Nonlinear Dynamics in Conditional Volatility*

Friedrich Lorenz
University of Münster

Karl Schmedders
IMD Lausanne and
University of Zurich

Malte Schumacher
University of Zurich

April 14, 2020

Abstract

Investors pay a substantial premium to hedge against fluctuations in volatility—the variance risk premium (VRP). The asset-pricing literature has presented numerous models with jumps in economic fundamentals to reproduce the properties and the time variation of the VRP. This paper shows that these quantitative results are almost exclusively driven by an inaccurate measure of conditional volatility. Solved accurately, conditional volatility exhibits—counterfactually—a strong procyclical pattern and the models do not deliver a sizeable VRP in response to jumps in state variables. The notion that the VRP is purely a compensation for fluctuations in macroeconomic uncertainty does not hold.

Keywords: asset pricing, conditional volatility, variance risk premium.

JEL codes: G11, G12.

*We are indebted to Marlon Azinovic, Nicole Branger, Winston Dou, Thomas Grünthaler, Ken Judd, Martin Lettau, Monika Piazzesi, Walt Pohl, Gregor Reich, David Schreindorfer, Roméo Tédongap, Julian Thimme, Ole Wilms, and Nancy Xu for helpful discussions on the subject. We are grateful to Ole Wilms for detailed comments on an earlier draft and to Dave Brooks for excellent editorial support. And we thank seminar audiences at the University of Zurich, the University of Münster, the joint Asset Pricing Workshop with the Goethe University Frankfurt, the 2019 ESSEC Workshop on Monte Carlo Methods and Approximate Dynamic Programming with Applications in Finance, and the 2019 SITE Summer Workshop.

1 Introduction

Financial markets display a "variance risk premium" (VRP) for assets whose payoffs are high when return variation is large. Due to its economically significant size and its predictive power for asset returns, the VRP has generated a large empirical and theoretical literature. One important strand of this literature argues that the VRP is a consequence of macroeconomic uncertainty. In particular, long-run risks (LRR) models with jumps in economic state variables have been shown to quantitatively capture moments of the VRP and its return predictability. In this paper, we show that these quantitative results are almost exclusively driven by an inaccurate measure of conditional volatility stemming from the Campbell–Shiller log-linearization solution approach. When capturing nonlinear dynamics in fundamentals, these models no longer deliver a sizeable VRP as a result of jumps—so, infrequent but large shocks to state variables. Put differently, the notion that the VRP is purely a compensation for fluctuations in macroeconomic uncertainty does not hold.

The log-linearization solution method of Campbell and Shiller (1988) leads to severe errors not only in the approximation of the VRP but also in the calculation of conditional return volatilities and their dynamics. To make matters worse, the true solution of the LRR model with jumps finds that conditional return volatilities exhibit—contrary to financial-market data—a strong procyclical behavior. Moreover, due to the strong correlation between risk premia and conditional volatility in these models (Nagel and Xu (2019)), the model also predicts counterfactually procyclical premia.[1] The procyclical behavior arises due to a time-varying weighting of cash flow and discount-rate volatility. In states with a higher price–dividend ratio, discount-rate volatility becomes more dominant, which leads to a higher conditional volatility of returns. Although a log-linear function may accurately approximate (logarithmic) price–dividend ratios, conditional volatility still exposes significant nonlinearities, which are not captured by the log-linear approximation. In a standard LRR model without jumps, we show that conditional volatility is increasing in expected consumption growth if the agent has a preference for early resolution of uncertainty. Furthermore, we show that approximation errors increase dramatically for models with jumps in economic fundamentals; additionally, jumps increase the magnitude of the counterfactual procyclical return volatility.

The VRP is one of the most prominent asset-pricing puzzles because of its economically significant size (Coval and Shumway (2001), Carr and Wu (2009)) and predictive power for

[1] See, for example, Fama and French (1989), who find that risk premia are especially high during recessions. On the contrary, risk premia decrease in LRR models if expected consumption growth falls.

excess returns (Bollerslev, Tauchen, and Zhou (2009)). Bollerslev, Tauchen, and Zhou (2009) introduced the notion of the variance risk premium into equilibrium models of asset returns on financial markets. They add a volatility process into the consumption growth volatility in the LRR framework of Bansal and Yaron (2004). In their theoretical model development, Bollerslev, Tauchen, and Zhou (2009) show that this "volatility-of-volatility" factor is among the drivers of the model-implied equity premium. Moreover, they manage to isolate this factor as the sole driver of the difference between the risk-neutral and the actual expected variance of one-period asset returns, which they term the variance risk premium. These two theoretical results lead to the testable model implication that the variance risk premium can forecast equity returns. The key empirical finding of Bollerslev, Tauchen, and Zhou (2009) is that, indeed, the variance risk premium can explain a sizable fraction of the market excess return.

As the VRP is a measure derived from option prices, borrowing from the numerous reduced-form models on the option-implied volatility smirk seems like a natural way to go. Besides stochastic volatility, there is a substantial literature building on jumps included into state variables.[2] One of the first general equilibrium models to adopt this approach is that of Liu, Pan, and Wang (2005). They show that to generate an option smirk, rare events and recursive utility are not sufficient to match the empirical option-implied volatility. They claim that no matter the sophistication with which the physical dynamics are modeled, the generated smirks are not able to fit the empirical ones. Benzoni, Collin-Dufresne, and Goldstein (2011) build on the LRR framework of Bansal and Yaron (2004) and extend it with rare jumps and learning about the probability of future jumps to explain the behavior of the option-implied volatility smirk around the 1987 market crash. Eraker and Shaliastovich (2008) explain how models with recursive preferences, in which the state variables follow affine jump–diffusion processes, can be solved after applying a continuous-time version of the Campbell–Shiller log-linearization. They present an example of how a model with jumps in volatility can produce an implied volatility smirk as in the data.

In a landmark paper, Drechsler and Yaron (2011) transfer these ideas to discrete time and claim to explain the puzzles regarding the VRP in an LRR model with stochastic volatility and jumps in state variables. The key ingredient for explaining the VRP is the inclusion of jumps with a time-varying probability into both the expected growth rate and the stochastic volatility process. In their setup, the VRP is tightly linked to the stochastic volatility. Hence, the paper supports the line of argument that the VRP is a compensation for time-varying macroeconomic uncertainty (Bollerslev, Tauchen, and Zhou (2009), Drechsler and Yaron (2011)). On the

[2]See, for example, Merton (1976), Bakshi, Cao, and Chen (1997), Duffie, Pan, and Singleton (2000)

contrary, another strand of literature argues that the VRP rather reflects time-varying risk aversion (Bekaert and Hoerova (2016), Bekaert, Engstrom, and Xu (2019)). We contribute to the literature by excluding the time-varying macroeconomic uncertainty in its current interpretation in LRR models as an explanation for the VRP. Our results challenge the widely adopted notion that the VRP is a compensation for time-varying macroeconomic uncertainty as in the model of Drechsler and Yaron (2011).[3]

Furthermore, the results in this paper show that the Campbell–Shiller log-linearization—employed by Drechsler and Yaron (2011) to solve their model—does lead to extreme errors if a model includes jumps. The effect of log-linearizing consumption-based asset-pricing models without jumps has been studied by Pohl, Schmedders, and Wilms (2018). The accuracy of the solutions deteriorates with higher persistence of state variables or risk aversion. Even if the elasticity of intertemporal substitution is equal to 1—resulting in a constant wealth–consumption ratio—the errors stemming from the log-linearization of the price–dividend ratio can still be substantial.[4] Pohl, Schmedders, and Wilms (2018) focus on the equity premium and predictability tests in models with conditional Gaussian shocks. The present paper is the first to examine the impact of jumps on the volatility approximation—under the physical as well as the risk-neutral measure. To this end, we show that even if the wealth–consumption and price–dividend ratios are log-linear, log-linearization massively overstates the VRP. If both ratios exhibit significant nonlinearities—as in the model of Drechsler and Yaron (2011)—the VRP is almost exclusively driven by errors induced by the log-linearization. Hence we show that including asymmetric shocks—through either jumps or non-Gaussian distributions—into endowment models necessitates the use of adequate solution techniques.

The remainder of this paper is organized as follows. Section 2 introduces the model of Drechsler and Yaron (2011), and documents how the log-linearization distorts the dynamics of conditional volatility. In Section 3 we first analyze the errors in a stylized model with jumps in the expected growth rate only but without a stochastic volatility process. In Section 4 we compare the results of the full model when solved with log-linearization against the results stemming from a global solution and examine the sources of errors in the VRP. Section 5 shows that overcoming the shortcomings of the model is a nontrivial task and points to some further implications of the model, which only occur if the model is solved accurately. Section 6 concludes.

[3]See, for example, Bollerslev and Todorov (2011), Kelly and Jiang (2014), Dew-Becker, Giglio, Le, and Rodriguez (2017), Kilic and Shaliastovich (2018), or Pyun (2019), all of which who cite Drechsler and Yaron (2011) accordingly.

[4]See Tsai and Wachter (2018) or Pohl, Schmedders, and Wilms (2018).

2 Macroeconomic Uncertainty and the Variance Risk Premium

Following Drechsler and Yaron (2011), we consider a standard consumption-based asset-pricing model in discrete time. The representative agent has recursive preferences as in Epstein and Zin (1989) and Weil (1990). Utility is defined as

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

where C_t is consumption at time t, $0 < \delta < 1$ is the time discount factor, γ determines the level of risk aversion, ψ is the elasticity of intertemporal substitution, and $\theta = \frac{1-\gamma}{1-\frac{1}{\psi}}$. The agent has a preference for the early resolution of uncertainty whenever $\gamma > 1/\psi$ and a preference for the late resolution for $\gamma < 1/\psi$. The (unobserved) value of the agent's wealth, W_t, satisfies the budget equation

$$W_{t+1} = (W_t - C_t)\, R_{t+1}^c,$$

where R_t^c denotes the return on wealth.

In this economy, the gross return on any asset i with ex-dividend price $P_{i,t}$ and dividend $D_{i,t}$ is given by $R_{t+1}^i = \frac{P_{i,t+1}+D_{i,t+1}}{P_{i,t}}$. We denote the natural logarithm of this return by $r_{t+1}^i = \ln\left(R_{t+1}^i\right)$.

Epstein and Zin (1989) derive necessary and sufficient optimality conditions for the agent's constrained utility maximization problem. The first-order condition for the return on any asset i states

$$E_t\left[\exp\left(m_{t+1} + r_{t+1}^i\right)\right] = 1, \tag{2}$$

where m_{t+1} denotes the natural logarithm of the stochastic discount factor,

$$m_{t+1} = \theta \ln(\delta) - \frac{\theta}{\psi}\Delta c_{t+1} + (\theta - 1) r_{t+1}^c, \tag{3}$$

with $\Delta c_{t+1} = \ln(C_{t+1}) - \ln(C_t)$. Using the logarithm of the wealth–consumption ratio, $z_t^c = \ln((W_t - C_t)/C_t)$, and the transformation

$$r_{t+1}^c = \ln\left(\exp\left(z_{t+1}^c\right) + 1\right) - z_t^c + \Delta c_{t+1}$$

we can rewrite the first-order condition (2) for the return on wealth as

$$E_t\left[\exp\left(\theta\Big(\ln(\delta)+\left(1-\frac{1}{\psi}\right)\Delta c_{t+1}-z_t^c+\ln\left(\exp\big(z_{t+1}^c\big)+1\right)\Big)\right)-1\right]=0. \tag{4}$$

Using the log price–dividend ratio, $z_t^i=\ln\left(\frac{P_{i,t}}{D_{i,t}}\right)$, we can rewrite Equation (2) for asset i as

$$\begin{aligned}E_t\Bigg[\exp\Bigg(&\theta\ln(\delta)-\frac{\theta}{\psi}\Delta c_{t+1}+(\theta-1)\left(\ln\left(\exp\big(z_{t+1}^c\big)+1\right)-z_t^c+\Delta c_{t+1}\right)\\&-z_t^i+\Delta d_{i,t+1}+\ln\left(\exp\big(z_{t+1}^i\big)+1\right)\Bigg)-1\Bigg]=\ 0,\end{aligned} \tag{5}$$

where $\Delta d_{i,t+1}=\ln(D_{i,t+1})-\ln(D_{i,t})$.

The VRP is defined as the difference between the risk-neutral and physical conditional variance of the dividend claim:[5]

$$VRP_t=\ Var_t^{\mathbb{Q}}[r_{t+1}^d]-\ Var_t[r_{t+1}^d]. \tag{6}$$

Whenever there is no superscript we refer to the physical measure, while the superscript $\mathbb{Q}$ stands for the risk-neutral measure. To compute the risk-neutral measure, we make use of the Radon–Nikodym derivative $\frac{d\mathbb{Q}}{d\mathbb{P}}=\frac{M_{t+1}}{E_t[M_{t+1}]}$. Hence, under the risk-neutral measure probability mass is shifted to bad states—that is, states with a high marginal utility are more likely than under the physical measure.

2.1 Dynamics of the State Vector

We now describe the dynamics of the state variables specified by Drechsler and Yaron (2011). We follow Branger, Rodrigues, and Schlag (2018) and express the explicit state dynamics.[6] Although the matrix representation of Drechsler and Yaron (2011) offers great flexibility and we use their exact calibration, we want to show the dynamics as clearly as possible. First, consumption and dividends—which are a levered consumption claim following Abel (1999)—are modeled containing a small long-run predictable growth rate x_t and time-varying macroeco-

[5]This definition corresponds to what Drechsler and Yaron (2011) coin as the level difference. We discuss the choice of this definition in Appendix A.3.

