.01374). Table 4 further shows for x-risk the total diffusive and jump premia are about the same (1.33% for diffusive risk, and 1.22% for jump risk). Jumps in the uncertainty factors α and σ^2 still command higher premia than diffusive risks. For example, for α the jump-related premium is more than ten times the one for diffusion risk (1.51% vs. 0.13%), so that the overall weight of α in explaining time-variation in expected excess returns is greater than that of σ^2.

The variance risk premium is mainly driven by jump risk, i.e. by α. Here, jumps in α are most important, and the second largest component is the premium for the variance caused by jumps in x with a similar size as in the model with proportional intensity. Since we have separated diffusion and jump risk we can now pin down the main driver of the variance risk premium. It is almost exclusively paid for jumps in the intensity process α, while jumps in σ^2 hardly matter.

Concerning option-related moments we find that decoupling intensities from conditional variance, as also previously suggested in the option pricing literature by Santa-Clara and Yan (2010), indeed represents a significant step forward. The correlation between the changes in the level and the slope of the smile is now perfectly explained by the model. Concerning the smile level the model also produces very good results. As stated above, matching the numerical value for the slope is a problem common to all specifications analyzed in this paper. In further work one could try to mitigate this problem by allowing for correlations between the state variables and by introducing a negative average jump size in x. We opted for leaving the state variables uncorrelated in our parametrization to be able to clearly identify the structural roles of jump intensity and diffusive variance.

3.3 Predictability

3.3.1 Prediction via the Price-Dividend Ratio

We regress excess returns and cash flow growth (dividends and consumption) on the lagged value of the p-d ratio. Note that the p-d ratio is computed similar to the way it is computed in the real world in that the current price of the dividend claim is divided by the average of the twelve monthly dividends paid over the previous year.

The first column in Table 7 shows the values in the data, while the other columns present the results for different models.

We first analyze the DY model. Consumption growth is predictable with R^2 values ranging from 10.4% for a one year horizon to 9.8% for a five year horizon. The same pattern is observed for dividend growth, where the goodness of fit decreases from 17% for a one year horizon to 8% for five years. The intuition for the existence of predictability is that both consumption and dividend growth depend on x and thus vary over time. This part can be predicted due to the autocorrelation in x. Since the p-d ratio also depends on x, it can be used as a predictor for the two cash flow variables. A higher x implies both a greater p-d ratio and a higher future consumption and dividend growth, so that the coefficients in these regressions should be (and actually are) positive.

More importantly, excess return are also predictable. The values for the R^2 in the predictive regressions range from 7% for a one year horizon to 12% for a prediction horizon of five years. The fact that excess returns are predictable is rooted in their dependence on σ^2, which follows an autocorrelated process and is thus itself predictable. The p-d ratio shares this dependence on σ^2 and is thus a sensible predictor (of course, it also depends on x and $\bar{\sigma}^2$, so it can only be a noisy predictor of future excess returns). Excess returns increase in σ^2, while the p-d ratio decreases with σ^2, so the regression coefficients should be (and actually are) negative.

When the prediction horizon increases, so does R^2. The mechanics behind this result are as follows. Over the short run the predictability of excess returns via the p-d ratio is driven by the fact that future values of σ^2 strongly depend on its current level. Over the medium term the high speed of mean reversion in σ^2 makes it less powerful as a predictor, and predictability comes more from the level of long-run uncertainty $\bar{\sigma}^2$ which is much more persistent than σ^2.[15]

For the model with a square-root specification for $\bar{\sigma}^2$ (Extension 1) we observe a significant increase in the predictability of excess returns by more than 50% for longer horizons. The predictive regression for three and five year excess returns now exhibit values for R^2 of more than 15% and 18%, respectively, compared to 10.5% and 11.9% in the DY case. This massive improvement in longer-term predictability is a further key contribution in economic terms of our modification of the DY setup. It takes the model much closer to the data, since Table 7 shows an empirical R^2 of 17% and 23% for a three and a five year horizon, respectively.

The intuition behind this result is that according to Equation (14) the expected excess return now also depends on $\bar{\sigma}^2$.[16] This means that the predictable part of the excess return now represents a larger share of the total, so that there is more potential for predictability.

For the cash flow variables the degree of predictability decreases, most pronounced for shorter horizons. For example, the R^2 for the regression of one year dividend growth on the current p-d ratio drops from 17.1% in the DY case to 12.4%, which brings us again closer to the data (where the corresponding value is 8.6%). The reason for this effect is that the impact of $\bar{\sigma}^2$ on the predictive variable p-d ratio has increased compared to the OU case. Table 3 shows that the coefficient $A_{d1,\bar{\sigma}}$ goes up in absolute terms from -0.19 to -0.26. On the other hand the coefficient

[15]For very long horizons also this variable would become useless as a predictor, and predictability would vanish completely.

[16]Remember that, with a square-root process for $\bar{\sigma}^2$, $H_{\bar{\sigma}}$ in Equation (6) is no longer the zero matrix.

for x remains basically unchanged. Predictability of cash flow growth is based on the dependence of the p-d ratio on x, but it has now become harder to 'extract' x from the p-d ratio due to the stronger impact of $\bar{\sigma}^2$. For longer horizons one has to admit that the model has a hard time to match the data. Despite the fact that R^2 is decreasing slightly, it is still way too high compared to the data. This is a problem common to all the specifications analyzed in our paper and remains a challenge for the further development of asset pricing models.

We now look at our 'Extension 2' model. The main effects of a separation of jump intensity and diffusive variance have already been discussed above in Section 3.2. Concerning the predictability of excess returns the separation of $\tilde{l}$ from σ^2 causes lower values for R^2. Both the p-d ratio and excess returns are now driven by both α and σ^2, but with different weights for the two variables. Table 3 shows that the relative importance of α versus σ^2 is much higher for the p-d ratio than for the expected excess return. So the current p-d ratio is a noisier signal for future excess returns than in the case when both quantities did not depend on α, which causes the decrease in R^2. Note, however, that the average values for R^2 are still above those produced by DY for horizons of 3 and 5 years.

The predictive regressions for the cash flow variables now exhibit a higher R^2 than for Extension 1. The mechanics behind this results are that the sum of the coefficients for σ^2 and α in the expression for the p-d ratio is smaller than the coefficient for σ^2 in Extension 1. The degree of predictability of the cash flow variables is generally reduced by the fact that the p-d ratio depends on other variables than x. Since this dependence is now less pronounced, the explanatory power of the p-d ratio increases.

3.3.2 Prediction via the Variance Risk Premium

We now regress future cash flow growth and excess returns on the current value of the variance risk premium. Note that we now consider much shorter prediction horizons of one to five *months* instead of one to five *years* as before. As usual, the first column of the table (here Table 8) shows the empirical evidence, and the model results are presented in the remaining columns.

In the DY case the R^2 for the prediction of excess returns ranges from 3% (one month) to 10% (five months). Similar to the discussion for predictions based on the p-d ratio the fundamental reason for a non-zero R^2 in these regressions is that both the variance risk premium and expected excess return depend on σ^2. However, due to the high speed of mean-reversion in σ^2 it is only over the short run that this variable can potentially have predictive power at all. In contrast to predictions based on the p-d ratio $\bar{\sigma}^2$ cannot 'take over' as predictor over longer horizons, since the impact of the long-run mean on the variance risk premium is basically negligible with a coefficient of 0.00057 (see Table 3).

For the regressions with cash flow growth on the left-hand side we observe very small values for R^2, with a maximum of 1.7% for consumption growth and less than 1% for dividend growth. This is perfectly in line with intuition, since cash flow growth rates depend on x, while the variance risk premium does not.

For 'Extension 1' predictability of excess returns decreases slightly. The argument here is analogous to the explanations given above. With a square-root specification expected excess return are now also driven $\bar{\sigma}^2$ instead of just by σ^2 and x, but the impact of $\bar{\sigma}^2$ on the variance risk premium is still negligible (and even smaller than in the DY case). All the other findings for this case are basically the same as for DY.

'Extension 2' delivers smaller R^2 values for the predictive regressions than the DY model. Their results are somewhat on the high side compared to the data, while our models tend to understate the predictive power of the variance risk premium slightly. The main driver behind the predictive power of the variance risk premium is the intensity component α, and not the conditional variance σ^2, as it may seem given the DY results. The advantage of our extension is that we can now clearly identify which of the two roles of σ^2 in the DY model is actually responsible for predictability, and it is the fact that it drives the intensity of jumps.

4 Conclusion

In this paper we analyzed an LRR model with a stochastic growth rate and stochastic uncertainty, driven by a stochastic variance, a stochastic intensity, and a stochastic mean-reversion level of the latter two. Our benchmark model is Drechsler and Yaron (2010), who introduced a stochastic long-run mean for the conditional variance process as well as jumps in the Bansal and Yaron (2004) framework. This extension of the fundamental LRR model was able to match stylized facts in the data, like the predictability of excess returns by the current p-d ratio and the current variance risk premium pretty well.

We generalize and extend the DY model in two important dimensions. First, we separate diffusive risk from jump risk by introducing an autonomous process driving the intensity of jumps. In DY the conditional variance process plays a double role: It obviously and naturally represents the amount of diffusive risk in the economy, but at the same time the jump intensity is a constant multiple of it, so that the evolutions of the amounts of jump and diffusive risk are perfectly correlated. Based on empirical results from the option pricing literature we relax this assumption.

The main advantage of our more general setup is that it enables us to analyze the true driving force behind the results generated by the DY model explicitly. In our model it is straightforward to decide whether a result based on the presence of stochastic uncertainty in DY is due to this uncertainty representing diffusive stochastic volatility or time-variation in jump intensity. One of the key new insights we gain from our analyses is that it is almost exclusively jump risk which generates a substantial variance risk premium. For the equity premium we also find that the jump intensity is the more important factor, although here stochastic volatility is also relevant.

Furthermore, disentangling stochastic intensity and stochastic variance is key to explaining the empirically very low correlation between changes in the level and in the slope of the volatility smile. While models with a proportional intensity result in correlations well above 0.5, a model with independent intensity can perfectly explain the empirical number.

Second, we replace the OU specification chosen by DY for the long-run mean of the conditional variance by a square-root process. From a technical point of view this modification can be easily justified by the fact that it guarantees non-negativity of the process, while an OU process has normal transition densities, resulting in a positive probability for negative future values.

The even more important aspect of this new choice for the dynamics of the long-run mean is that it also has far-reaching economic implications. With the volatility of the long-run mean depending on its own level in the square-root case, expected excess returns also depend on it. This implies that they not only increase when the (short-run) variance or the jump intensity goes up, but also when the economy becomes more risky in the long run (with a higher long-run mean for the variance process). Due to this additional dependence of expected excess returns on the long-run mean of the conditional variance, future excess returns can be predicted much better than in the DY model. At the same time the larger impact of the long-run mean on the p-d ratio reduces its predictive power for future consumption and dividend growth. These gains bring the model closer to the data.

The next step in the development of the LRR model could be take the unobservability of conditional variances, their long-run means and the long-run risk factor explicitly into account in equilibrium and to study the implications of this new setup for risk premia, expected returns, and volatilities.

A Solving for the Equilibrium

A.1 Consumption Claim

The log wealth consumption ratio is

$$v_t = A_0 + A_1' Y_t,$$

and the Campbell-Shiller approximation for the log return on the consumption claim is

$$d\ln(R_{c,t}) = k_0 dt + k_1 dv_t - (1-k_1)v_t dt + dc_t.$$

The coefficients A_0 and A as well as the linearizing coefficients k_0 and k_1 solve the non-linear system of equations

$$\begin{aligned}
0 &= K'\chi - \theta(1-k_1)A_1 + 0.5\Xi + l_1'(\Psi(\chi) - e)' \\
0 &= \theta\left(\ln(\delta) + k_0 - (1-k_1)A_0\right) + M'\chi + 0.5\chi' h\chi + l_0'(\Psi(\chi) - e)' \\
0 &= A_0 + A_1'\mu_Y - \ln k_1 + \ln(1-k_1) \\
0 &= -\ln(k_1) + (1-k_1)A_1'\mu_Y - k_0 - (k_1 - 1)A_0
\end{aligned}$$

where K is from Equation (3), $\chi = \theta\left(\left(1 - \frac{1}{\psi}\right) e_c + k_1 A_1\right)$, $e_c = (1,0,0,0,0,0)'$, and $e = (1,1,1)$. Ξ is a 6×1-vector with entries given by $\Xi_3 = \chi' H_\sigma \chi$, $\Xi_4 = \chi' H_\alpha \chi$, $\Xi_5 = \chi' H_{\bar{\sigma}} \chi$, and $\Xi_1 = \Xi_2 = \Xi_6 = 0$. Ψ denotes the moment-generating function of the jump-size distribution. For a vector $u \in \mathbb{R}^6$, it is defined as

$$\Psi(u) = E^{\mathbb{P}}\left[e^{u'\xi}\right]$$

where the exponent $u'\xi$ of the exponential function is a 1×3-vector and where the exponent is evaluated componentwise. $\Psi(u)$ is then also a 1×3-vector. Finally, μ_Y denotes the mean of the state variables Y, where the entries are set to zero for the non-stationary state variables consumption and dividends.

We now look in more detail at the first set of equations, which give the coefficients A_1 for the state variables as a function of k_1. Plugging in K, the variance-covariance matrices H_σ, H_α, and $H_{\bar\sigma}$ and the specification of the jumps gives

$$\begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & -\kappa_x & 0 & 0 & 0 & \phi \\ 0 & 0 & -\kappa_\sigma & 0 & 0 & 0 \\ 0 & 0 & 0 & -\kappa_\alpha & 0 & 0 \\ 0 & 0 & k_{\sigma,\bar\sigma} & k_{\alpha,\bar\sigma} & -\kappa_{\bar\sigma} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \theta \begin{pmatrix} 1-\frac{1}{\psi}+k_1A_{1c} \\ k_1A_{1x} \\ k_1A_{1\sigma} \\ k_1A_{1\alpha} \\ k_1A_{1\bar\sigma} \\ k_1A_{1\delta} \end{pmatrix} - \theta(1-k_1)\begin{pmatrix} A_{1c} \\ A_{1x} \\ A_{1\sigma} \\ A_{1\alpha} \\ A_{1\bar\sigma} \\ A_{1\delta} \end{pmatrix}$$

$$+0.5\theta^2 \begin{pmatrix} 0 \\ 0 \\ \left(1-\frac{1}{\psi}\right)^2 \sigma_c^2 w_c + (k_1A_{1x})^2\sigma_x^2 w_x + (k_1A_{1\sigma})^2\sigma_\sigma^2 w_\sigma \\ (k_1A_{1\alpha})^2\sigma_\alpha^2 w_\alpha \\ (k_1A_{1\bar\sigma})^2\sigma_{\bar\sigma}^2 w_{\bar\sigma} \\ 0 \end{pmatrix}$$

$$+\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ l_{1x\sigma} & l_{1\sigma\sigma} & l_{1\alpha\sigma} \\ l_{1x\alpha} & l_{1\sigma\alpha} & l_{1\alpha\alpha} \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} E\left[e^{\theta k_1 A_{1x}\xi^x}-1\right] \\ E\left[e^{\theta k_1 A_{1\sigma}\xi^\sigma}-1\right] \\ E\left[e^{\theta k_1 A_{1\alpha}\xi^\alpha}-1\right] \end{pmatrix} = 0$$

where the quantities l_{1ij} represent the part of the intensity of jumps in variable i which depends on the current level of variable j.