[6]To map them into the matrix notation of Drechsler and Yaron (2011) we refer to the online appendix of Branger, Rodrigues, and Schlag (2018).

nomic uncertainty $\sigma_{i,t},\ i \in \{c,d\}$:

$$\begin{aligned}\Delta c_{t+1} &= \mu_c + x_t + \sigma_{c,t} e_{c,t+1} \\ \Delta d_{t+1} &= \mu_d + \Phi x_t + \sigma_{d,t}\left(\rho_t e_{c,t+1} + \sqrt{1-\rho_t^2}\, e_{d,t+1}\right),\end{aligned} \tag{7}$$

where innovations to dividend growth are correlated with innovations to consumption growth with a time-varying correlation coefficient ρ_t. Next, the state variables follow the dynamics

$$\begin{aligned}x_{t+1} &= \rho x_t + \phi_x \sigma_t e_{x,t+1} + J^x_{t+1} - E[\xi^x]\lambda_{x,t} \\ \sigma^2_{t+1} &= \bar{\sigma}^2_t + \nu_\sigma(\sigma^2_t - \bar{\sigma}^2_t) + \phi_\sigma \sigma_t e_{\sigma,t+1} + J^\sigma_{t+1} - E[\xi^\sigma]\lambda_{\sigma,t} \\ \bar{\sigma}^2_{t+1} &= \mu_{\bar{\sigma}} + \nu_{\bar{\sigma}} \bar{\sigma}^2_t + \sigma_{\bar{\sigma},t} e_{\bar{\sigma},t+1} \\ & e_{c,t+1}, e_{d,t+1}, e_{x,t+1}, e_{\sigma,t+1}, e_{\bar{\sigma},t+1} \sim \mathcal{N}(0,1),\end{aligned} \tag{8}$$

where $J^k_{t+1},\ k \in \{x,\sigma\}$, are compound Poisson processes with i.i.d. random jump sizes, ξ_k, and intensity $\lambda_{k,t}$. Compensating the Poisson processes by adjusting the drift through the conditional expected value, $E_t[J^k_{t+1}] = E[\xi^k]\lambda_{k,t}$, ensures that the expected jumps do not enter the drift of the state variables. The volatilities of Δc_{t+1}, Δd_{t+1}, and $\bar{\sigma}_{t+1}$ depend on the parameters $w_c, w_d,$ and $w_{\bar{\sigma}}$, respectively:

$$\begin{aligned}\sigma^2_{i,t} &= \phi_i(w_i \sigma^2_t + 1 - w_i), \quad i \in \{c,d\} \\ \sigma^2_{\bar{\sigma},t} &= \phi_{\bar{\sigma}}(w_{\bar{\sigma}} \bar{\sigma}^2_t + 1 - w_{\bar{\sigma}}).\end{aligned} \tag{9}$$

Setting $w_{\bar{\sigma}} = 1$ as in the calibration of Drechsler and Yaron (2011) makes the process the equivalent of an Ornstein–Uhlenbeck process in discrete time, which is the benchmark calibration employed by Drechsler and Yaron (2011). Since the conditional volatility of consumption growth as well as the conditional volatility of dividend growth depend solely on σ^2_t, we refer to the latter as macroeconomic uncertainty. The conditional correlation ρ_t between consumption and dividend shocks is given by the conditional covariance $\sigma_{cd,t}$, which is given by

$$\sigma_{cd,t} = \omega_{cd}\phi_c\phi_d\left(\sqrt{\omega_c\omega_d}\sigma^2_t + \sqrt{(1-\omega_c)(1-\omega_d)}\right). \tag{10}$$

For the distributions of the jump sizes, ξ_k, we use the specification of Drechsler and Yaron (2011) with non-Gaussian (jump) shocks in the volatility process σ^2_t and normally distributed jump sizes for the long-run risk process x_t,

$$\xi_\sigma \sim \Gamma(a_\sigma, \frac{b_\sigma}{a_\sigma}), \quad \xi_x \sim \mathcal{N}(0, \phi_{\xi^x}). \tag{11}$$

With this parameterization of the Gamma distribution, the expected jump size in σ_{t+1} is simply b_σ. In Equation (8), $\lambda_{k,t}$ denotes the intensity of the Poisson counting process N^k, for $k \in \{x, \sigma\}$, which is affine in the variance:

$$\begin{aligned} J_{t+1}^k &= \sum_{j=1}^{N_{t+1}^k} \xi_k^j, \quad k \in \{x, \sigma\} \\ \lambda_{k,t} &= l_{0,k} + l_{1,k}\sigma_t^2. \end{aligned} \tag{12}$$

2.2 Conditional Volatility Approximation

As typically done in LRR models, Drechsler and Yaron (2011) use the Campbell–Shiller approximation to solve the model:

$$\begin{aligned} r_{t+1}^i &= \log\big(\exp\big(z_{t+1}^i\big) + 1\big) - z_t^i + \Delta i_{t+1} \\ r_{t+1}^i &\approx \kappa_0^i + \kappa_1^i z_{t+1}^i - z_t^i + \Delta i_{t+1}, \quad i \in \{c, d\}, \end{aligned} \tag{13}$$

where z_t^c denotes the log wealth–consumption ratio and z_t^d is the log price–dividend ratio of the levered consumption claim in Equation (7). Assuming that the approximation in Equation (13) holds, the wealth–consumption ratio as well as the price–dividend ratio are log-linear functions of the state variables:

$$\begin{aligned} z_t^c &= A_0 + A_x x_t + A_\sigma \sigma_t^2 + A_{\bar\sigma} \bar\sigma_t^2 \\ z_t^d &= A_{0,d} + A_{x,d} x_t + A_{\sigma,d} \sigma_t^2 + A_{\bar\sigma,d} \bar\sigma_t^2. \end{aligned} \tag{14}$$

Only by applying this approximation does the model fall into the affine class, the solution to this approximation being exact for all points in the state space.[7] By way of contrast, the exact return definition leads to nonlinearities that give rise to significant approximation errors in the log-linear solution. In this paper, we focus mainly on the resulting errors for conditional volatilities and the VRP. The log-linearization coefficients k_0 and k_1 are defined as

$$\begin{aligned} \kappa_1^i({z_t}^i) &= \frac{\exp({z_t}^i)}{1 + \exp({z_t}^i)} \\ \kappa_0^i({z_t}^i) &= \log\big(1 + \exp\big({z_t}^i\big)\big) - \kappa_1^i {z_t}^i, \quad i \in \{c, d\}, \end{aligned} \tag{15}$$

where ${z_t}^i$ is the point of expansion. To solve the model using log-linearization, the point of expansion has to be fixed at some level—typically the steady state, which we denote by $\bar z^i$.

[7] For details of the log-linear solution, we refer to the papers of Eraker (2008), Eraker and Shaliastovich (2008), and Drechsler and Yaron (2011).

In this case, the log-linearization constants are fixed at $\kappa_0^i(z_t{}^i) = \kappa_0^i(\bar{z}^i)\ \forall\ t$ and $\kappa_1^i(z_t{}^i) = \kappa_1^i(\bar{z}^i)\ \forall\ t$, in which case we drop the dependence on the point of expansion and simply write κ_0^i and κ_1^i. Hence, the log-linearized conditional volatility is given by

$$Var_t[r_{t+1}^i] = (\kappa_1^i)^2\, Var_t[z_{t+1}^i] + Var_t[\Delta i_{t+1}], \quad i \in \{c, d\}. \tag{16}$$

Consequently, the time variation in the conditional volatility is solely driven by the conditional volatilities of the price–dividend ratio and the cash-flow volatility. The level of the price–dividend ratio, on the other hand, does not affect the conditional volatility. If, however, we do not fix the point of expansion but instead take the Taylor expansion around the conditional value of the price–dividend ratio, the variance of returns would be a nonlinear function of the point of expansion, of the conditional volatility of the price–dividend ratio, and of the conditional volatility of dividend growth:[8]

$$Var_t[r_{t+1}^i] = (\kappa_1^i(z_t^i))^2\, Var_t[z_{t+1}^i] + Var_t[\Delta i_{t+1}], \quad i \in \{c, d\}. \tag{17}$$

Hence, the weight of the discount-rate variation, $Var_t(z_{t+1}^i)$, increases for higher values of the price–dividend ratio. Only if expected discount-rate variation and expected cash flow variation, $Var_t(\Delta i_{t+1})$, are negatively correlated might the effect cancel out. On the other hand, news about the long-run growth rate affects the level of the price–dividend ratio, but not the conditional volatility of the price–dividend ratio or dividends. Consequently, such news changes the relative weight of the expected discount rate variation in the conditional return variation equation, Equation (17), and leads to the conditional volatility of returns depending on the level of the price–dividend ratio.[9] Notably, the expected consumption growth rate x_t in LRR models has exactly this effect. To see the effect in closed form, we derive the conditional volatility in the model of Bansal and Yaron (2004) and allow it to be state dependent.[10] In Appendix A.1 we show that if the agent has a preference for the early resolution of uncertainty, the conditional variance of returns is increasing in the long-run growth rate x_t. As the macroeconomic uncertainty σ_t^2 enters the conditional variance of the price–dividend ratio as well as that of dividend growth positively, the conditional return variance is—as expected—increasing in σ_t^2. For a more detailed examination, however, we

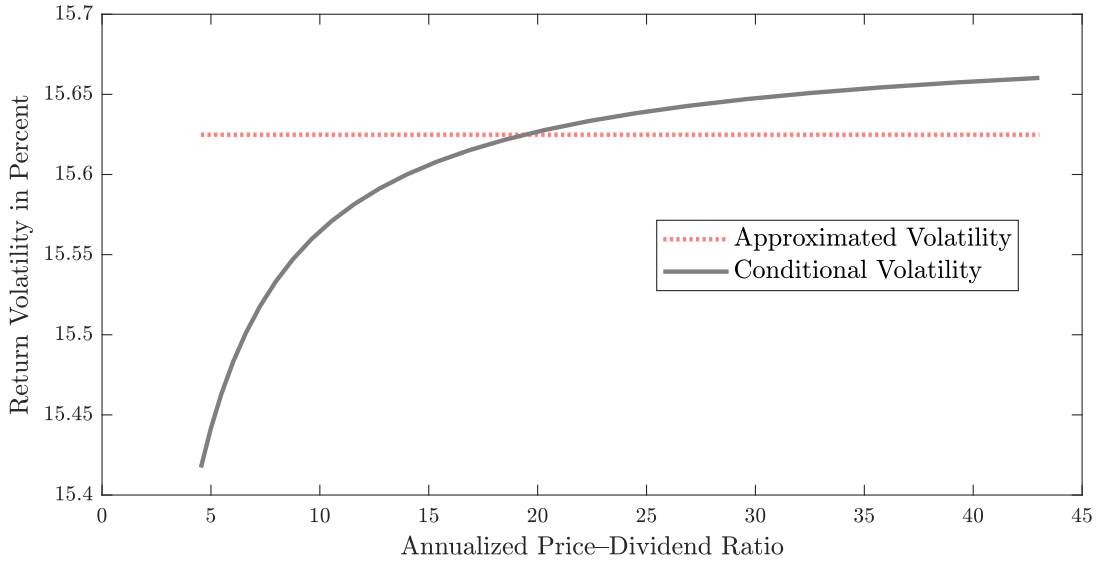
[8] We drop $Cov_t[z_{t+1}^i, \Delta d_{t+1}^i]$ as dividend and state variable shocks are assumed to be mutually independent (see Bansal and Yaron (2004) or Drechsler and Yaron (2011)).

[9] This effect holds for the levered as well as the unlevered consumption claim.

[10] Note that the model of Bansal and Yaron (2004) is nested by the model of Drechsler and Yaron (2011). As the derivations only assume preference for the early resolution of uncertainty, we expect the various partial derivatives of conditional volatility to be of the same sign in both models.

derive the second derivative of conditional return variance with respect to σ_t^2. Adding up the parts shows that the second derivative is negative and not zero as implied by the log-linearization—that is, volatility is, ceteris paribus, an increasing and concave function of σ_t^2. As LRR models emphasize the importance of discount-rate news compared to cash flow news, the discount-rate channel in conditional volatility contributes to a larger extent to the conditional volatility of returns. The discount-rate channel's weight is, however, reduced as macroeconomic uncertainty rises, due to a lower $\kappa_1^i(z_t^i)$.

Figure 1: Volatility Approximation



This figure shows the conditional volatility of the return on the dividend claim depending on the current level of the price–dividend ratio. The dotted red line denotes the volatility with a constant value of the price–dividend ratio as the point of expansion. The solid black line denotes the conditional volatility where the current level of the price–dividend ratio is used as the point of expansion. We use $Var_t[\Delta z_{t+1}^d] = 0.154$, $Var_t[\Delta d_{t+1}] = 0.03$, and $E[z^d] = 5.45$, which are close to the unconditional values of the model of Bansal and Yaron (2004).

To show how the Cambell–Shiller approximation can lead to errors in the variance of returns, we look at a simple example. In Figure 1 we plot the conditional volatility as a function of the current level of the price–dividend ratio. In this example, we only vary the level of the price–dividend ratio, keeping the conditional volatility of both the price–dividend ratio and dividends unchanged. The dotted line denotes the volatility if we follow the typical approach solving LRR models and keep the point of expansion fixed at the unconditional mean of the price–dividend ratio $E[z^i]$ and consequently with constant κ_1^i. On the other hand, the solid line plots the true conditional volatility where we change the point of expansion to the actual value of the price–dividend ratio z_t^i—that is, $\kappa_1^i(z_t^i)$. We find that the latter leads to a conditional volatility that is increasing with the price–dividend ratio. A fixed point of expansion, on the other hand, implies a constant conditional volatility. For bad states—that is, for a low

price–dividend ratio—keeping the point of expansion fixed overestimates conditional volatility compared to its true value. On the other hand, a fixed point of expansion underestimates the conditional volatility in good states.[11] Consequently, the physical volatility is already badly approximated by log-linearization. Due to the change of measure, which is also prone to approximation errors, the error might even be further amplified in the risk-neutral conditional volatility.

To see how these effects accumulate in general equilibrium, we first solve a stylized LRR model with jumps and then the more complicated model of Drechsler and Yaron (2011) globally using projection methods. While the approximation of Gaussian shocks is standard, the integration of Gamma-distributed variables including one or more compound Poisson processes is new. The approximation of the jump processes is, in particular, at the heart of our paper. In contrast to Gaussian innovations, which are well studied, the solution to models involving jumps is still at an early stage; see, for example, Fernández-Villaverde and Levintal (2018). To integrate the Gamma-distributed shocks, we rely on Gauss–Laguerre quadrature. For the Gaussian shocks, we use monomials to keep the problem tractable. For details on the model solution and accuracy checks see Appendix B.

3 Results of a Stylized Model

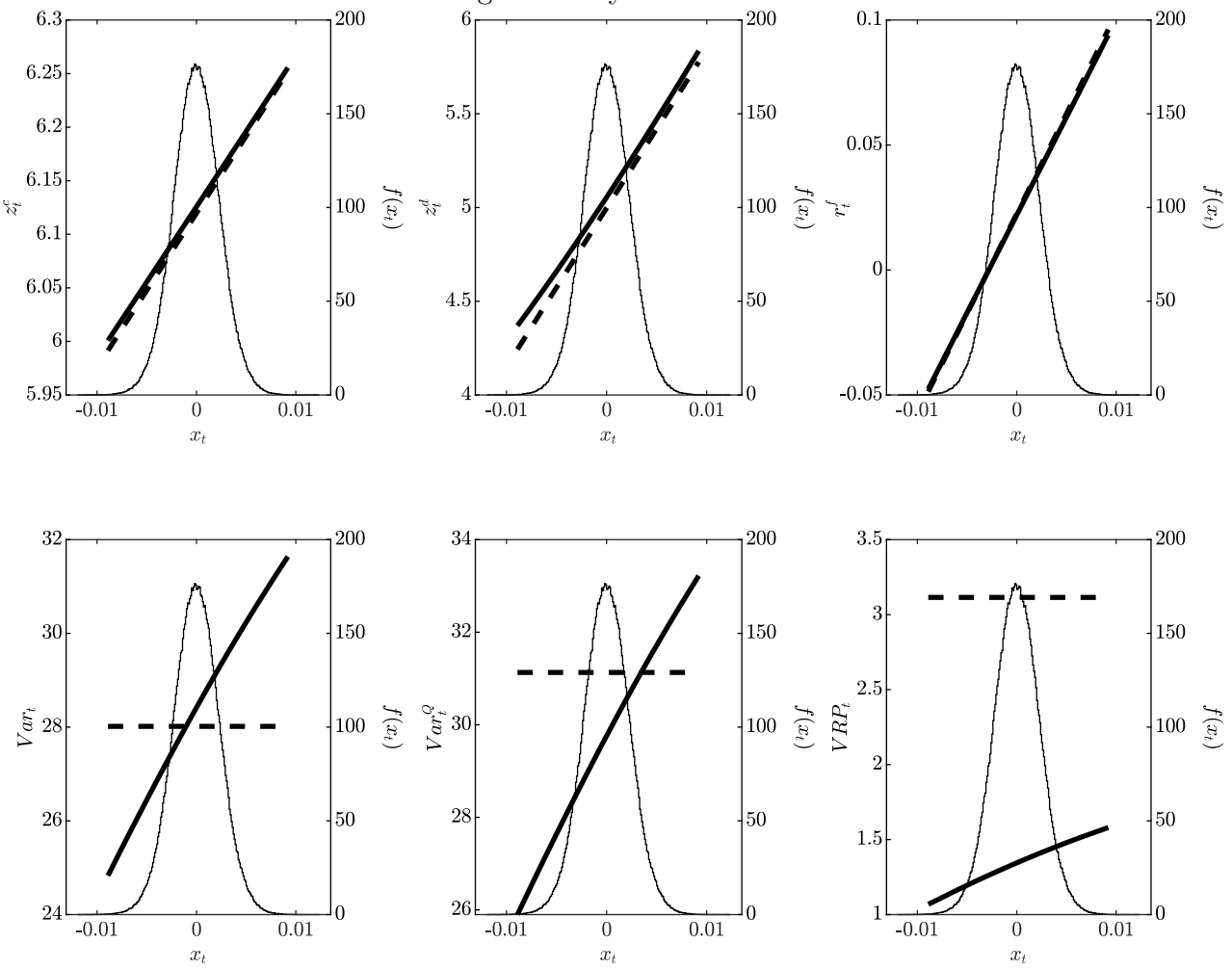
We start the analysis by examining the effects of the nonlinearities in conditional volatility when the underlying log price–dividend ratio is well approximated by the log-linearized solution. First, we reduce the model to one without stochastic volatility and time-varying jump intensity. To this end, we use the calibrated Case I of Bansal and Yaron (2004), which is nested by the dynamics of Section 2.1, but add a compound Poisson process with constant jump intensity and normally distributed shocks as in Drechsler and Yaron (2011). Hence, the only time variation in the conditional volatility must come from the aforementioned time-varying weighting in Equation (17). To arrive at this setting, we set $\delta = 0.998$, $\gamma = 10$, $\psi = 1.5$, $\mu_c = 0.0016$, $\rho = 0.979$, $\mu_d = 0.0016$, $\Phi = 3$, $\sigma_t = 0.0078 \; \forall \; t$, $\omega_{cd} = 0$, $\omega_c = 1$, $\omega_d = 1$, $\phi_c = 1$, $\phi_d = 4.5$, $\phi_x = 0.044$, $l_{0,x} = 0.8/12$, and $l_{1,x} = 0$. In contrast to the full model of Drechsler and Yaron (2011), this parameterization also allows for closed-form log-linearized solutions, which we report in Appendix A.2.

In the upper row of Figure 2, we show the log wealth–consumption ratio, the log price–

[11] Note that this example serves only as an illustration of how the mechanism works and that the quantitative effects should not be interpreted. We examine the quantitative effect in the fully calibrated model in the next sections.

dividend ratio, and the annualized risk-free rate as a function of the single state variable x_t together with a histogram of x_t obtained with the global (solid line) and log-linear (dashed line) solution. For these three quantities, we observe that all are close to linear even for the global solution and that the differences between the log-linearized and the global solution are small across the most relevant areas of the state space. This might lead one to conclude that log-linearization is appropriate in this setup and to continue solving for other quantities. The bottom row shows the physical conditional variance, the risk-neutral conditional variance, and the VRP as functions of the single state variable x_t together with a histogram of x_t. For these quantities, the picture changes dramatically compared to the graphs in the upper row. While the volatility obtained by the log-linearized solution (dashed line) is constant, the global solution (solid line) captures the nonlinearities and therefore the conditional volatility is strongly procyclical. The middle plot shows that this also holds true for the conditional volatility of the return on equity under the risk-neutral measure. The log-linearized solution implies a constant conditional volatility while the true conditional volatility is procyclical. Most notably, the true VRP—depicted in the bottom right plot—is not only procyclical but also significantly lower than the constant VRP implied by the log-linearized solution. Hence, the log-linearized solution not only inflates the VRP but also conceals the qualitatively wrong dynamics of conditional volatility and the VRP.

Figure 2: Stylized Model



On the left axes, the graphs show the log-linearized (dashed line) and the global (solid line) solution for the log wealth–consumption ratio, the log price–dividend ratio, the annualized risk-free rate, the physical and risk-neutral expected variance, and the VRP in the stylized model for different values of the single state variable x_t. On the right axes, the graphs show the simulated density function of x_t. All parameters except for J_{t+1}^x, which we calibrate following Drechsler and Yaron (2011), are from Bansal and Yaron (2004) and are given by $\delta = 0.998$, $\gamma = 10$, $\psi = 1.5$, $\mu_c = 0.0016$, $\rho = 0.979$, $\mu_d = 0.0016$, $\Phi = 3$, $\sigma_t = 0.0078 \; \forall \; t$, $\omega_{cd} = 0$, $\omega_c = 1$, $\omega_d = 1$, $\phi_c = 1$, $\phi_d = 4.5$, $\phi_x = 0.044$, $l_{0,x} = 0.8/12$, $l_{1,x} = 0$, and $\phi_{\xi_x} = 3.645\phi_x$.

Next we examine the quantitative predictions of the log-linearized solution against the true solution. Table 1 reports annualized summary statistics for the log-linear and the global solution. First, the risk-free rate, the equity premium, and the log price–dividend ratio are—as expected from the plots in Figure 2—approximated rather well by the log-linear solution. The same holds for the physical conditional volatility, which is slightly understated by the log-linearized solution. In contrast, the risk-neutral conditional volatility is slightly overstated (31.13 instead of 29.70). This slight error in the level of the conditional volatilities, however, leads to a decrease in the VRP of more than 50 percent. Furthermore, while the conditional volatilities are constant using the log-linear solution, they are time varying using the correct solution. Put together, log-linearization does still deliver accurate point estimates for most

Table 1: Summary Statistics of the Stylized Model

Moment	Data		Model					
			Log-linear			Global		
	Est.	SE.	5%	50%	95%	5%	50%	95%
$E[r^f]$	0.72	0.33	1.37	2.27	3.08	1.33	2.22	3.00
$\sigma[r^f]$	3.11	0.38	1.26	1.66	2.17	1.23	1.63	2.12
$E[r^d - r^f]$	5.28	2.13	4.81	7.95	11.11	4.36	7.56	10.80
$\sigma[r^d]$	19.46	1.70	16.19	18.61	21.18	15.93	18.30	20.85
$E[z^d]$	3.41	0.05	2.41	2.52	2.63	2.46	2.57	2.68
$\sigma[z^d]$	45.21	3.07	17.69	22.70	29.26	17.19	21.96	28.03
$AC1[z^d]$	0.88	0.05	0.53	0.69	0.81	0.52	0.69	0.81
$skew[z^d]$	0.23	0.18	-0.56	-0.01	0.56	-0.51	0.03	0.61
$E[Var_t]$	31.72	3.30	28.02	28.02	28.02	27.93	28.36	28.74
$\sigma[Var_t]$	50.44	14.57	0.00	0.00	0.00	0.64	0.83	1.07
$E[Var_t^{\mathbb{Q}}]$	41.07	2.86	31.13	31.13	31.13	29.23	29.70	30.11
$\sigma[Var_t^{\mathbb{Q}}]$	43.76	5.27	0.00	0.00	0.00	0.69	0.89	1.15
$E[VRP_t]$	9.35	2.07	3.11	3.11	3.11	1.31	1.34	1.37
$\sigma[VRP_t]$	31.62	7.23	0.00	0.00	0.00	0.05	0.06	0.08

The table reports the summary statistics of the risk-free rate, return on equity, and the log price–dividend ratio for the stylized model. We follow Beeler and Campbell (2012) and Drechsler and Yaron (2011) to calculate and time-aggregate all moments. The first two columns display the estimated mean and standard error of US data for the sample period from 1930 to 2016, the next three columns the moments implied by the model solved with log-linearization, and the last three columns the moments implied by the global solution. The model-implied moments are the 5%, 50%, and 95% quantile values from 1,000 simulations with the same length as the data sample, where we use the same seed for each calibration to ensure comparability. All parameters except for J^x_{t+1}, which we calibrate following Drechsler and Yaron (2011), are from Bansal and Yaron (2004) and are given by $\delta = 0.998$, $\gamma = 10$, $\psi = 1.5$, $\mu_c = 0.0016$, $\rho = 0.979$, $\mu_d = 0.0016$, $\Phi = 3$, $\sigma_t = 0.0078\ \forall\ t$, $\omega_{cd} = 0$, $\omega_c = 1$, $\omega_d = 1$, $\phi_c = 1$, $\phi_d = 4.5$, $\phi_x = 0.044$, $l_{0,x} = 0.8/12$, $l_{1,x} = 0$, and $\phi_{\xi_x} = 3.645\phi_x$.

quantities even if the underlying state variables can jump. Nevertheless, log-linearization does not capture the dynamics of conditional volatility and leads to severe errors in the VRP. In contrast to the conclusion of Bansal and Yaron (2004), even the model without time-varying economic uncertainty leads to time-varying risk premia and conditional volatilities.

4 Results of the Full Model

The stylized model already exemplifies that jumps in state variables can lead to a deteriorating accuracy of log-linearization for conditional volatility. We now turn to the more complicated parameterization of Drechsler and Yaron (2011). This allows us to examine the additional effect of nonlinearities in the underlying function on the approximation of the VRP. To do so, we continue the analysis by comparing summary statistics implied by the calibration of Drechsler and Yaron (2011) solved with log-linearization to the results of the true solution, which is based on highly accurate numerical methods. We closely follow Drechsler and Yaron (2011) to solve, simulate, and time-aggregate the model.