From the first and the last line, it immediately follows that $A_{1c} = A_{1\delta} = 0$. Rewriting the second line and dividing by θ gives

$$1-\frac{1}{\psi}-\kappa_x k_1 A_{1x} - (1-k_1)A_{1x} = 0$$

which can be solved for A_{1x}:

$$A_{1x} = \frac{1-\frac{1}{\psi}}{1-k_1+k_1\kappa_x}.$$

After dividing by θ the fifth line yields

$$k_{\sigma,\bar\sigma}k_1A_{1\sigma} + k_{\alpha,\bar\sigma}k_1A_{1\alpha} - \kappa_{\bar\sigma}k_1A_{1\bar\sigma} - (1-k_1)A_{1\bar\sigma} + 0.5\theta\,(k_1A_{1\bar\sigma})^2\sigma_{\bar\sigma}^2 w_{\bar\sigma} = 0$$

which is equivalent to

$$(1-k_1+k_1\kappa_{\bar\sigma})\,A_{1\bar\sigma} - 0.5\theta\,(k_1A_{1\bar\sigma})^2\sigma_{\bar\sigma}^2 w_{\bar\sigma} = k_1\left(k_{\sigma,\bar\sigma}A_{1\sigma} + k_{\alpha,\bar\sigma}A_{1\alpha}\right). \quad \text{(A.1)}$$

If $\bar\sigma^2$ follows an OU-process, $w_{\bar\sigma}=0$, and we obtain

$$A_{1\bar\sigma} = \frac{k_1}{1-k_1+k_1\kappa_{\bar\sigma}}\left(k_{\sigma,\bar\sigma}A_{1\sigma} + k_{\alpha,\bar\sigma}A_{1\alpha}\right). \quad \text{(A.2)}$$

Since the only impact of $\bar{\sigma}$ on the consumption and dividend dynamics is via the mean-reversion levels of σ^2 and α, the coefficient $A_{1\bar{\sigma}}$ depends on the coefficients $A_{1\sigma}$ and $A_{1\alpha}$. We will show below that $A_{1\sigma}$ and $A_{1\alpha}$ are negative, so that $A_{1\bar{\sigma}}$ is also negative.

If $\bar{\sigma}^2$ follows a square-root process, $w_{\bar{\sigma}} = 1$. Then, the left hand side of Equation (A.1) increases, and the negative value of $A_{1\bar{\sigma}}$ for which the (negative) left hand side of Equation (A.1) equals the given (negative) term $k_1 \left(k_{\sigma,\bar{\sigma}} A_{1\sigma} + k_{\alpha,\bar{\sigma}} A_{1\alpha}\right)$ on the right hand side decreases further, i.e. becomes more negative. This is also in line with intuition, since $\bar{\sigma}$ now has an impact on its own volatility. The additional channel makes $\bar{\sigma}$ more important, and the coefficient $A_{1\bar{\sigma}}$ increases in absolute terms.

To get an approximate solution for $A_{1\sigma}$ and $A_{1\alpha}$, we first use a second order Taylor expansion of the exponential function. This gives

$$E\left[e^{\theta k_1 A_{1x} \xi^x} - 1\right] \approx \theta k_1 A_{1x} E\left[\xi^x\right] + 0.5 \left(\theta k_1 A_{1x}\right)^2 E\left[\left(\xi^x\right)^2\right]$$

and analogous approximations for the other two exponential terms.

The equation for $A_{1\sigma}$ then becomes

$$\begin{aligned}
& -\theta\kappa_\sigma k_1 A_{1\sigma} - \theta(1-k_1) A_{1\sigma} \\
& +0.5\theta^2 \left(\left(1 - \frac{1}{\psi}\right)^2 \sigma_c^2 w_c + \left(k_1 A_{1x}\right)^2 \sigma_x^2 w_x + \left(k_1 A_{1\sigma}\right)^2 \sigma_\sigma^2 w_\sigma \right) \\
& +l_{1x\sigma} \left(\theta k_1 A_{1x} E\left[\xi^x\right] + 0.5 \left(\theta k_1 A_{1x}\right)^2 E\left[\left(\xi^x\right)^2\right]\right) \\
& +l_{1\sigma\sigma} \left(\theta k_1 A_{1\sigma} E\left[\xi^\sigma\right] + 0.5 \left(\theta k_1 A_{1\sigma}\right)^2 E\left[\left(\xi^\sigma\right)^2\right]\right) \\
& +l_{1\alpha\sigma} \left(\theta k_1 A_{1\alpha} E\left[\xi^\alpha\right] + 0.5 \left(\theta k_1 A_{1\alpha}\right)^2 E\left[\left(\xi^\alpha\right)^2\right]\right) = 0.
\end{aligned}$$

Dividing by θ and sorting terms gives

$$\begin{aligned}
& \left(1 - k_1 + k_1 \left(\kappa_\sigma - l_{1\sigma\sigma} E\left[\xi^\sigma\right]\right)\right) A_{1\sigma} - 0.5\theta k_1^2 \left(\sigma_\sigma^2 w_\sigma + l_{1\sigma\sigma} E\left[\left(\xi^\sigma\right)^2\right]\right) A_{1\sigma}^2 \\
= \;& 0.5\theta \left(1 - \frac{1}{\psi}\right)^2 \sigma_c^2 w_c + l_{1x\sigma} \left(k_1 A_{1x} E\left[\xi^x\right] + 0.5\theta \left(k_1 A_{1x}\right)^2 E\left[\left(\xi^x\right)^2\right]\right) \\
& + l_{1\alpha\sigma} \left(k_1 A_{1\alpha} E\left[\xi^\alpha\right] + 0.5\theta \left(k_1 A_{1\alpha}\right)^2 E\left[\left(\xi^\alpha\right)^2\right]\right)
\end{aligned} \tag{A.3}$$

In a similar way, we get the equation for $A_{1\alpha}$:

$$\begin{aligned}
& \left(1 - k_1 + k_1 \left(\kappa_\alpha - l_{1\alpha\alpha} E\left[\xi^\alpha\right]\right)\right) A_{1\alpha} - 0.5\theta k_1^2 \left(\sigma_\alpha^2 w_\alpha + l_{1\alpha\alpha} E\left[\left(\xi^\alpha\right)^2\right]\right) A_{1\alpha}^2 \\
= \;& l_{1x\alpha} \left(k_1 A_{1x} E\left[\xi^x\right] + 0.5\theta \left(k_1 A_{1x}\right)^2 E\left[\left(\xi^x\right)^2\right]\right) \\
& + l_{1\sigma\alpha} \left(k_1 A_{1\sigma} E\left[\xi^\sigma\right] + 0.5\theta \left(k_1 A_{1\sigma}\right)^2 E\left[\left(\xi^\sigma\right)^2\right]\right)
\end{aligned} \tag{A.4}$$

If the intensity of the jumps is proportional to σ^2, $l_{1x\alpha} = l_{1\sigma\alpha} = 0$, and the right-hand side of Equation (A.4) is equal to zero, so that $A_{1\alpha} = 0$. The state variable α thus only has an impact on the wealth-consumption ratio if it influences the intensity of jumps in the other state variables x or σ.

In the simplest case, both the volatility and the jump intensity of α are constant, i.e. $w_\alpha = l_{1\alpha\alpha} = 0$. Then

$$\begin{aligned} A_{1\alpha} &= \frac{1}{1-k_1+k_1\kappa_\alpha}\Big[l_{1x\alpha}\left(k_1 A_{1x} E\left[\xi^x\right] + 0.5\theta\left(k_1 A_{1x}\right)^2 E\left[\left(\xi^x\right)^2\right]\right) \\ &\qquad + l_{1\sigma\alpha}\left(k_1 A_{1\sigma} E\left[\xi^\sigma\right] + 0.5\theta\left(k_1 A_{1\sigma}\right)^2 E\left[\left(\xi^\sigma\right)^2\right]\right)\Big] \end{aligned}$$

If the average jump size in x is non-positive and if $A_{1\sigma} < 0$ (which we will show below), the right hand side of Equation (A.4) is negative. This implies that $A_{1\alpha}$ is negative, too.

If the volatility of α or the intensity of jumps in α depends on α, we have $w_\alpha > 0$ and $l_{1\alpha\alpha} > 0$. Then, the left hand side of Equation (A.4) increases, and by a similar argument as above, $A_{1\alpha}$ decreases further, i.e. becomes more negative.

Finally, we look at the equation for $A_{1\sigma}$. The term on the right hand side differs from zero if σ has an impact on the diffusion volatility of consumption or on the intensity of jumps in the other state variables x and α. The first term which arises due to the impact of σ on the dividend volatility is negative for sure, and the other two terms are negative, too, if the average jump size in x is non-positive. Again, we then look at the simplest case first in which σ has no impact on its own diffusion volatility and on the intensity of jumps in its own level. Then, it holds that

$$\begin{aligned} A_{1\sigma} &= \frac{1}{1-k_1+k_1\kappa_\sigma}\Big[0.5\theta\left(1-\frac{1}{\psi}\right)^2 \sigma_c^2 w_c \\ &\qquad + l_{1x\sigma}\left(k_1 A_{1x} E\left[\xi^x\right] + 0.5\theta\left(k_1 A_{1x}\right)^2 E\left[\left(\xi^x\right)^2\right]\right) \\ &\qquad + l_{1\alpha\sigma}\left(k_1 A_{1\alpha} E\left[\xi^\alpha\right] + 0.5\theta\left(k_1 A_{1\alpha}\right)^2 E\left[\left(\xi^\alpha\right)^2\right]\right)\Big], \end{aligned}$$

which is negative.