4.1 Parameterization of Drechsler and Yaron (2011)

Table 2 reports the parameter values of Drechsler and Yaron (2011). The model is solved at a monthly frequency. They calibrate their model to match consumption and dividend data, the equity premium, the risk-free rate, the VRP and its predictability, and predictability by the price–dividend ratio.

Table 2: Parameters of Drechsler and Yaron (2011)

	Parameters				
Preferences	δ	ψ	γ		
	0.999	2	10		
Economy	μ_c	ϕ_c	ω_c	ρ	ϕ_x
	0.0016	0.0066	0.5	0.976	2.1e-04
	$l_{1,x}$	ϕ_{ξ^x}	ν	ϕ_σ	a_σ
	0.8/12	7.7e-04	0.87	0.35	1
	b_σ	$l_{1,\sigma}$	$\mu_{\bar{\sigma}}$	$\nu_{\bar{\sigma}}$	$\phi_{\bar{\sigma}}$
	2.55	0.8/12	0.015	0.985	0.1
	μ_d	Φ	ϕ_d	ω_d	ω_{cd}
	0.0016	2.5	0.0376	0.125	0.2

The table reports the parameters of the model of Drechsler and Yaron (2011). All parameters are in monthly terms as the model is solved with a monthly decision horizon.

Looking only at the persistence and the risk aversion, the parameterization does not imply high inaccuracies (see Pohl, Schmedders, and Wilms (2018)). The persistence of x is slightly lower than the persistence in the calibration by Bansal and Yaron (2004). Furthermore, neither volatility nor the long-run mean of volatility are as persistent as in the calibration by Bansal, Kiku, and Yaron (2012). Although the additional process for the long-run mean of volatility might, even without a high persistence, lead to higher nonlinearities, that is not necessarily the

case.[12] Hence, any potential inaccuracies must stem from the addition of jumps to the model. To interpret the jump intensity more easily, Drechsler and Yaron (2011) set the unconditional means of both $\bar{\sigma}_t^2$ and σ_t^2 to 1. Thus, both the volatility, σ_t^2, as well as the expected growth rate, x_t, have on average $0.8/12 = 0.0667$ jumps per month or almost one jump per year. Consequently, the frequency of jumps is far higher than in the disaster risk literature.[13] The mean jump size in volatility is set to 2.55 and the volatility of jumps in expected growth to 7.7e-04. By setting a_σ to 1, the gamma distribution reduces to an exponential distribution which implies a variance of 6.5 for the jump size in σ_t^2.

Table 3 shows that the model can match the mean, volatility, and autocorrelation of consumption and dividends well. As the jumps do not enter consumption or dividends directly, but only via the expected growth rate x_t and the macroeconomic uncertainty σ_t^2, the effect on higher moments of consumption is not visible in contrast to disaster-risk models. This results in zero skewness for both consumption and dividend growth. All in all, except for skewness, the model does match the macroeconomic moments of the extended time series quite well.

Table 3: Summary Statistics of Macroeconomic Moments

Moment	**Data**		**Model**		
	Est.	SE.	5%	50%	95%
$E[\Delta c]$	1.83	0.23	1.12	1.94	2.75
$\sigma[\Delta c]$	2.11	0.30	1.83	2.32	3.02
$AC1[\Delta c]$	0.48	0.14	0.25	0.44	0.61
$skew[\Delta c]$	-1.48	0.58	-0.82	0.00	0.81
$E[\Delta d]$	1.44	1.20	-0.92	1.88	4.69
$\sigma[\Delta d]$	11.14	1.62	9.67	11.18	13.05
$AC1[\Delta d]$	0.19	0.15	0.11	0.28	0.45
$skew[\Delta d]$	-0.85	0.78	-0.46	0.01	0.47

The table reports the summary statistics of the annualized consumption and dividend growth moments for the parameter values denoted in Table 2. We follow Beeler and Campbell (2012) and Drechsler and Yaron (2011) in the calculation and aggregation of all moments. The first two columns display the estimated mean and standard error of US data for the sample period from 1930 to 2016, and the last three columns the moments implied by the model. The model-implied moments are the 5%, 50%, and 95% quantile values from 1,000 simulations with the same length as the data sample, where we use the same seed for each calibration to ensure comparability.

[12] See the model by Bollerslev, Xu, and Zhou (2015), who add volatility-of-volatility instead of a long-run mean of volatility. Still, their calibration, with similar values of persistence as those of Drechsler and Yaron (2011), does not result in significant errors due to log-linearization (see Pohl, Schmedders, and Wilms (2018)).

[13] For example, the mean intensity in Wachter (2013) is set to 0.0355 yearly, which is 0.003 monthly.

4.2 Asset Pricing

As a first check, we examine the implications of the solution method for standard asset-pricing moments. In Table 4, we compare selected moments of the risk-free rate, the return on equity, and the price–dividend ratio for the log-linear and the true solution. If solved by log-linearization, the model can create a small and slightly volatile risk-free rate. Solved correctly, the risk-free rate, in contrast, is more than 30% higher and 23% less volatile. The equity premium, the volatility of the return on equity, and its monthly skewness match their empirical counterparts quite well when solved with log-linearization, although the picture changes for the true solution. The equity premium decreases by 44% and the unconditional volatility of the equity premium by (a modest) 8%. Consistent with the intuition of Figure 1, the log-linearization leads to a higher volatility of returns than an accurate solution. The too-high volatility in bad states relative to the true volatility outweighs the too low volatility in good states. Moreover, the sign of monthly return skewness even switches from negative to positive, hence being even further away from the data. The negative skewness of returns in the data has been a puzzle as it cannot be rationalized within conditional normal LRR, habit, or disaster-risk models (Wachter and Zhu (2019)). We show that non-Gaussian LRR models also lack an explanation for this, as jumps in state variables do not create negatively skewed returns. As for the price–dividend ratio, the mean level significantly increases, by more than 20 percent, from 2.97 to 3.59. The volatility and the skewness of the price–dividend ratio react qualitatively similarly to the return—that is, the volatility drops modestly while the skewness changes substantially (in absolute terms). In summary, the parameterization of Drechsler and Yaron (2011) with multiple jumps in state variables leads to larger approximation errors than even the most extreme persistence or risk aversion levels considered in the literature on conditional Gaussian LRR models so far.[14]

[14]Note that if we shut down jump shocks, both solution methods result in the same basic asset-pricing moments. They still, however, differ significantly in the VRP as the log-linearization would imply a small but positive VRP while the true solution results in a slightly negative VRP.

Table 4: Summary Statistics of Asset-Pricing Moments

Moment	**Data**		**Model**					
			Log-linear			**Global**		
	Est.	SE.	5%	50%	95%	5%	50%	95%
$E[r^f]$	0.72	0.33	0.32	0.94	1.39	0.73	1.23	1.61
$\sigma[r^f]$	3.11	0.38	0.88	1.43	2.28	0.72	1.10	1.63
$AC[r^f]$	0.66	0.08	0.28	0.48	0.67	0.37	0.56	0.73
$E[r^d - r^f]$	5.28	2.13	3.64	6.24	8.89	0.84	3.47	6.07
$\sigma[r^d]$	19.46	1.70	14.72	17.41	20.78	13.68	16.01	18.78
$AC[r^d]$	-0.07	0.12	-0.24	-0.06	0.11	-0.21	-0.01	0.16
$skew[r^d]$	-0.83	0.21	-0.61	0.06	0.75	-0.57	0.03	0.66
$skew[r^d](m)$	-0.55	0.54	-0.97	-0.30	0.10	-0.26	0.03	0.33
$E[z^d]$	3.41	0.05	2.88	2.97	3.04	3.53	3.59	3.65
$\sigma[z^d]$	0.45	0.03	0.13	0.17	0.23	0.11	0.15	0.20
$AC[z^d]$	0.85	0.05	0.36	0.57	0.74	0.33	0.55	0.72
$skew[z^d]$	0.23	0.18	-1.33	-0.45	0.20	-1.09	-0.28	0.39

The table reports the summary statistics of the risk-free rate, return on equity, and the log price–dividend ratio for the parameter values denoted in Table 2. Monthly moments are indicated by (m); all other moments are annual. We follow Beeler and Campbell (2012) and Drechsler and Yaron (2011) to calculate and time-aggregate all moments. The first two columns display the estimated mean and standard error of US data for the sample period from 1930 to 2016, the next three columns the moments implied by the model solved with log-linearization, and the last three columns the moments implied by the global solution. The model-implied moments are the 5%, 50%, and 95% quantile values from 1,000 simulations with the same length as the data sample, where we use the same seed for each calibration to ensure comparability.

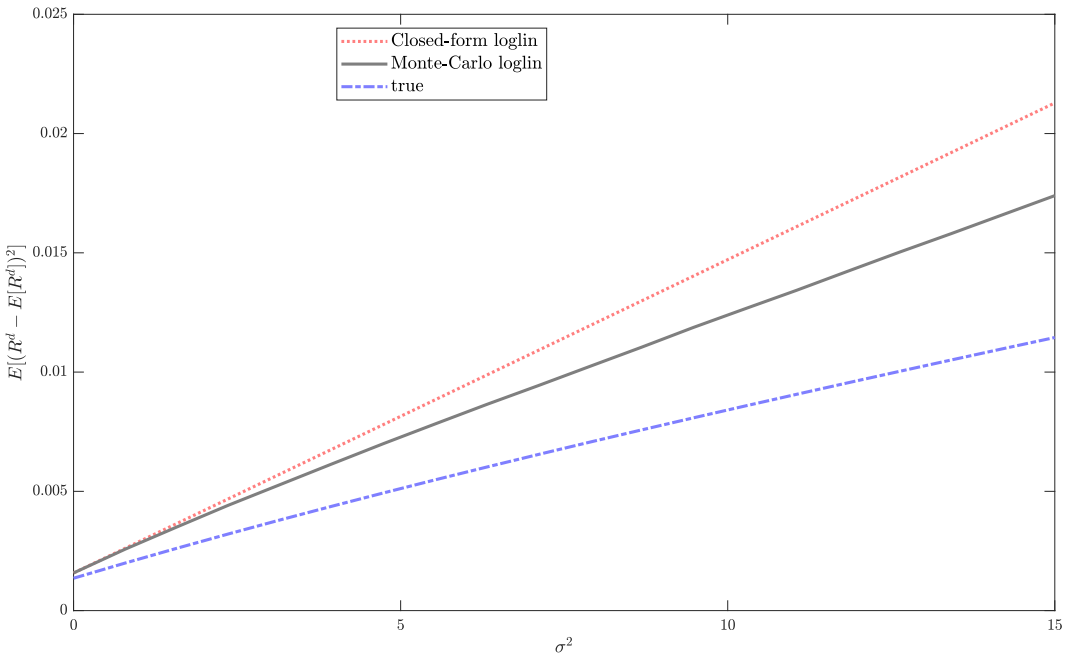
4.3 The Variance Risk Premium

Given that log-linearization suppresses the nonlinear behavior of volatility, the VRP and its constituents—that is, the physical and the risk-neutral expectation of the return variation—are especially prone to approximation errors. First, we analyze the approximation of the physical expectation. Figure 3 shows three different approximations of the conditional variance. The red dotted line plots the linear approximation $E_t[\,Var_t(r^d_{t+1})] = A_0^{Var_t} + A_x^{Var_t} x_t + A_\sigma^{Var_t}\sigma_t^2 + A_{\bar\sigma}^{Var_t}\bar\sigma^2$. Next, we stay in the log-linear framework but instead of using the log-linear approximation of conditional variance, we only use the log-linear function of the price–dividend ratio, run a Monte Carlo simulation conditionally on the level of σ_t^2, calculate the total variation, and take the expected value. This allows for a nonlinear volatility and is similar to dynamically adjusting the point of expansion as in Equation (17). As we can see from the gray solid line, the log-linear approximation and the simulated variance only coincide for small values of σ_t^2 and then diverge. In total, the volatility approximation performs poorly even if we assume the price–dividend ratio to be log-linear. Further accounting for nonlinearities in the price–dividend ratio leads to an even lower conditional volatility, which is drawn as a blue dash-dotted line in Figure 3. The latter corresponds to the conditional volatility that we approximate with Chebyshev polynomials after solving for the price–dividend ratio with global solution methods.

The summary statistics in Table 5 based on the simulation of the model confirm this observation. The *log-linear* model corresponds to the dotted line in Figure 3. The *nonlinear var* model corresponds to the black solid line—that is, we assume the log price–dividend ratio is linear but let the conditional variance be nonlinear by using a Monte Carlo integration at each point in time. The solution reported under the *global* model also corrects for nonlinearities in the price–dividend and wealth–consumption ratios. This step-by-step correction allows us to disentangle the effects stemming from a nonlinear underlying function from those stemming from a nonlinear conditional variance.

For the conditional variance under the physical measure, we observe that taking into account the nonlinearities in conditional variance reduces the mean only slightly. As the linear approximation is exact in the steady state and higher for good states, this small effect is to be expected. Taking into account the nonlinear price–dividend ratio using global solution methods gives an accurate mean value. The nonlinear conditional variance shows, however, exactly the behavior predicted by the intuition of Figures 1 and 3. Due to the constant slope, the conditional variance is much more volatile and exhibits more extreme values, as can be seen from the skewness and the kurtosis of the linear variance in comparison to the nonlinear

Figure 3: Variance Approximation



The figure shows the physical expectation of the return variation as a function of σ^2 keeping $\bar{\sigma}^2 = 1$ and $x_t = 0$. The red dashed line depicts the log-linear solution. The black solid line represents the conditional variance based on a Monte Carlo simulation with 1,000,000 draws using only the log-linear solution for the price–dividend ratio. The blue dash-dotted line shows the true expectation based on the global solution.

variance. Both higher moments, on the other hand, change only slightly when allowing for nonlinearities in the underlying price–dividend ratio. Hence, correctly approximating the price–dividend ratio is more important for the unconditional mean and variance, while the nonlinearity in the conditional variance already corrects for the higher moments.