If the volatility of σ or the intensity of jumps in σ depends on σ, we have $w_\sigma > 0$ and $l_{1\sigma\sigma} > 0$. Then, the left hand side of Equation (A.3) increases, and by a similar argument as above, $A_{1\sigma}$ decreases further, i.e. becomes more negative.

A.2 Market Prices of Risk and Risk-Free Rate

Plugging the expression for the log return on the consumption claim into the pricing kernel gives

$$d\ln M_t = \ldots dt - \Lambda' G(Y_t) dW_t - \Lambda' \xi_t dN_t$$

where the vector Λ denotes the market prices of risk and is given by

$$\Lambda = (1-\theta)k_1 A + \gamma e_c.$$

From the expected change of the pricing kernel, we also get the local risk-free rate

$$r_t = r_0 + r_1' Y_t$$

where

$$\begin{aligned} r_0 &= -\theta \ln \delta - (1-\theta)\ln k_1 + (1-\theta)(1-k_1)A'\mu_Y + \Lambda' M - 0.5\Lambda' h \Lambda - l_0'(\Psi(-\Lambda) - e)' \\ r_1 &= (1-\theta)(1-k_1)A + K'\Lambda - 0.5\Xi_r - l_1'(\Psi(-\Lambda) - e)' \end{aligned}$$

where Ξ_r is a 6×1-vector with entries given by $\Xi_{r,3} = \Lambda' H_\sigma \Lambda$, $\Xi_{r,4} = \Lambda' H_\alpha \Lambda$, $\Xi_{r,5} = \Lambda' H_{\bar{\sigma}} \Lambda$, and $\Xi_{r,1} = \Xi_{r,2} = \Xi_{r,6} = 0$.

A.3 Dividend Claim

The log price dividend ratio is

$$v_{d,t} = A_{d0} + A_{d1}' Y_t,$$

and the Campbell-Shiller approximation for the log return on the dividend claim is

$$d\ln(R_{d,t}) = k_{d0}dt + k_{d1}dv_{d,t} - (1-k_{d1})v_{d,t}dt + d\delta_t.$$

The coefficients A_{d0} and A_{d1} as well as the linearizing coefficients k_{d0} and k_{d1} solve the non-linear system of equations

$$\begin{aligned} 0 &= K'\chi_d + (1-\theta)(1-k_1)A_1 - (1-k_{d1})A_{d1} + 0.5\Xi_d + l_1'(\Psi(\chi_d) - e)' \\ 0 &= \theta\ln(\delta) + (1-\theta)\ln k_1 - (1-\theta)(1-k_1)A_1'\mu_Y - \ln k_{1d} + (1-k_{1d})A_{1d}'\mu_Y \\ &\quad + M'\chi_d + 0.5\chi_d' h \chi_d + l_0'(\Psi(\chi_d) - e)' \\ 0 &= A_{d0} + A_{d1}'\mu_Y - \ln k_{d1} + \ln(1-k_{d1}) \\ 0 &= -\ln(k_{d1}) + (1-k_{d1})(A_{d0} + A_{d1}'\mu_Y) - k_{d0} \end{aligned}$$

where $\chi_d = \theta k_{d1} A_{d1} + e_d - \Lambda$ and where $e_d = (0,0,0,0,0,1)'$. Ξ_d is a 6×1-vector with entries given by $\Xi_{d,3} = \chi_d' H_\sigma \chi_d$, $\Xi_{d,4} = \chi_d' H_\alpha \chi_d$, $\Xi_{d,5} = \chi_d' H_{\bar{\sigma}} \chi_d$, and $\Xi_{d,1} = \Xi_{d,2} = \Xi_{d,6} = 0$. Ψ denotes the moment-generating function of the jump-size distribution.

A.4 Variance Risk and Variance Risk Premium

The variance swap rate is the expectation of the quadratic variation from t to $t+\tau$ under the risk-neutral measure $\mathbb{Q}$

$$
\begin{aligned}
&E_t^{\mathbb{Q}}\left[\int_t^{t+\tau}(d\ln R_{d,s})^2\right] \\
&= (k_{d1}A_{d1}+e_\delta)'\left(h\,\tau + H_\sigma E_t^{\mathbb{Q}}\left[\int_t^{t+\tau}\sigma_s^2 ds\right] + H_\alpha E_t^{\mathbb{Q}}\left[\int_t^{t+\tau}\alpha_s ds\right] + H_{\bar\sigma}E_t^{\mathbb{Q}}\left[\int_t^{t+\tau}\bar\sigma_s^2 ds\right]\right) \\
&\qquad \cdot (k_{d1}A_{d1}+e_\delta) \\
&\quad + \left((k_{d1}A_{d1,x})^2 E_t^{\mathbb{Q}}\left[(\eta_t^x)^2\right] + (k_{d1}A_{d1,\sigma})^2 E_t^{\mathbb{Q}}\left[(\eta_t^\sigma)^2\right] + (k_{d1}A_{d1,\alpha})^2 E_t^{\mathbb{Q}}\left[(\eta_t^\alpha)^2\right]\right) \\
&\qquad \cdot \left(\tilde l_0^{\mathbb{Q}}\tau + \tilde l_1^{\mathbb{Q}}\left(\varphi_\alpha E_t^{\mathbb{Q}}\left[\int_t^{t+\tau}\alpha_s ds\right] + (1-\varphi_\alpha)E_t^{\mathbb{Q}}\left[\int_t^{t+\tau}\sigma_s^2 ds\right]\right)\right)
\end{aligned}
$$

where the jump intensities and the distribution of the jump sizes under the risk-neutral measure are given in Section 2. Both the expected quadratic variation and the variance swap rate depend on the expectations of the integrals over the state variables. We now solve for this expectation under $\mathbb{P}$, and the calculation under $\mathbb{Q}$ proceeds analogously. First, the dynamics of the state variables give an ode for $E_t^{\mathbb{P}}[Y_s]$

$$
\begin{aligned}
dE_t^{\mathbb{P}}[Y_s] &= \left(M + KE_t^{\mathbb{P}}[Y_s]\right)ds + E_t^{\mathbb{P}}[\xi_s]\left(l_0 + l_1 E_t^{\mathbb{P}}[Y_s]\right)ds \\
&= \left(M + E_t^{\mathbb{P}}[\xi_s]\,l_0\right)ds + \left(K + E_t^{\mathbb{P}}[\xi_s]\,l_1\right)E_t^{\mathbb{P}}[Y_s]\,ds.
\end{aligned}
$$

Solving this ode gives

$$
\begin{aligned}
E_t^{\mathbb{P}}[Y_s] &= \exp\left\{\left(K + E_t^{\mathbb{P}}[\xi_s]\,l_1\right)(s-t)\right\}\left[Y_t + \left(K + E_t^{\mathbb{P}}[\xi_s]\,l_1\right)^{-1}\left(M + E_t^{\mathbb{P}}[\xi_s]\,l_0\right)\right] \\
&\quad - \left(K + E_t^{\mathbb{P}}[\xi_s]\,l_1\right)^{-1}\left(M + E_t^{\mathbb{P}}[\xi_s]\,l_0\right).
\end{aligned}
$$

The integral over the expectation from t to $t+\tau$ can then be calculated as

$$
\begin{aligned}
&E_t^{\mathbb{P}}\left[\int_t^{t+\tau} Y_s ds\right] \\
&= \left(K + E_t^{\mathbb{P}}[\xi_s]\,l_1\right)^{-1}\left(e^{\left(K + E_t^{\mathbb{P}}[\xi_s] l_1\right)\tau} - I\right)\left(Y_t + \left(K + E_t^{\mathbb{P}}[\xi_s]\,l_1\right)^{-1}\left(M + E_t^{\mathbb{P}}[\xi_s]\,l_0\right)\right) \\
&\quad - \left(K + E_t^{\mathbb{P}}[\xi_s]\,l_1\right)^{-1}\left(M + E_t^{\mathbb{P}}[\xi_s]\,l_0\right)\tau,
\end{aligned}
$$

where I denotes the identity matrix.

References

Bansal, R., R. Dittmar, and C. Lundblad, 2005, Consumption, Dividends, and the Cross-Section of Equity Returns, *Journal of Finance* 60, 1639–1672.