Still, the VRP could exist if the change of measure puts enough weight on the tails of the distribution induced by jumps. But, as the nonlinear variance solution already has less extreme values, the change of measure does not imply the strong reweighting as the log-linearized solution does. Consequently, the risk-neutral conditional volatility is even more strongly influenced by the nonlinearities in the conditional variance. Hence, the drop in the mean and volatility between the log-linearized and the nonlinear conditional variance approximation is more significant and decreases further when we take the nonlinearities in the price–dividend ratio into account in the global solution. Similar to the physical conditional variance, the higher moments change only moderately if we move from the nonlinear conditional variance to the global solution. In total, the VRP already drops by more than 40 percent just by correctly approximating the conditional variance—even if the log price–dividend ratio would be a linear function. All in all, due to the tight link between the physical and the risk-neutral

Table 5: Summary Statistics of the Variance Risk Premium

Moment	Data		Model								
			Log-linear			Nonlinear var			Global		
	Est.	SE.	5%	50%	95%	5%	50%	95%	5%	50%	95%
$E[Var_t]$	31.72	3.30	24.69	29.65	36.01	23.71	27.84	32.92	19.38	22.17	25.53
$\sigma[Var_t]$	50.44	14.57	14.59	23.86	36.16	12.21	19.17	27.91	8.38	13.10	18.85
$skew[Var_t]$	7.43	0.90	3.01	4.00	5.79	2.73	3.60	5.09	2.72	3.56	4.99
$kurt[Var_t]$	79.06	22.43	13.49	22.77	46.83	11.39	18.79	36.56	11.33	18.35	35.28
$E[Var_t^{\mathbb{Q}}]$	41.07	2.86	29.03	36.39	45.82	26.48	31.79	38.20	19.87	22.82	26.32
$\sigma[Var_t^{\mathbb{Q}}]$	43.76	5.27	21.64	35.38	53.62	15.74	24.08	34.49	8.88	13.70	19.49
$skew[Var_t^{\mathbb{Q}}]$	3.09	0.33	3.01	4.00	5.79	2.58	3.39	4.79	2.63	3.41	4.79
$kurt[Var_t^{\mathbb{Q}}]$	14.60	2.95	13.49	22.77	46.83	10.48	17.05	32.82	10.61	17.13	32.52
$E[VRP_t]$	9.35	2.07	4.34	6.73	9.80	2.77	3.93	5.27	0.47	0.64	0.81
$\sigma[VRP_t]$	31.62	7.23	7.05	11.52	17.46	3.51	4.99	6.63	0.53	0.65	0.77

This table reports descriptives of the VRP, IV, and RV measured in monthly percentage terms for the parameter values denoted in Table 2. The first column displays the data moments of US data calculated following Drechsler and Yaron (2011) and Bollerslev, Tauchen, and Zhou (2009) for the period 1996 to 2016. The model-implied results are calculated following Drechsler and Yaron (2011) and use only the last 20 years of each simulated path with the 5% and 95% quantiles of the simulations in square brackets.

measure enforced by Epstein–Zin utility (Andersen, Fusari, and Todorov (2015)), the VRP only amounts to 0.64 if we solve the model globally—that is, it almost completely vanishes.

Notably, the risk-neutral conditional volatility shrinks by a larger fraction than the physical one. This observation indicates a second effect besides the approximation error in the conditional variance. Due to jumps in state variables, even the wealth–consumption ratio displays large nonlinearities, which the log-linearization cannot capture. Hence, the log-linearization overstates the return on the consumption claim, which feeds into the pricing kernel. This overstatement of the pricing kernel affects the Radon–Nikodym derivative and consequently implies a stronger change of measure than the global solution. To clearly show this point, we repeat this exercise for the stylized model of Section 3, in which the wealth–consumption and price–dividend ratios are almost log-linear. We report the results in Table 8 in Appendix C. In line with our intuition, using a nonlinear function for the conditional variance leads to almost the same VRP as the global solution because the errors in the stylized model are (mainly) due to the nonlinear dynamics in conditional volatility.

4.4 Sources of Errors

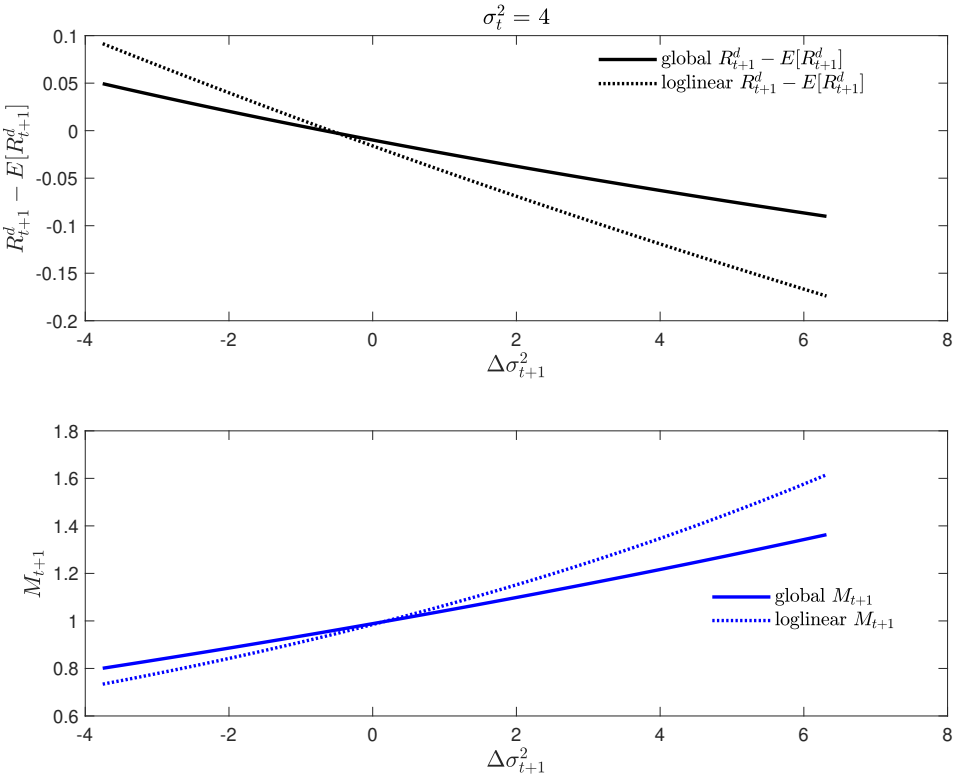
To further explore why the VRP is substantially smaller when the model is solved globally, we further decompose it and look at each part separately. First, note that we can rewrite the VRP similar to the equity premium as the covariance of the conditional variance with the

pricing kernel:

$$VRP_t = exp(r_t^f)Coskew_t[M_{t+1}, (R_{t+1}^d - E_t(R_{t+1}^d))]. \tag{18}$$

That is, the VRP will be high if absolute extreme returns are positively correlated with the pricing kernel. To examine how the log-linearization affects both quantities, we plot them

Figure 4: VRP Constituents



This plot illustrates the pricing kernel M_{t+1} and the unexpected return $R_{t+1}^d - E_t(R_{t+1}^d)$ for $\sigma_t^2 = 4$. The horizontal axis denotes the change in σ_{t+1}^2 ($\Delta\sigma_{t+1}^2 = \sigma_{t+1}^2 - \sigma_t^2$). We set all other state variables as well as both consumption and dividend growth to their unconditional means. Then we keep all innovations except $e_{\sigma,t+1}$ and J_{t+1}^{σ} at 0, impose the shocks on σ_t^2, and calculate the pricing kernel and the unexpected return for the respective $\Delta\sigma_{t+1}^2$.

as a function of changes in σ^2, in Figure 4. The upper panel shows the unexpected return $R_{t+1}^d - E_t(R_{t+1}^d)$ as a function of $\Delta\sigma_t^2 = \sigma_{t+1}^2 - \sigma_t^2$.[15] In order to isolate the effects of the volatility process, we set the shocks of all other endogenous and exogenous variables to zero. We choose the range of shocks in σ_t^2 by $\pm 1\phi_\sigma$ and values up to the 90% quantile of the jump size distribution. As the initial value, we choose $\sigma_t^2 = 4$, but the picture is (qualitatively) unchanged for smaller initial values; see Figure 9 in Appendix C. In the upper panel of Figure 4, we see that the log-linear solution overestimates returns in absolute terms for gains as well as for losses. More importantly, the lower panel shows that not only quantities based on

[15]Note that σ_t^2 possesses a mean-reverting component. Hence, a value of $\Delta\sigma_t^2 = 0$ resembles a positive shock by the size of the mean reversion $\sigma_{t+1}^2 - \bar{\sigma}_t^2 + \nu_\sigma(\sigma_t^2 - \bar{\sigma}_t^2) - E[\xi^\sigma]\lambda_{\sigma,t} = 4 - 1 + 0.87(4-1) - 2.55\frac{0.8}{12}4 = 1.07$.

the price–dividend ratio—which contains more nonlinearities than the wealth–consumption ratio due to the leverage—but also quantities based on the wealth–consumption ratio display severe approximation errors. The solid blue line denotes the level of the pricing kernel using the global solution while the dashed blue line denotes the pricing kernel stemming from the log-linearization. Recall from Equation (3) that both quantities only differ due to the return on the consumption claim. The return on the consumption claim is—similar to the return on the dividend claim in the upper panel—always overstated in absolute terms. As the return on the consumption claim directly enters the pricing kernel, the latter is also overstated by the log-linear solution. This effectively translates into an inflated change of measure from the physical to the risk-neutral distribution.

5 Do We Reject the Model or Only the Calibration?

As the model predictions deviate significantly from the data even in terms of the equity premium, we now test whether the VRP is only non-existent due to the lack of sufficient risk. To do so, we recalibrate the model first by changing one moment at a time. Then, we show how the dynamics of the conditional volatility and risk premia behave inside the model.

5.1 Alternative Calibrations

First, we identify three major parameters that might have an impact on the VRP. The first is the risk aversion γ. Increasing risk aversion has the advantage that it does not distort the state vector dynamics. Hence, we increase it until the model matches the equity premium and the risk-free rate. The fourth column of Table 6 documents that even for $\gamma = 13$ the model-implied VRP is not comparable to that in the data. For this reason, we next examine the effect of increasing the persistence of the expected growth rate ρ to 0.981. A higher persistence leads to an extended half-life of jumps and should have an impact on the VRP. This is especially true for the persistence of the long-run growth rate, which is the main channel through which the model works. As can be seen from the fifth column, the VRP does increase slightly due to the persistence of the expected growth rate ρ, but not enough to consider the puzzle solved.[16] Third, we increase the parameter ϕ_{ξ^x}, which drives the variability of jumps in the expected growth rate. As jumps in x carry the highest price of risk in absolute terms, we vary the jump size to maximize the effect on the VRP without overly affecting the equity premium.[17] Still,

[16] Note that a value of 0.981 for ρ also leads to counterfactual implications w.r.t. other model moments, especially the predictability by the price–dividend ratio (Beeler and Campbell (2012)).

[17] We thank Nicole Branger for raising this point.

the model can explain less than 30 percent of the size of the VRP.[18] We also observe that neither calibration can explain the negative skewness of monthly returns.

Table 6: Alternative Calibrations

Moment	**Data**	**DY**	γ =13	ρ =0.981	ϕ_{ξ^x} =1.2e-3
$E[r^f]$	0.72	1.23	0.81	0.94	0.89
$E[r^d - r^f]$	5.28	3.47	5.76	5.53	5.84
$skew[r^d](m)$	-0.55	0.03	-0.03	-0.02	-0.01
$E[z^d]$	3.41	3.59	3.08	3.11	3.06
$E[VRP]$	9.35	0.64	1.85	2.31	2.67

The table displays selected empirical moments, the respective values of the model by Drechsler and Yaron (2011), and three alternative calibrations of the model. For each alternative calibration, we change one parameter, keeping all others to their respective values denoted in Table 2. The column names display the respective parameter value for each calibration. All results are obtained with global solution methods.

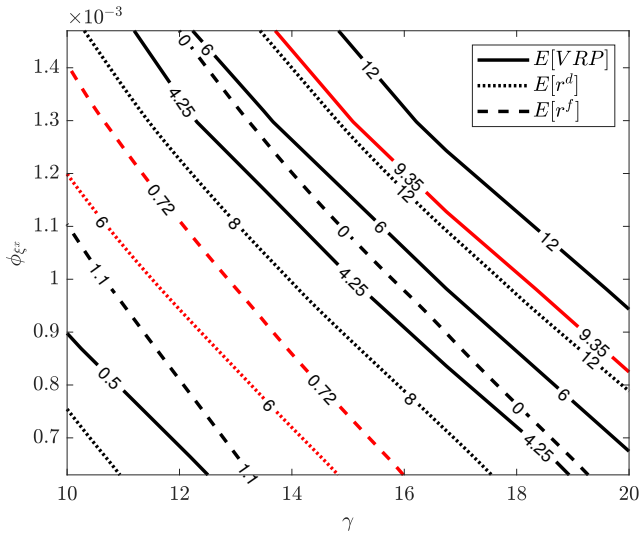
As altering one parameter, ceteris paribus, does not help to create a sizeable VRP, we now conduct a sensitivity analysis with respect to two parameters—the risk aversion and the variability of jumps in the expected growth rate x. To do so, we solve the model for a wide range of each parameter. We plot the resulting values of the VRP, the return on the dividend claim, and the risk-free rate in Figure 5 as isolines. That is, the solid line with the label *10* marks combinations of γ and ϕ_{ξ^x} that result in a VRP of 10. We further color the isolines corresponding to the empirical counterpart of each target moment red. The three red isolines would need to intersect for there to be a calibration that could jointly generate a low—but not too low—risk-free rate, a high return on equity, and a high VRP. We observe, however, that the isolines for each moment shift in an almost parallel manner. Hence, there exists no possible combination of those two parameters to match the three moments at the same time.

We do not report changes of parameters affecting the volatility process σ_t^2. Due to the relatively low level of persistence of σ_t^2, a slight increase of the persistence does not result in significantly better results. Increasing the persistence of volatility does, however, increase the domains over which we have to solve.[19] Furthermore, using the existence condition of Pohl, Schmedders, and Wilms (2019) we find that higher values of the persistence of σ_t^2 or altering the cumulant-generating function of the jumps in σ_t^2 make the existence condition for the model of Drechsler and Yaron (2011) significantly more demanding. As Drechsler and Yaron

[18]Note that our data estimate, which we calculate following Drechsler and Yaron (2011), is a conservative estimate of the VRP. Using the time series provided by Hao Zhou, which dates back to 1990, the VRP would be even higher, amounting to 15.8.

[19]In Figure 8 we plot the ergodic sets for the state variables given the calibration of Drechsler and Yaron (2011). Note that as jumps are more likely the higher the volatility and the jump size in volatility has a large tail, the domain is already quite large with regions of very low probability. Hence, increasing the jump size or the persistence of volatility leads to even more extreme values, which the global solution has to cover.

Figure 5: Sensitivity Analysis



This plot shows the return on equity, the risk-free rate, and the VRP as a function of ϕ_{ξ^x} (y-axis) and γ (x-axis), keeping all remaining parameters unchanged. The solid lines are isolines for the VRP, the dotted lines are isolines for the return on the dividend claim, and the dashed lines are isolines for the risk-free rate. The isolines are calculated by solving and simulating the model for five different values of each parameter—that is, we solve and simulate the model 125 times in total. The red isolines denote the empirical point estimates that are targeted by the calibration.

(2011) show, however, the main mechanism of generating jumps in the model works through jumps in the long-run growth rate, as shutting down this channel results in a VRP of almost zero, even in the log-linearized model (see Table 14 in Drechsler and Yaron (2011)).

In sum, the results in Section 4.3 document that the original parametrization of the Drechsler and Yaron (2011) model—see Table 2—does not deliver a sizable VRP. Moreover, no simple variation of the original parametrization allows the model to match the VRP. While finding parameter values for the model to match the VRP certainly appears possible, the resulting parametrized model would deliver other key moments, such as the risk-free rate or the equity premium, which would deviate significantly from observed values. Since we did not, however, perform a complete sweep through the parameter universe—due to the large number of parameters and the strong interdependencies between the state variables—we cannot claim to reject the model beyond any doubt. Nevertheless, we conjecture that jumps in fundamentals do not yield a variance risk premium that can match the data. The narrative that the VRP is purely a compensation for time-varying macroeconomic risks seems highly unlikely.

Instead, we conjecture that alternative explanations are needed for the VRP. Recent empirical evidence from Bekaert and Hoerova (2016) and Londono and Xu (2019) suggests that the VRP appears to be driven much more by time-varying risk aversion than by fundamen-

tals. Consequently, models featuring a preference-based explanation might be a promising alternative with which to explain the VRP.[20] Moreover, recent literature has also found that frictions might play a role in financial markets, and especially in the option market. Hence, theories such as demand-based option pricing or constrained financial intermediaries might also help to explain the size of the VRP.[21]

5.2 Procyclicality

The most puzzling feature of the VRP stems from the procyclical effect on conditional volatility in long-run risks models discussed earlier. In this section, we examine if this effect also translates into risk premia in the full model of Drechsler and Yaron (2011). The volatility σ_t^2 directly enters the conditional volatility of cash-flow as well as of discount-rate news. Hence, the intuition for its influence on the dynamics of premia inside the model is largely in line with log-linearization as we have already shown in Figure 3—that is, the VRP increases in σ_t^2, although not linearly. On the other hand, shocks to the long-run mean of volatility and the expected growth rate only affect the level of the wealth–consumption and price–dividend ratios, while the conditional volatility of consumption and dividends stays unchanged. Hence, conditional volatility is higher for better states of both variables. To examine the effect on the VRP, we plot selected quantities in the recalibrated model with higher persistence of expected growth, although the effect is the same for all calibrations.

To further clarify the term procyclicality, we keep both volatilities constant and examine the effect of the expected consumption growth.[22] Empirically, we would expect that if risk premia vary with expected consumption growth at all, they should be high in times of low expected consumption growth—that is, in recessions—and low in times of high expected consumption growth (Fama and French (1989)). This not only holds for the equity risk premium but also for the VRP, which is especially high during times of economic turmoil and financial crises (see for example Zhou (2018)). From the perspective of valuation ratios—that is, the wealth–consumption and the price–dividend ratio—this would imply that risk premia are negatively correlated with them. Furthermore, Lettau and Ludvigson (2010) show that, empirically, the conditional Sharpe Ratio is also strongly countercyclical and that the

[20] See e.g. Lorenz and Schumacher (2018), Baele, Driessen, Ebert, Londono, and Spalt (2018), Bekaert and Engstrom (2017), or Schreindorfer (forthcoming).

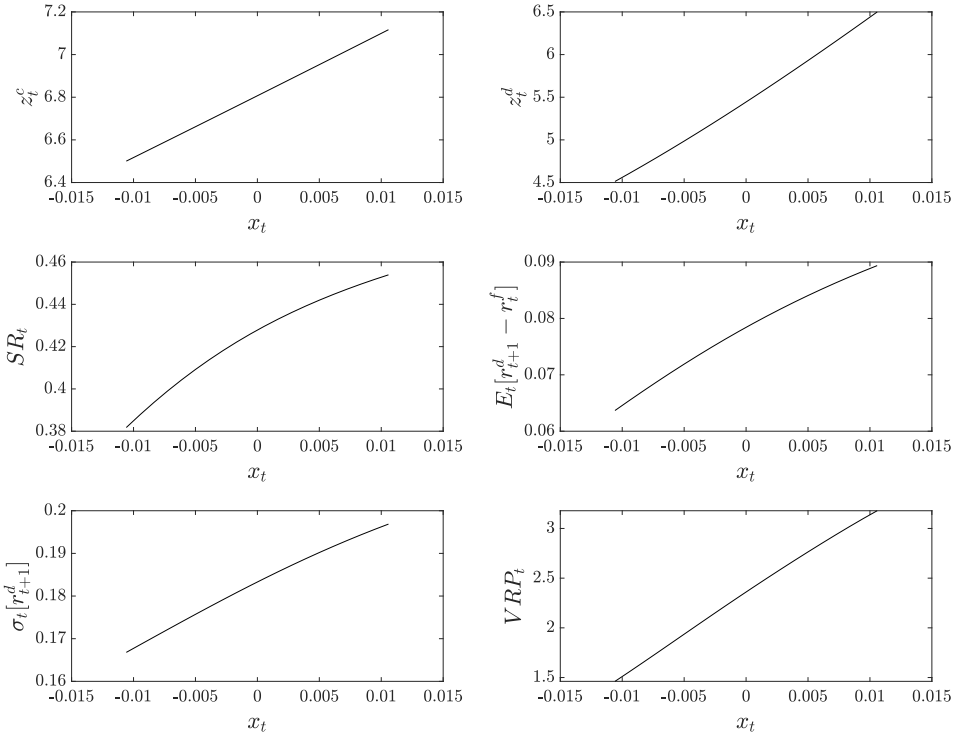
[21] See e.g. Garleanu, Pedersen, and Poteshman (2008), Drechsler, Moreira, and Savov (2019), or He, Kelly, and Manela (2017).

[22] We also show the effect of the long-run mean of volatility, in Figure 10. Since the long-run mean of volatility does not increase the conditional volatility directly, the instant effects are the same as for the expected consumption growth. For ease of exposition, however, we only look at the expected consumption growth, which is the most important state variable in the model, in this section.

log-linearized model by Bansal and Yaron (2004) is not able to explain the dynamics of the Sharpe Ratio.

Figure 6 shows the model-implied valuation ratios, the Sharpe Ratio, conditional volatility, and risk premia across different values of the expected consumption growth. The valuation ratios display the expected behavior—that is, they are higher with higher expected consumption growth. Volatility and risk premia, on the other hand, increase with expected consumption growth. This is the case for the equity risk premium as well as for the variance risk premium.[23] As both the conditional volatility and expected returns increase with expected consumption growth, both could cancel each other out in the Sharpe Ratio. This is, however, not the case, as the left plot in the middle row shows. Contrary to the empirical evidence, the Sharpe Ratio is procyclical too.

Figure 6: Valuation Ratios, Conditional Volatility, and Expected Risk Premia w.r.t. x_t



This plot shows the model-implied functions w.r.t. x_t for the calibration $\rho = 0.981$, keeping $\sigma_t^2 = 1$ and $\bar{\sigma}_t^2 = 1$. In the upper row, the left plot denotes the wealth–consumption ratio and the right plot the price–dividend ratio. In the middle row, the left plot shows the conditional Sharpe Ratio $SR_t = \frac{E_t[r_{t+1}^d - r_t^f]}{\sigma_t[r_{t+1}^d]}$ and the right plot the expected equity premium. In the bottom row, the left plot shows the conditional standard deviation and the right plot the VRP. All functions are computed using global solution methods.

[23] Pohl, Schmedders, and Wilms (2018) find that the equity premium is procyclical in the model of Bansal and Yaron (2004) (see their earlier working-paper version) but cannot explain the origin of this seemingly puzzling behavior. We deliver the exact mechanism inside the model, which stems entirely from the conditional volatility approximation and not from nonlinear effects in the valuation ratios.

All in all, the procyclicality of conditional variance induced by the time-varying weighting carries over to the equity premium, the conditional Sharpe Ratio, and the VRP. The log-linearization, on the other hand, would imply that those three quantities do not change if σ_t^2 does not change.

6 Conclusion

In this paper, we have presented an analysis of the dynamics of conditional volatility in long-run risk models. The common view of LRR models is that they only generate time varying risk premia through stochastic volatility. We show that by the construction of the model, risk premia are time varying even in the absence of stochastic volatility—but in a counterfactual way. This effect has been overlooked in the literature so far due to log-linearized model solutions. The time varying weighting of cash-flow and discount-rate variance leads to different insights from the models if solved correctly. We show that even if the underlying function is well approximated by the log-linear solution as in our stylized model, log-linearization fails for higher-order moments—especially for the dynamics of conditional volatility and the size of the VRP.

In particular, the procyclicality of conditional volatility and risk premia with respect to the expected growth rate is hard to reconcile with the empirical facts. This mechanism lies at the heart of LRR models—the high importance of the small but persistent expected growth rate x_t. Even in models with more complex volatility dynamics, as in Drechsler and Yaron (2011), the main mechanism works through the expected growth rate. A potential candidate model linking fluctuating uncertainty to future asset returns should thus not only include time-varying risk—as in Drechsler and Yaron (2011)—but also time-varying *prices* of risk. This mechanism can either be preferences or constraint based.

We further show that approximation errors increase dramatically for models with jumps in fundamentals. In the calibrated model of Drechsler and Yaron (2011), log-linearization overstates the equity premium by 79.8 percent and the VRP by more than 950 percent.

Finally, the procyclical behavior potentially also has an impact on model-implied option prices and distorts the shape of the implied volatility smile. The model of Drechsler and Yaron (2011) was, however, not calibrated to match this pattern. LRR models that try to capture the behavior of options are mostly solved in continuous time; see, for example, Drechsler (2013) or Branger, Rodrigues, and Schlag (2018). Using log-linearization in continuous time has recently been shown to lead to severe errors in risk premia (Tsai and Wachter (2018)). The

consequences for the conditional volatility approximation in continuous times are, however, still unknown. Given the integral role of continuous-time models in explaining option prices and the VRP, exploring potential errors is an important task for future research.