Bansal, R., and I. Shaliastovich, 2010, Confidence Risk and Asset Prices, *American Economic Association Papers and Proceedings* 100, 537–541.

Bansal, R., and I. Shaliastovich, 2011, Learning and Asset-Price Jumps, forthcoming *Review of Financial Studies.*

Bansal, R., and A. Yaron, 2004, Risks For the Long Run: A Potential Resolution of Asset Pricing Puzzles, *Journal of Finance* 59, 1481–1509.

Bates, D., 2000, Post-'87 crash fears in the S&P 500 futures options market, *Journal of Econometrics* 94, 181–238.

Beeler, J., and J. Y. Campbell, 2009, The Long-Run Risks Model and Aggregate Asset Prices: An Empirical Assessment, NBER Working Paper.

Benzoni, L., P. Collin-Dufresne, and R. S. Goldstein, 2010, Explaining Asset Pricing Puzzles Associated with the 1987 Market Crash, forthcoming *Journal of Financial Economics.*

Bollerslev, T., N. Sizova, and G. Tauchen, 2011, Volatility in Equilibrium: Asymmetries and Dynamic Dependencies, forthcoming *Journal of Finance.*

Bollerslev, T., G. Tauchen, and H. Zhou, 2010, Expected Stock Returns and Variance Risk Premia, *Review of Financial Studies* 22, 4463–4492.

Campbell, J.Y., and R. J. Shiller, 1988, Stock Prices, Earnings, and Expected Dividends, *Journal of Finance* 43, 661–676.

Carr, P., and L. Wu, 2009, Variance Risk Premiums, *Review of Financial Studies* 22, 1311–1341.

Drechsler, I., 2009, Uncertainty, Time-Varying Fear, and Asset Prices, Working Paper.

Drechsler, I., and A. Yaron, 2010, What's Vol Got to Do With It, forthcoming *Review of Financial Studies.*

Egloff, D., M. Leippold, and L. Wu, 2010, Variance Risk Dynamics, Variance Risk Premia, and Optimal Variance Swap Investments, forthcoming *Journal of Financial and Quantitative Analysis.*

Epstein, L. G., and S. E. Zin, 1989, Substitution, Risk Aversion, and the Temporal Behavior of Consumption Growth and Asset Returns I: A Theoretical Framework, *Econometrica* 57, 937–969.

Eraker, B., and I. Shaliastovich, 2008, An Equilibrium Guide to Designing Affine Pricing Models, *Mathematical Finance* 18, 519–543.

Pan, J., 2002, The jump-risk premia implicit in options: evidence from an integrated time-series study, *Journal of Financial Economics* 63, 3–50.

Santa-Clara, P., and S. Yan, 2010, Crashes, Volatility, and the Equity Premium: Lessons from S&P 500 Options, *Review of Economics and Statistics* 92, 435–451.

Todorov, V., 2010, Variance Risk Premium Dynamics, *Review of Financial Studies* 23, 345–383.

Wachter, J., 2010, Can time-varying risk of rare disasters explain aggregate stock market volatility?, Working Paper.

Wang, Z., and P. V. Bidarkota, 2010, A Long-Run Risks Model of Asset Pricing with Fat Tails, *Review of Finance* 14, 409–449.

Yan, S., 2011, Jump risk. stock returns, and the slope of the implied volatility smile, *Journal of Financial Economics* 99, 216–233.

Zhou, G., and Y. Zhu, 2009, A Long-run Risks Model with Long- and Short-run Volatilities: Explaining Predictability and Volatility Risk Premium, Working Paper.

Table 1:
Parameters

Common parameters for all models									
γ	ψ	β							
10	2	0.9881							
μ_c			σ_c	w_c					
0.0192			0.022863	0.5000					
μ_δ	ϕ		σ_δ	w_δ	$w_{c\delta}$				
0.0192	2.5000		0.130320	0.1250	0.2				
		κ_x	σ_x	w_x	$E[\xi^x]$	$Var[\xi^x]$	$\tilde{l}_{x,0}$	$\tilde{l}_{x,1}$	φ_x
		0.2880	0.008779	1.0000	0.0000	0.000085	0.0000	0.8000	0.0000
		$\kappa_{\bar{\sigma}}$	$\sigma_{\bar{\sigma}}$						
		0.1800	0.346410						
DY									
$k_{\sigma,\bar{\sigma}}$		κ_σ	σ_σ	w_σ	$E[\xi^\sigma]$		$\tilde{l}_{\sigma,0}$	$\tilde{l}_{\sigma,1}$	φ_σ
1.5600		3.6000	1.212436	1.0000	2.5500		0.0000	0.8000	0.0000
				$w_{\bar{\sigma}}$					
				0.0000					
Extension 1									
$k_{\sigma,\bar{\sigma}}$		κ_σ	σ_σ	w_σ	$E[\xi^\sigma]$		$\tilde{l}_{\sigma,0}$	$\tilde{l}_{\sigma,1}$	φ_σ
1.5600		3.6000	1.212436	1.0000	2.5500		0.0000	0.8000	0.0000
				$w_{\bar{\sigma}}$					
				1.0000					
Extension 2									
$k_{\sigma,\bar{\sigma}}$		κ_σ	σ_σ	w_σ	$E[\xi^\sigma]$		$\tilde{l}_{\sigma,0}$	$\tilde{l}_{\sigma,1}$	φ_σ
1.4400		6.0800	1.385641	1.0000	5.8000		0.0000	0.8000	1.0000
$k_{\alpha,\bar{\sigma}}$		κ_α	σ_α	w_α	$E[\xi^\alpha]$		$\tilde{l}_{\alpha,0}$	$\tilde{l}_{\alpha,1}$	φ_α
1.4400		2.0200	1.039230	1.0000	5.8000		0.0000	0.1000	1.0000
				$w_{\bar{\sigma}}$					
				1.0000					

The table gives the parameters for the four models we consider in our analysis. 'DY' is the model suggested by Drechsler and Yaron (2010) with an OU process for $\bar{\sigma}^2$, the long-run mean of variance, and a jump intensity which is perfectly correlated with the stochastic volatility process. In 'Extension 1' the OU process for $\bar{\sigma}^2$ is replaced by a square-root process, while 'Extension 2' refers to the model where we additionally introduce the autonomous intensity process α.

Table 2:
Cash Flows: Descriptive Statistics and Moments from the Models

	Data	**DY**				**Extension 1**				**Extension 2**			
	S&P 500	Mean	5%	50%	95%	Mean	5%	50%	95%	Mean	5%	50%	95%
$E(\Delta\delta)$	0.0090	0.0196	-0.0081	0.0202	0.0467	0.0192	-0.0091	0.0186	0.0469	0.01990	-0.01034	0.01952	0.04928
$\sigma(\Delta\delta)$	0.1346	0.1092	0.0935	0.1090	0.1262	0.1086	0.0922	0.1082	0.1250	0.10866	0.09310	0.10826	0.12511
$AC1(\Delta\delta)$	0.0855	0.2674	0.0843	0.2711	0.4437	0.2602	0.0726	0.2676	0.4248	0.26337	0.07517	0.26291	0.44752
$E(\Delta c)$	0.0198	0.0191	0.0122	0.0188	0.0263	0.0191	0.0123	0.0193	0.0262	0.01911	0.01139	0.01903	0.02651
$\sigma(\Delta c)$	0.0222	0.0214	0.0171	0.0210	0.0271	0.0212	0.0166	0.0208	0.0269	0.02113	0.01665	0.02085	0.02676
$AC1(\Delta c)$	0.4860	0.3592	0.1639	0.3646	0.5526	0.3585	0.1586	0.3596	0.5559	0.35760	0.15086	0.36135	0.54255
$corr(\Delta c, \Delta\delta)$	0.5901	0.2713	0.0537	0.2795	0.4673	0.2683	0.0646	0.2728	0.4676	0.26752	0.06178	0.26973	0.47221

The first column of the table shows the values in the data, represented by the S&P 500 index from 1930 to 2006. The statistics for the models are computed as in Drechsler and Yaron (2010) from 1,000 samples of length 77 years each. $\Delta\delta$ (Δc) is the growth rate of log dividends (log consumption), $\sigma(\cdot)$ denotes the standard deviation of log changes, and $AC1(\cdot)$ represents the first-order autocorrelation. The statistics are given for yearly growth rates. Deviations of the numbers shown in the table from those shown in Drechsler and Yaron (2010) are due our model being formulated in continuous rather than in discrete time and to sampling variation. 'DY' stands for the model presented in Drechsler and Yaron (2010), 'Extension 1' denotes the model where the OU process for $\bar{\sigma}^2$ is replaced by a square-root specification, 'Extension 2' is the model where additionally the jump intensity is an affine function of α instead of σ^2.