Appendix

A Derivations

A.1 Conditional Variance in an LRR model

Take, for example, the model of Bansal and Yaron (2004), where

$$\begin{aligned} \Delta c_{t+1} &= \mu_c + x_t + \phi_c \sigma_t e_{c,t+1} \\ x_{t+1} &= \rho x_t + \phi_x \sigma_t e_{x,t+1} \\ \sigma_t^2 &= \bar{\sigma}^2 + \nu(\sigma_t^2 - \bar{\sigma}^2) + \phi_\sigma e_{\sigma,t+1} \\ \Delta d_{t+1} &= \mu_d + \Phi x_t + \phi_d \sigma_t e_{d,t+1}. \end{aligned} \tag{19}$$

Given their calibration, the model can be solved by assuming a linear function for the wealth–consumption and price–dividend ratios (Bansal and Yaron (2004), Pohl, Schmedders, and Wilms (2018)):

$$\begin{aligned} z_t^c &= A_0 + A_x x_t + A_\sigma \sigma_t^2 \\ z_t^d &= A_{0,d} + A_{x,d} x_t + A_{\sigma,d} \sigma_t^2. \end{aligned} \tag{20}$$

Taking the conditional volatility of dividends and the price–dividend ratio results in

$$\begin{aligned} Var_t[\Delta d_{t+1}] &= \phi_d^2 \sigma_t^2 \\ Var_t[z_{t+1}^d] &= (A_{x,d})^2 \phi_x^2 \sigma_t^2 + (A_{\sigma,d})^2 \phi_\sigma^2. \end{aligned} \tag{21}$$

The conditional volatility of the return on the dividend claim using the Campbell–Shiller approximation equals

$$Var_t[r_{t+1}^d] = \left(\left(\kappa_1^d\right)^2 \left(A_{x,d}\right)^2 \phi_x^2 + \phi_d^2 \right) \sigma_t^2 + \left(\kappa_1^d\right)^2 \left(A_{\sigma,d}\right)^2 \phi_\sigma^2, \tag{22}$$

where $\kappa_1^d = \frac{\exp\left(\bar{z}^d\right)}{1+\exp\left(\bar{z}^d\right)}$. Assuming a preference for the early resolution of uncertainty, $A_{x,d} < 0$ and thus positive shocks to long-run growth increase the price–dividend ratio. If κ_1^d is constant—as assumed in the solution procedure—a positive shock to long-run growth does not enter the conditional variance. This, however, is only an approximation, due to the dependence of $\kappa_1^d(z_t^d)$ on the point of expansion. To hold exactly, we have to vary $\kappa_1^d(z_t^d)$ in accordance with the current level of the price–dividend ratio—that is, we plug in $z_t^d = A_{0,d} + A_{x,d} x_t + A_{\sigma,d} \sigma_t^2$. Plugging in the equation for $k_1^{z_t^d}$, the sensitivities of the variance w.r.t. the respective variables

are

$$\begin{aligned}
\frac{\mathrm{d}z_t^d}{\mathrm{d}x_t} &> 0 \\
\frac{\mathrm{d}\kappa_1^d}{\mathrm{d}x_t} &> 0 \\
\frac{\mathrm{d}\,Var_t[r_{t+1}^d]}{\mathrm{d}x_t} &> 0 \\
\frac{\mathrm{d}\,Var_t[r_{t+1}^d]}{\mathrm{d}\sigma_t^2} &> 0 \\
\frac{\mathrm{d}^2\,Var_t[r_{t+1}^d]}{\mathrm{d}(\sigma_t^2)^2} &< 0.
\end{aligned} \tag{23}$$

The exact formula for $\frac{\mathrm{d}\,Var_t[r_{t+1}^d]}{\mathrm{d}\sigma_t^2}$ is

$$\begin{aligned}
\frac{\mathrm{d}\,Var_t[r_{t+1}^d]}{\mathrm{d}\sigma_t^2} = \phi_d^2 + \sigma_t^2 \Bigg(& \frac{(2A_2 B_x^2 exp(2A_{0,d} + 2A_{\sigma,d}\sigma_t^2 + 2A_{x,d}x_t))}{(\exp(A_{0,d} + A_{\sigma,d}\sigma_t^2 + A_{x,d}x_t) + 1)^2} \\
& - \frac{(2A_{\sigma,d}B_x^2 \exp(A_{0,d} + A_{\sigma,d}\sigma_t^2 + A_{x,d}x_t)\exp(2A_{0,d} + 2A_{\sigma,d}\sigma_t^2 + 2A_{x,d}x_t))}{(\exp(A_{0,d} + A_{\sigma,d}\sigma_t^2 + A_{x,d}x_t) + 1)^3} \Bigg) \\
& + \frac{(B_x^2 \exp(2A_{0,d} + 2A_{\sigma,d}\sigma_t^2 + 2A_{x,d}x_t))}{(\exp(A_{0,d} + A_{\sigma,d}\sigma_t^2 + A_{x,d}x_t) + 1)^2} + \frac{(2A_{\sigma,d}B_\sigma^2 \exp(2A_{0,d} + 2A_{\sigma,d}\sigma_t^2 + 2A_{x,d}x_t))}{(\exp(A_{0,d} + A_{\sigma,d}\sigma_t^2 + A_{x,d}x_t) + 1)^2} \\
& - \frac{(2A_{\sigma,d}B_\sigma^2 \exp(A_{0,d} + A_{\sigma,d}\sigma_t^2 + A_{x,d}x_t)\exp(2A_{0,d} + 2A_{\sigma,d}\sigma_t^2 + 2A_{x,d}x_t))}{(\exp(A_{0,d} + A_{\sigma,d}\sigma_t^2 + A_{x,d}x_t) + 1)^3}.
\end{aligned} \tag{24}$$

The exact formula for $\frac{\mathrm{d}^2\, Var_t[r^d_{t+1}]}{\mathrm{d}(\sigma_t^2)^2}$ is

$$\begin{aligned}
\frac{\mathrm{d}^2\, Var_t[r^d_{t+1}]}{\mathrm{d}(\sigma_t^2)^2} = \sigma_t^2 \Bigg(& \frac{4(A_{\sigma,d})^2 B_x^2 \exp(2A_{0,d} + 2A_{\sigma,d}\sigma_t^2 + 2A_{x,d}x_t)}{\left(\exp(A_{0,d} + A_{\sigma,d}\sigma_t^2 + A_{x,d}x_t) + 1\right)^2} \\
& + \frac{6(A_{\sigma,d})^2 B_x^2 \exp(4A_{0,d} + 4A_{\sigma,d}\sigma_t^2 + 4A_{x,d}x_t)}{\left(\exp(A_{0,d} + A_{\sigma,d}\sigma_t^2 + A_{x,d}x_t) + 1\right)^4} \\
& - \frac{10(A_{\sigma,d})^2 B_x^2 \exp(A_{0,d} + A_{\sigma,d}\sigma_t^2 + A_{x,d}x_t) \exp(2A_{0,d} + 2A_{\sigma,d}\sigma_t^2 + 2A_{x,d}x_t)}{\left(\exp(A_{0,d} + A_{\sigma,d}\sigma_t^2 + A_{x,d}x_t) + 1\right)^3} \Bigg) \\
& + \frac{4(A_{\sigma,d})^2 B_\sigma^2 \exp(2A_{0,d} + 2A_{\sigma,d}\sigma_t^2 + 2A_{x,d}x_t)}{\left(\exp(A_{0,d} + A_{\sigma,d}\sigma_t^2 + A_{x,d}x_t) + 1\right)^2} \\
& + \frac{6(A_{\sigma,d})^2 B_\sigma^2 \exp(4A_{0,d} + 4A_{\sigma,d}\sigma_t^2 + 4A_{x,d}x_t)}{\left(\exp(A_{0,d} + A_{\sigma,d}\sigma_t^2 + A_{x,d}x_t) + 1\right)^4} \\
& + \frac{4A_{\sigma,d} B_x^2 \exp(2A_{0,d} + 2A_{\sigma,d}\sigma_t^2 + 2A_{x,d}x_t)}{\left(\exp(A_{0,d} + A_{\sigma,d}\sigma_t^2 + A_{x,d}x_t) + 1\right)^2} \\
& - \frac{10(A_{\sigma,d})^2 B_\sigma^2 \exp(A_{0,d} + A_{\sigma,d}\sigma_t^2 + A_{x,d}x_t) \exp(2A_{0,d} + 2A_{\sigma,d}\sigma_t^2 + 2A_{x,d}x_t)}{\left(\exp(A_{0,d} + A_{\sigma,d}\sigma_t^2 + A_{x,d}x_t) + 1\right)^3} \\
& - \frac{4A_{\sigma,d} B_x^2 \exp(A_{0,d} + A_{\sigma,d}\sigma_t^2 + A_{x,d}x_t) \exp(2A_{0,d} + 2A_{\sigma,d}\sigma_t^2 + 2A_{x,d}x_t)}{\left(\exp(A_{0,d} + A_{\sigma,d}\sigma_t^2 + A_{x,d}x_t) + 1\right)^3}.
\end{aligned} \tag{25}$$

A.2 Closed-Form Solutions for the Stylized Model

While the calibration by Drechsler and Yaron (2011) necessitates solving a system of equations for the coefficients of the log wealth–consumption ratio and the log price–dividend ratio, the coefficients are given in closed form for the stylized model. Following Drechsler and Yaron (2011), we denote the moment-generating function of the jump size in x_t as $m_\xi(u)$ and its second derivative by $m_\xi^{(2)}(u)$. For ease of notation, we define $\sigma_{x,t} \equiv \phi_x \sigma_t$ and drop the time subscript for $\sigma_{x,t}, \sigma_{c,t}, \sigma_{d,t}$, and λ_t as $\sigma_t = 0.0078 \ \forall \ t$. Following Bansal, Kiku, and Yaron (2012), we solve for the log-linearization constants endogenously. To calculate the wealth–consumption ratio, we plug Equations (13) and (3) into (2) and sort for coefficients, which

gives

$$
\begin{aligned}
m_\xi(u) &= \exp\left(\frac{\phi_{\xi_x}^2 u^2}{2}\right) \\
k_1^c &= \exp(z_t^c)(1+\exp(z_t^c))^{-1} \\
k_0^c &= \log(1+\exp(z_t^c)) - k_1^c z_t^c \\
A_x &= \frac{(1-\frac{1}{\psi})}{(1-k_1^c\rho)} \\
A_0 &= \frac{\left(\theta\log(\delta) + (\theta-\frac{\theta}{\psi})\mu + \theta k_0^c + 0.5(\theta-\frac{\theta}{\psi})^2\sigma_c^2 + 0.5(\theta k_1^c A_x\sigma_x)^2 + \lambda\left(m_\xi(\theta k_1^c A_x)-1\right)\right)}{(\theta-\theta k_1^c)}.
\end{aligned}
\tag{26}
$$

By a similar argument, the coefficients for the price–dividend ratio are given by

$$
\begin{aligned}
k_{1,d} =& \exp\left(z_t^d\right)\left(1+\exp\left(z_t^d\right)\right)^{-1} \\
k_0^d =& \log\left(1+\exp\left(z_t^d\right)\right) - k_1^d z_t^d \\
A_{x,d} =& \frac{(\Phi-\frac{1}{\psi})}{(1-k_1^d\rho)} \\
A_{0,d} =& \left(\theta\log(\delta) + \left(\theta-1-\frac{\theta}{\psi}\right)\mu + (\theta-1)\left(k_0^c + (k_1^c-1)A_0\right) + k_0^d + \mu_d + 0.5\left(\theta-1-\frac{\theta}{\psi}\right)^2\sigma_c^2\right. \\
&+ 0.5\left((\theta-1)k_1^c A_x + k_1^d A_{x,d}\right)^2\sigma_x^2 + 0.5\sigma_d^2 + \lambda\left(m_\xi\left((\theta-1)k_1^c A_x + k_1^d A_{x,d}\right)-1\right)\Bigg) \\
&\frac{1}{(1-k_1^d)}.
\end{aligned}
\tag{27}
$$

We use the fact that the monthly risk-free rate is given by $-\log(B_{0,t})$, where $B_{0,t} = E_t[M_{t+1}]$,

$$
\begin{aligned}
A_{0,B} =& \theta\log(\delta) + \left(\theta-1-\frac{\theta}{\psi}\right)\mu + (\theta-1)\left(k_0^c + k_1^c A_0 - A_0\right) + 0.5\left(\theta-1-\frac{\theta}{\psi}\right)^2\sigma_c^2 \\
&+ 0.5\left((\theta-1)k_1^c A_x\sigma_x\right)^2 + \lambda\left(m_\xi\left((\theta-1)k_1^c A_x\right)-1\right) \\
A_{x,B} =& \left(1 - A_x + k_1^c A_x\rho\right)(\theta-1) - \frac{\theta}{\psi}.
\end{aligned}
\tag{28}
$$

The conditional variance under the physical measure is given by using Equation (17). Sorting for coefficients gives

$$
\begin{aligned}
A_{0,\mathbb{P}} &= \left(k_1^d A_{x,d}\right)^2\sigma_x^2 + \left(k_1^d A_{x,d}\right)^2\lambda\phi_{\xi_x}^2 + \sigma_d^2 \\
A_{x,\mathbb{P}} &= 0.
\end{aligned}
\tag{29}
$$

To compute the conditional variance under the risk-neutral measure, we follow Drechsler and Yaron (2011) and apply the change of measure separately to the jump and the Gaussian part.

This leaves us with

$$\begin{aligned} A_{0,\mathbb{Q}} &= \left(k_1^d A_{x,d}\right)^2 \sigma_x^2 + \left(k_1^d A_{x,d}\right)^2 \lambda m_\xi^{(2)} \left((\theta - 1) k_1^c A_x\right) + \sigma_d^2 \\ A_{x,\mathbb{Q}} &= 0. \end{aligned} \tag{30}$$

Comparing the coefficients for the risk-neutral and the physical conditional variance directly shows that the contributions of the Gaussian risk cancel out in the VRP. Furthermore, both conditional volatilities are not state-dependent, which leads to a constant VRP as well, which is solely driven by the jump part.