Table 3:
Coefficients of the State Variables in Asset Pricing Moments

	DY	**Ext. 1**	**Ext. 2**		**DY**	**Ext. 1**	**Ext. 2**
A: Log Wealth-Consumption Ratio				**B: Log Price-Dividend Ratio**			
average	4.53328	4.46227	4.58057	average	2.94984	2.82906	3.06257
k_0	0.05888	0.06236	0.05667	k_{d0}	0.19776	0.21518	0.18254
k_1	0.98937	0.98860	0.98985	k_{d1}	0.95026	0.94423	0.95532
const.	4.57149	4.51090	4.62256	const.	3.16825	3.11693	3.30682
x	1.69165	1.68850	1.69363	x	6.18395	6.10294	6.25370
σ^2	-0.00416	-0.00411	-0.00042	σ^2	-0.02831	-0.02715	-0.00167
α	—	—	-0.00353	α	—	—	-0.02292
$\bar{\sigma}^2$	-0.03404	-0.04452	-0.03803	$\bar{\sigma}^2$	-0.19010	-0.26072	-0.21966
C: Expected Excess Return Dividend Claim				**D: Variance Risk Premium**			
average	0.06050	0.06767	0.05493	average	0.00960	0.00872	0.01298
const.	0.00843	-0.00349	-0.00349	const.	0.00001	0.00000	0.00000
x	0	0	0	x	0	0	0
σ^2	0.05208	0.05009	0.01374	σ^2	0.00903	0.00817	0.00000
α	—	—	0.02930	α	—	—	0.01223
$\bar{\sigma}^2$	0	0.02106	0.01537	$\bar{\sigma}^2$	0.00057	0.00055	0.00075

The table gives the coefficients of the affine functions for the asset pricing moments. For the wealth-consumption ratio and the price-dividend ratio, the table shows the average value of the respective moment (based on the stationary means of the state variables), the linearizing constants k_0, k_1, k_{d0}, and k_{d1}, and the coefficients A_0, A_1, A_{d0}, and A_{d1}. For the expected excess return on the dividend claim and the variance risk premium, the table also gives the average and the coefficients. The state variables are the long-run growth rate x, the current variance σ^2, the additional intensity risk factor α, and the long-run mean reversion level $\bar{\sigma}^2$ for the variance and the intensity. An entry of '0' means that the number is exactly equal to zero. 'DY' stands for the model presented in Drechsler and Yaron (2010), 'Extension 1' denotes the model where the OU process for $\bar{\sigma}^2$ is replaced by a square-root specification, 'Extension 2' is the model where additionally the jump intensity is an affine function of α instead of σ^2.

Table 4:
Decomposition of Expected Excess Returns on the Dividend Claim

		Constant	σ^2	α	$\bar{\sigma}^2$	Sum
DY						
Diffusion Risk	x	0	1.31	—	0	1.31
	σ^2	0	0.26	—	0	0.26
	$\bar{\sigma}^2$	1.19	0	—	0	1.19
	dividends	-0.35	0.04	—	0	-0.31
Sum (diffusion risk)		0.84	1.61	—	0	2.45
Jump Risk	x	0	1.20	—	0	1.20
	σ^2	0	2.40	—	0	2.40
Sum (jump risk)		0	3.60	—	0	3.60
Sum (diffusion and jump risk)		0.84	5.21	—	0	6.05
Extension 1						
Diffusion Risk	x	0	1.28	—	0	1.28
	σ^2	0	0.24	—	0	0.24
	$\bar{\sigma}^2$	0	0	—	2.11	2.11
	dividends	-0.35	0.04	—	0	-0.31
Sum (diffusion risk)		-0.35	1.57	—	2.11	3.32
Jump Risk	x	0	1.18	—	0	1.18
	σ^2	0	2.27	—	0	2.27
Sum (jump risk)		0	3.44	—	0	3.44
Sum (diffusion and jump risk)		-0.35	5.01	—	2.11	6.77
Extension 2						
Diffusion Risk	x	0	1.33	0	0	1.33
	σ^2	0	0.00	0	0	0.00
	α	0	0	0.13	0	0.13
	$\bar{\sigma}^2$	0	0	0	1.54	1.54
	dividends	-0.35	0.04	0.00	0.00	-0.31
Sum (diffusion risk)		-0.35	1.37	0.13	1.54	2.69
Jump Risk	x	0	0	1.22	0	1.22
	σ^2	0	0	0.07	0	0.07
	α	0	0	1.51	0	1.51
Sum (jump risk)		0	0	2.80	0	2.80
Sum (diffusion and jump risk)		-0.35	1.37	2.93	1.54	5.49

The table gives the decomposition of the expected excess returns on the dividend claim, where we distinguish between the premia for diffusive risk and the premia for jump risk. The numbers are stated in percent. The rows give the risk factor which causes the risk premium, while the columns decompose the premium into a constant part and a linear combination of the state variables. For example, the row related to the long-run risk factor x in the first panel states the risk premium for diffusion risk in x is given by $0.0131\sigma^2$. 'DY' stands for the model presented in Drechsler and Yaron (2010), 'Extension 1' denotes the model where the OU process for $\bar{\sigma}^2$ is replaced by a square-root specification, 'Extension 2' is the model where additionally the jump intensity is an affine function of α instead of σ^2.

Table 5:
Decomposition of Variance Risk Premium

		Constant	σ^2	α	$\bar{\sigma}^2$	Sum
DY						
Diffusion Risk	x	0.00	1.47	—	0.06	1.54
	σ^2	0.00	0.59	—	0.03	0.61
	$\bar{\sigma}^2$	0	0	—	0	0
	dividends	0.00	1.06	—	0.05	1.11
Sum (diffusion risk)		0.01	3.11	—	0.14	3.26
Jump Risk	x	0.01	4.76	—	0.28	5.05
	σ^2	0.05	82.38	—	5.26	87.69
Sum (jump risk)		0.05	87.14	—	5.54	92.74
Sum (diffusion and jump risk)		0.06	90.25	—	5.68	95.99
Extension 1						
Diffusion Risk	x	0.00	1.40	—	0.07	1.47
	σ^2	0.00	0.53	—	0.03	0.56
	$\bar{\sigma}^2$	0.00	0	—	0.34	0.34
	dividends	0.00	1.04	—	0.05	1.09
Sum (diffusion risk)		0.00	2.97	—	0.48	3.45
Jump Risk	x	0.00	4.59	—	0.28	4.87
	σ^2	0.02	74.14	—	4.76	78.92
Sum (jump risk)		0.02	78.73	—	5.04	83.79
Sum (diffusion and jump risk)		0.03	81.70	—	5.51	87.25
Extension 2						
Diffusion Risk	x	0.00	0.01	0.63	0.03	0.67
	σ^2	0.00	0.00	0.00	0.00	0.00
	α	0.00	0	0.23	0.01	0.24
	$\bar{\sigma}^2$	0.00	0	0	0.21	0.21
	dividends	0.00	0.01	0.44	0.02	0.47
Sum (diffusion risk)		0.00	0.02	1.30	0.27	1.59
Jump Risk	x	0.00	0	4.62	0.26	4.88
	σ^2	0.00	0	0.29	0.02	0.30
	α	0.03	0	116.06	6.96	123.06
Sum (jump risk)		0.04	0	120.96	7.24	128.24
Sum (diffusion and jump risk)		0.04	0.02	122.26	7.50	129.83

The table gives the decomposition of the annualized variance risk premium on the dividend claim, where we distinguish between the premia for diffusive risk and for jump risk. The numbers are multiplied by 10,000, i.e. they are given in basis points. The rows give the risk factor which causes the variance of the return, while the columns decompose the premium into a constant part and a linear combination of the state variables. For example, the row related to the long-run risk factor x in the first panel states that the contribution of diffusion risk in x to the variance risk premium over the next 30 days is given by $0.00015\sigma^2 + 0.00001\bar{\sigma}^2$. 'DY' stands for the model presented in Drechsler and Yaron (2010), 'Extension 1' denotes the model where the OU process for $\bar{\sigma}^2$ is replaced by a square-root specification, 'Extension 2' is the model where additionally the jump intensity is an affine function of α instead of σ^2.

Table 6:
Asset Pricing: Descriptive Statistics and Moments from the Models

	Data	**DY**				**Extension 1**				**Extension 2**			
	S&P 500	Mean	5%	50%	95%	Mean	5%	50%	95%	Mean	5%	50%	95%
$E(r_m)$	0.0617	0.0721	0.0453	0.0722	0.0979	0.0783	0.0508	0.0775	0.1060	0.0663	0.0375	0.0660	0.0949
$\sigma(r_m)$	0.1960	0.1707	0.1423	0.1679	0.2087	0.1792	0.1456	0.1767	0.2185	0.1673	0.1372	0.1648	0.2030
$E(r_f)$	0.0111	0.0086	0.0020	0.0090	0.0135	0.0083	0.0009	0.0089	0.0137	0.0096	0.0036	0.0098	0.0143
$\sigma(r_f)$	0.0412	0.0173	0.0085	0.0162	0.0306	0.0170	0.0080	0.0155	0.0298	0.0143	0.0069	0.0129	0.0265
$E(r_m - r_f)$	0.0506	0.0636	0.0374	0.0635	0.0890	0.0701	0.0418	0.0691	0.0998	0.0568	0.0288	0.0568	0.0845
$E(p-d)$	3.2511	2.9441	2.8624	2.9471	3.0198	2.8357	2.7085	2.8442	2.9376	3.0720	2.9699	3.0778	3.1598
$\sigma(p-d)$	0.4346	0.1593	0.1212	0.1562	0.2078	0.1889	0.1307	0.1824	0.2620	0.1703	0.1182	0.1651	0.2435
$E(VP)$	0.0152	0.0103	0.0061	0.0098	0.0159	0.0088	0.0047	0.0084	0.0147	0.0130	0.0076	0.0126	0.0204
$\sigma(VP)$	0.0172	0.0170	0.0075	0.0159	0.0312	0.0150	0.0063	0.0137	0.0274	0.0184	0.0065	0.0152	0.0401
$Skew(VP)$	2.4550	3.8039	2.2579	3.5709	6.0777	3.8642	2.2314	3.6192	6.1510	3.7490	1.1230	3.7106	6.9128
$Kurt(VP)$	12.6879	20.6956	7.9183	17.4127	45.3676	21.3378	7.8653	17.7295	46.0840	22.7069	3.8846	19.4786	55.1211
$E(\nu)$	0.1917	0.1817	0.1664	0.1803	0.1998	0.1882	0.1668	0.1867	0.2142	0.1739	0.1582	0.1732	0.1932
$\sigma(\nu)$	0.0616	0.0597	0.0352	0.0579	0.0902	0.0588	0.0353	0.0574	0.0867	0.0428	0.0260	0.0406	0.0664
$E(s)$	0.0175	0.0067	0.0051	0.0067	0.0084	0.0055	0.0044	0.0055	0.0068	0.0048	0.0037	0.0048	0.0059
$\sigma(s)$	0.0045	0.0044	0.0038	0.0045	0.0050	0.0037	0.0031	0.0037	0.0043	0.0033	0.0025	0.0033	0.0041
$corr(\Delta\nu, \Delta s)$	0.0694	0.6620	0.5129	0.6665	0.8087	0.6931	0.5408	0.7003	0.8237	0.0711	-0.1233	0.0689	0.2779

The first column shows the empirical values, based on the S&P 500 index from 1930 to 2006 and on options on theS&P 500 for the period from 1990 to 2006. The statistics for the models are computed as in Drechsler and Yaron (2010) from 1,000 samples of length 77 years each. r_m is the market return, r_f is the risk-free rate, $p-d$ is the log price-dividend ratio, VP is the variance risk premium. The first block shows the numbers obtained from our model with an OU-process for long run variance and an intensity proportional to the conditional variance. Any deviations of these entries from the corresponding numbers in Drechsler and Yaron (2010) are due to to sampling variation and to the fact that our model is formulated in continuous rather than in discrete time. 'DY' stands for the model presented in Drechsler and Yaron (2010), 'Extension 1' denotes the model where the OU process for $\bar{\sigma}^2$ is replaced by a square-root specification, 'Extension 2' is the model where additionally the jump intensity is an affine function of α instead of σ^2.

Table 7:
Predictive Regressions of Excess Returns and Cash Flow Growth on the Price-Dividend Ratio

	Data	**DY**				**Extension 1**				**Extension 2**			
		Mean	5%	50%	95%	Mean	5%	50%	95%	Mean	5%	50%	95%
Return Dividend Claim													
$\beta(1y)$	-0.0823	-0.21544	-0.46423	-0.20916	0.00054	-0.23282	-0.45036	-0.21854	-0.04726	-0.18110	-0.43044	-0.16634	0.01794
$R^2(1y)$	0.0331	0.04927	0.00084	0.03611	0.14694	0.06424	0.00289	0.05403	0.15971	0.04226	0.00052	0.02959	0.12959
$\beta(3y)$	-0.2262	-0.14149	-0.31178	-0.13783	0.01387	-0.16768	-0.31226	-0.16700	-0.02846	-0.13838	-0.30513	-0.13868	0.02260
$R^2(3y)$	0.0989	0.07922	0.00070	0.05896	0.23699	0.11881	0.00472	0.10231	0.29491	0.08491	0.00071	0.06359	0.24798
$\beta(5y)$	-0.3493	-0.10904	-0.25111	-0.10532	0.02791	-0.13530	-0.26302	-0.13566	-0.00967	-0.11372	-0.25484	-0.11700	0.02479
$R^2(5y)$	0.1129	0.09279	0.00082	0.06610	0.29028	0.14627	0.00323	0.12202	0.37201	0.10772	0.00101	0.07986	0.30793
Excess Return Dividend Claim													
$\beta(1y)$	-0.1006	-0.26345	-0.52333	-0.25765	-0.03982	-0.27475	-0.50003	-0.26107	-0.08360	-0.22221	-0.47064	-0.20616	-0.01870
$R^2(1y)$	0.0481	0.06769	0.00251	0.05537	0.17942	0.08534	0.01038	0.07552	0.19715	0.05766	0.00144	0.04361	0.15868
$\beta(3y)$	-0.2969	-0.17232	-0.34262	-0.17075	-0.01651	-0.19639	-0.34284	-0.19477	-0.05998	-0.16780	-0.33184	-0.16653	-0.00980
$R^2(3y)$	0.1683	0.10515	0.00227	0.08842	0.28259	0.15292	0.01441	0.14040	0.34288	0.11200	0.00266	0.09251	0.29302
$\beta(5y)$	-0.4944	-0.13163	-0.27337	-0.13059	0.00722	-0.15713	-0.28243	-0.15813	-0.02957	-0.13616	-0.27972	-0.13938	0.00415
$R^2(5y)$	0.2319	0.11932	0.00192	0.09483	0.32691	0.18309	0.00734	0.16343	0.42781	0.13730	0.00180	0.11116	0.36098
Log Consumption Growth													
$\beta(1y)$	0.0146	0.03769	-0.00057	0.03817	0.07348	0.02543	-0.00628	0.02570	0.05910	0.03280	-0.00521	0.03346	0.06860
$R^2(1y)$	0.0876	0.10378	0.00196	0.08009	0.28050	0.07811	0.00060	0.05089	0.23846	0.09526	0.00114	0.07294	0.27424
$\beta(3y)$	0.0238	0.02605	-0.00739	0.02616	0.05752	0.01733	-0.01412	0.01828	0.04695	0.02248	-0.01045	0.02345	0.05473
$R^2(3y)$	0.0509	0.10605	0.00104	0.07412	0.31289	0.08977	0.00052	0.05914	0.30434	0.10261	0.00094	0.06624	0.33043
$\beta(5y)$	0.0235	0.01931	-0.00833	0.01929	0.05045	0.01302	-0.01538	0.01358	0.04117	0.01637	-0.01364	0.01646	0.04681
$R^2(5y)$	0.0369	0.09818	0.00054	0.06080	0.31852	0.09422	0.00033	0.05523	0.33807	0.09989	0.00046	0.05820	0.33843
Log Dividend Growth													
$\beta(1y)$	0.0921	0.27794	0.12657	0.27485	0.44433	0.19759	0.04603	0.19145	0.36362	0.25386	0.09435	0.24699	0.42952
$R^2(1y)$	0.0859	0.17102	0.03547	0.15997	0.34183	0.12359	0.00788	0.11112	0.28320	0.15836	0.02627	0.14746	0.32838
$\beta(3y)$	0.0768	0.12236	-0.01104	0.12398	0.25445	0.08369	-0.03365	0.08607	0.20994	0.11020	-0.01958	0.10728	0.24776
$R^2(3y)$	0.0254	0.09412	0.00151	0.07113	0.25803	0.07490	0.00047	0.05005	0.23736	0.09041	0.00112	0.06088	0.29117
$\beta(5y)$	0.0534	0.08103	-0.04193	0.08079	0.20138	0.05364	-0.05744	0.05447	0.17070	0.07208	-0.04658	0.07003	0.19906
$R^2(5y)$	0.0098	0.08195	0.00058	0.05236	0.25839	0.07285	0.00044	0.03980	0.26184	0.08210	0.00059	0.04698	0.27575

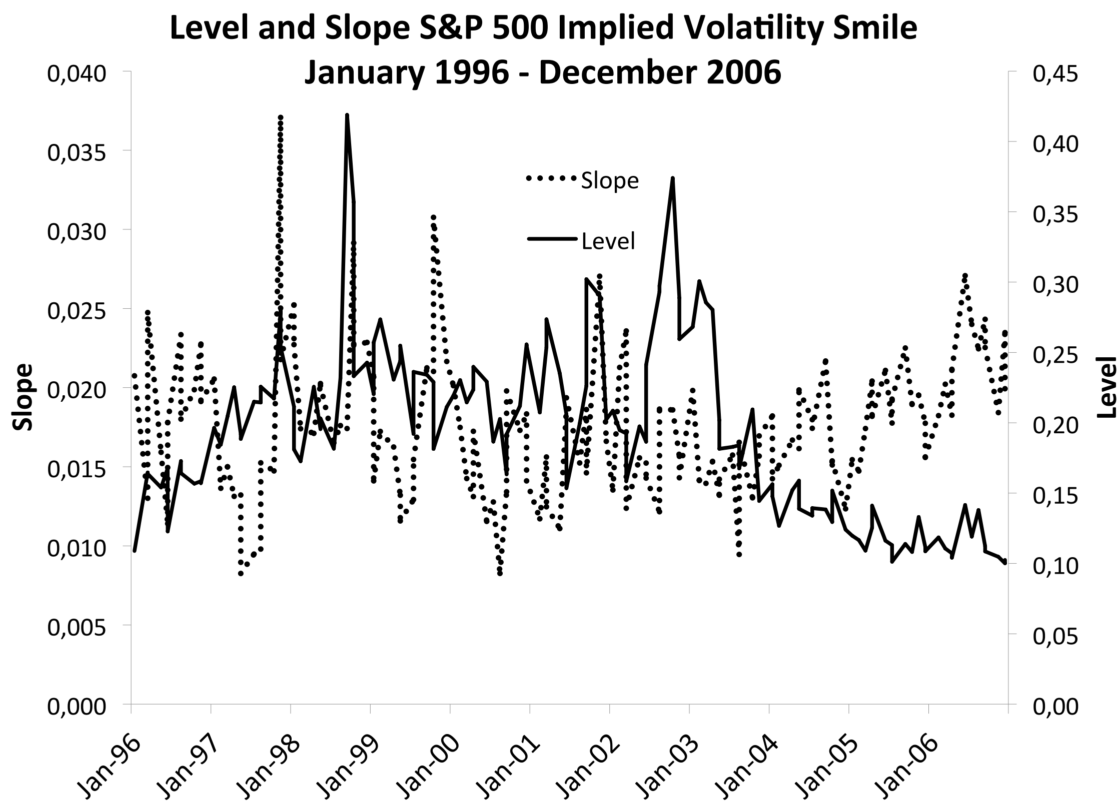
The first column shows the empirical coefficients for the predictive regressions based on the period from 1930 to 2006. The statistics for the models are computed as in Drechsler and Yaron (2010) from 1,000 samples of length 77 years each. The first block shows the numbers obtained from our model with an OU-process for long run variance and an intensity proportional to the conditional variance. Any deviations of these entries from the corresponding numbers in Drechsler and Yaron (2010) are due to to sampling variation and to the fact that our model is formulated in continuous rather than in discrete time. 'DY' stands for the model presented in Drechsler and Yaron (2010), 'Extension 1' denotes the model where the OU process for $\bar{\sigma}^2$ is replaced by a square-root specification, 'Extension 2' is the model where additionally the jump intensity is an affine function of α instead of σ^2.

Table 8:
Predictive Regressions of Excess Returns and Cash Flow Growth on the Variance Risk Premium

	Data	DY				Extension 1				Extension 2			
		Mean	5%	50%	95%	Mean	5%	50%	95%	Mean	5%	50%	95%
Return Dividend Claim													
$\beta(1m)$	0.4920	0.50303	0.19840	0.48648	0.86912	0.55213	0.21015	0.54098	0.97114	0.29920	0.05594	0.29099	0.58840
$R^2(1m)$	0.0450	0.02492	0.00477	0.02281	0.05339	0.02202	0.00350	0.01998	0.04809	0.01209	0.00049	0.00927	0.03351
$\beta(3m)$	1.0250	0.42894	0.16233	0.41784	0.73685	0.48061	0.19060	0.47436	0.81076	0.27577	0.05700	0.26340	0.53699
$R^2(3m)$	0.0680	0.05808	0.01016	0.05253	0.12824	0.05260	0.00778	0.04684	0.11840	0.03166	0.00155	0.02351	0.08688
$\beta(5m)$	1.3500	0.37613	0.14902	0.36430	0.63473	0.42630	0.16701	0.41941	0.71104	0.25367	0.04522	0.24190	0.50217
$R^2(5m)$	0.0720	0.07828	0.01218	0.07103	0.17284	0.07148	0.01058	0.06411	0.15947	0.04553	0.00197	0.03376	0.12330
Excess Return Dividend Claim													
$\beta(1m)$	0.4862	0.58310	0.27598	0.56550	0.95028	0.64128	0.30407	0.63177	1.06052	0.34833	0.10917	0.33815	0.63297
$R^2(1m)$	0.0442	0.03299	0.00893	0.03059	0.06594	0.02926	0.00698	0.02676	0.05836	0.01600	0.00123	0.01284	0.04143
$\beta(3m)$	1.0126	0.49820	0.23344	0.48882	0.80159	0.55814	0.27099	0.54974	0.88662	0.32178	0.10607	0.30962	0.58167
$R^2(3m)$	0.0667	0.07698	0.01914	0.07199	0.15722	0.06968	0.01536	0.06367	0.14602	0.04180	0.00362	0.03268	0.10557
$\beta(5m)$	1.3317	0.43669	0.21689	0.42501	0.68936	0.49445	0.24480	0.48571	0.77621	0.29612	0.08923	0.28382	0.53878
$R^2(5m)$	0.0714	0.10347	0.02503	0.09751	0.21206	0.09432	0.02055	0.08597	0.19893	0.05994	0.00499	0.04612	0.15516

The first column shows the empirical coefficients for the predictive regressions based on the period from 1990 to 2006. The statistics for the models are computed as in Drechsler and Yaron (2010) from 1,000 samples of length 77 years each. The first block shows the numbers obtained from our model with an OU-process for long run variance and an intensity proportional to the conditional variance. Any deviations of these entries from the corresponding numbers in Drechsler and Yaron (2010) are due to to sampling variation and to the fact that our model is formulated in continuous rather than in discrete time. 'DY' stands for the model presented in Drechsler and Yaron (2010), 'Extension 1' denotes the model where the OU process for $\bar{\sigma}^2$ is replaced by a square-root specification, 'Extension 2' is the model where additionally the jump intensity is an affine function of α instead of σ^2.

Figure 1:



The figure shows the level and the slope of the implied volatility smile for options on the S&P 500 index over the period from January 1996 to December 2006. The data are taken from OptionMetrics. 'Level' is defined as the implied volatility of an option with 30 days to maturity and a moneyness (defined as strike price divided by current stock price) of 1. 'Slope' is defined as the difference between the implied volatilities of two 30-day options, one with a moneyness of 0.975 and the other one with a moneyness of 1. The observation frequency is monthly.