A.3 VRP in the Data vs. Discrete-Time Models

The empirical VRP is closely related to a one-month variance swap that pays the difference between the total return variation over the next month and the variance swap rate. Demeterfi, Derman, Kamal, and Zou (1999), Britten-Jones and Neuberger (2000), Bakshi, Kapadia, and Madan (2003), Carr and Wu (2009), and Jiang and Tian (2005) show how the variance swap rate can be synthesized by an option portfolio in a model-free way. The ex post total return variation, on the other hand, is calculated by summing squared daily returns (see the CBOE website for a detailed description).

The ex ante conditional expected payoff being short in the variance swap is the premium investors pay to hedge against fluctuations in the variance—the variance risk premium:

$$VRP_t = E_t^{\mathbb{Q}}[\text{total return variation}_{t,t+1}] - E_t^{\mathbb{P}}[\text{total return variation}_{t,t+1}]. \tag{31}$$

As option prices already contain forward-looking information, the synthetic variance swap rate directly measures the risk-neutral expectation. Measuring the expected total return variation under the physical measure requires some explicit assumptions. First, the use of high-frequency data allows a more precise measurement of *ex post* variance (see Andersen, Bollerslev, Diebold, and Labys (2003) and Andersen, Bollerslev, Diebold, and Labys (2003)). Bollerslev, Tauchen, and Zhou (2009) further assume that variance follows a martingale and use the ex post realized variance between $t-1$ and t as an estimator for the conditional expectation at time t under the physical measure. Corsi (2009) and Drechsler and Yaron (2011), meanwhile, show that using additional predictors—for example, lagged realized variance or the risk-neutral expectation—significantly improves forecast accuracy compared to the martingale approach and therefore is a better approximation of the physical expectation.

How does this relate to discrete-time models? To be completely in line with the empirical

approach, one would need to solve a model at least at a daily frequency, simulate daily returns, and proceed with the same approach that one chose for the empirical physical expectation. Branger, Rodrigues, and Schlag (2018) show how this can be done in a continuous-time model. Most discrete-time models are, however, solved with a monthly decision horizon and consequently the approach based on realized variance is infeasible in this case. While Bonomo, Garcia, Meddahi, and Tédongap (2015) choose to change the decision horizon of the agent from monthly to daily, we want to be as close to the LRR literature—and especially Drechsler and Yaron (2011)—as possible. We therefore follow a different approach, which avoids the need to calculate the realized volatility and consequently allows us to keep the monthly decision horizon as in Drechsler and Yaron (2011). To do so, one can follow the arguments of Christoffersen, Fournier, Jacobs, and Karoui (2017), who define the price of co-skewness risk λ_t^{Coskew} as the uncentered second moment:

$$\lambda_t^{Coskew} = E_t^{\mathbb{P}}[(R_{t+1}^d)^2] - E_t^{\mathbb{Q}}[(R_{t+1}^d)^2]. \tag{32}$$

They further show that the price of co-skewness corresponds to the VRP empirically and that one can also use the second centered moment—as we do. We further switch the sign of λ_t^{Coskew} to be in line with Bollerslev, Tauchen, and Zhou (2009) and Drechsler and Yaron (2011). Consequently, we can express the VRP as the difference between the risk-neutral and physical conditional variance as in Equation (6), which can be calculated inside the model directly without resorting to another frequency. This can further be expressed as

$$\begin{aligned} VRP_t &= E_t\left[\frac{M_{t+1}}{E_t[M_{t+1}]}(R_{t+1}^d - E_t(R_{t+1}^d))^2\right] - E_t\left[(R_{t+1}^d - E_t(R_{t+1}^d))^2\right] \\ &= exp(r_t^f)\Big(E_t[M_{t+1}]E_t[(R_{t+1}^d - E_t(R_{t+1}^d))^2] + \\ &\quad Cov[M_{t+1}, (R_{t+1}^d - E_t(R_{t+1}^d))^2]\Big) - E_t[(R_{t+1}^d - E_t(R_{t+1}^d))^2] \\ &= exp(r_t^f)\, Cov_t[M_{t+1}, (R_{t+1}^d - E_t(R_{t+1}^d))^2] \\ &= exp(r_t^f)\, Coskew_t[M_{t+1}, (R_{t+1}^d - E_t(R_{t+1}^d))]. \end{aligned} \tag{33}$$

Drechsler and Yaron (2011) coin the expression in Equation (6) the *level difference* although they also argue that the VRP in a discrete-time model equals exactly this difference. They define "their" VRP, however, as the sum of the level difference and the drift difference, which

is given by

$$\begin{aligned} \text{drift difference} = \quad & \left\{ E_t^{\mathbb{Q}}[Var_{t+1}^{\mathbb{Q}}(r_{t+2}^d)] - Var_t^{\mathbb{Q}}(r_{t+1}^d) \right\} \\ & - \left\{ E_t^{\mathbb{P}}[Var_{t+1}^{\mathbb{P}}(r_{t+2}^d)] - Var_t^{\mathbb{P}}(r_{t+1}^d) \right\}. \end{aligned} \tag{34}$$

To be consistent with the previous literature and to isolate the effects of the nonlinear dynamics in conditional volatility as well as possible, we use the definition in Equation (6) and neglect the drift difference. Drechsler and Yaron (2011), however, show that the level difference strongly dominates their total VRP, which we also confirm in unreported results based on the global solution.

B Solution and Accuracy

B.1 Solution Details

We solve the model using projection methods (Judd (1992)). To keep the problem tractable, we use complete polynomials and a monomial integration rule. Judd, Maliar, and Maliar (2011) show that a quadrature integration with 2J nodes and weights delivers highly accurate solutions for models with rare disasters. As the jumps in the considered models are more frequent but smaller, the solutions have to be approximated on a much smaller domain and monomials will deliver a highly accurate solution. We also check the accuracy of our solution by comparing the results with a product rule, using the PALMA II high performance computing cluster of the University of Münster. Using five nodes for each dimension does not change the results.

For the Poisson process, we make use of its property that the probability of a jump within an ensuing short time interval between t and $t + h$ equals $\lambda h + o(h)$, while the probability of more than one jump is o(h); see, for example, Feller (1968) for a formal treatment of the Poisson process. For details on how to efficiently discretize a jump–diffusion process, see Piazzesi (2001) or Andersen, Benzoni, and Lund (2002). We also check the regularity of the Poisson processes. To this end, we make use of the superposition operation on point processes. We generate the original process X from the superposition of the independent point processes $X_1, \ldots, X_K$:

$$X = \bigcup_{k=1}^{K} X_i. \tag{35}$$

In robustness checks without a stochastic long-run mean of volatility, we split the original Poisson process into up to four independent point processes, which leaves us with a negligible

probability for more than one jump in each process. The results remain unchanged compared to those for one process.

B.2 Accuracy

Table 7 shows the mean, median, minimal, and maximal Euler Equation errors for the log-linear as well as the global solutions of the wealth–consumption and price–dividend ratio for the full model of Drechsler and Yaron (2011). Following Judd (1992), we calculate both measures over randomly chosen points across the state space. We first calculate the implied valuation ratios for each asset, denoted by $\tilde{z}_t^i$, which are given by rearranging Equation (5) to

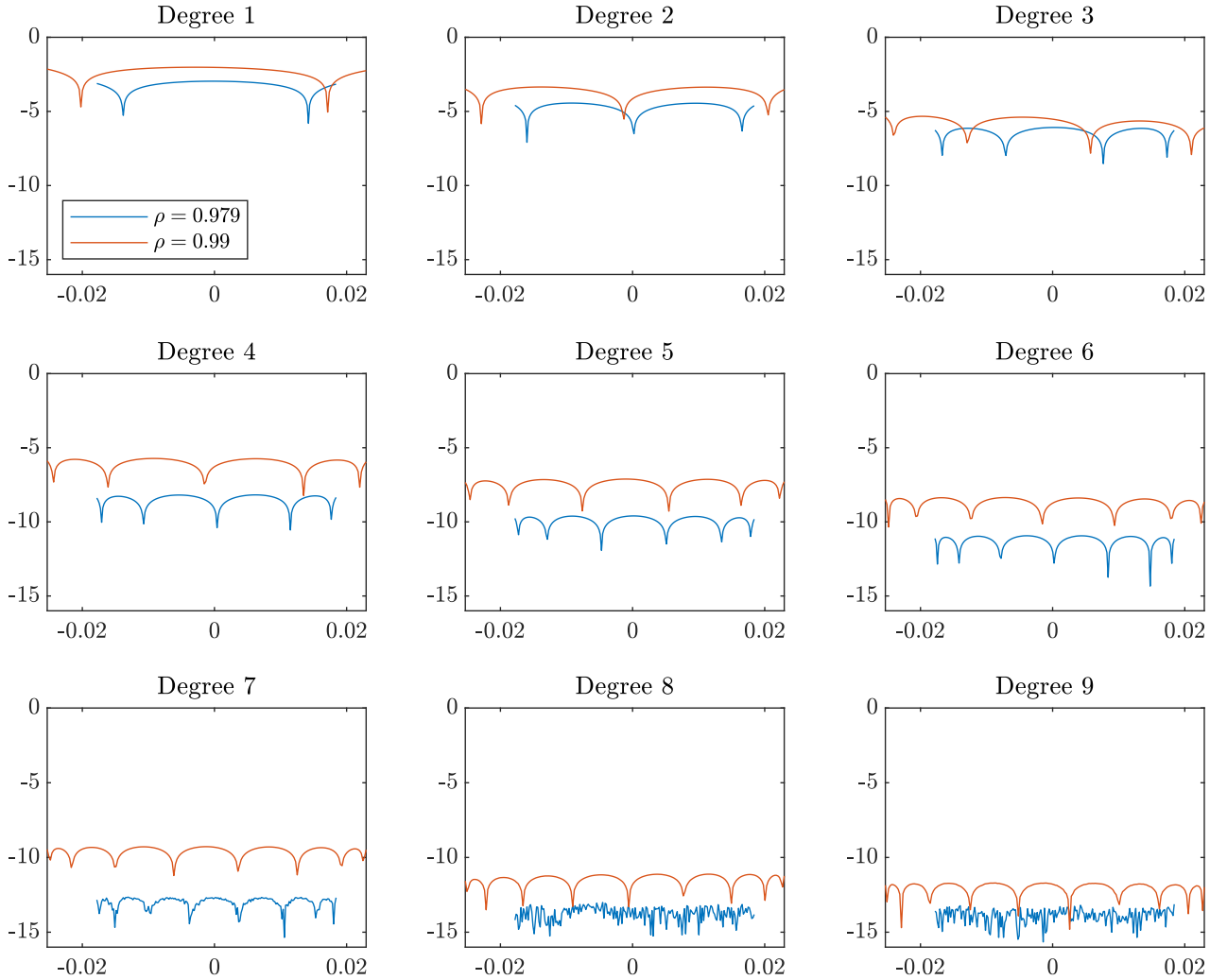
$$E_t\Bigg[\exp\Bigg(\theta\ln(\delta)-\frac{\theta}{\psi}\Delta c_{t+1}+(\theta-1)\left(\ln\left(\exp\left(\hat{z}_{t+1}^c\right)+1\right)-\hat{z}_t^c+\Delta c_{t+1}\right) \\ +\Delta d_{i,t+1}+\ln\left(\exp\left(\hat{z}_{t+1}^i\right)+1\right)\Bigg)\Bigg] = \exp\left(\tilde{z}_t^i\right). \tag{36}$$

The Euler Equation errors are defined as the relative difference between the implied values $\tilde{z}_t^i$ and the values given by our solution, denoted as $\hat{z}_t^i$, at each point:

$$\text{Euler error} = \log_{10}\left(\frac{|\exp(\tilde{z}_t^i)-\exp(\hat{z}_t^i)|}{\exp(\hat{z}_t^i)}\right). \tag{37}$$

To further check the accuracy, we plot the absolute log_{10} errors in the simplified model for different degrees of approximation, see Figure 7. The absolute errors are calculated for the price–dividend ratio by simply evaluating Equation (5) over randomly chosen points across the state space.

Figure 7: Absolute Errors



This figure shows the absolute errors in the stylized model across different values of the single state variable x_t. Each subplot shows the absolute errors for different degrees of approximation using the collocation method. All parameters except for J^x_{t+1}, which we calibrate following Drechsler and Yaron (2011), are from Bansal and Yaron (2004) and are given by $\delta = 0.998$, $\gamma = 10$, $\psi = 1.5$, $\mu_c = 0.0016$, $\mu_d = 0.0016$, $\Phi = 3$, $\sigma_t = 0.0078 \; \forall \; t$, $\omega_{cd} = 0$, $\omega_c = 1$, $\omega_d = 1$, $\phi_c = 1$, $\phi_d = 4.5$, $\phi_x = 0.044$, $l_{0,x} = 0.8/12$, $l_{1,x} = 0$, and $\phi_{\xi_x} = 3.645\phi_x$. The blue line denotes the errors for $\rho = 0.979$ (the benchmark value of Bansal and Yaron (2004)) and the orange line the errors for $\rho = 0.99$.

Table 7: Euler Errors

Euler Errors (log_{10})		
	Log-linear	**Global**
z^c		
mean	-0.0030	-9.7576
median	-0.0030	-9.6546
min	-0.0043	-13.4952
max	-0.0021	-9.0057
z^d		
mean	-1.4353	-10.2896
median	-1.3251	-10.1829
min	-5.3189	-14.3525
max	-0.7550	-9.2202

This table shows the mean, median, minimal, and maximal value of Euler Equation errors for both the log-linear and the global solution.

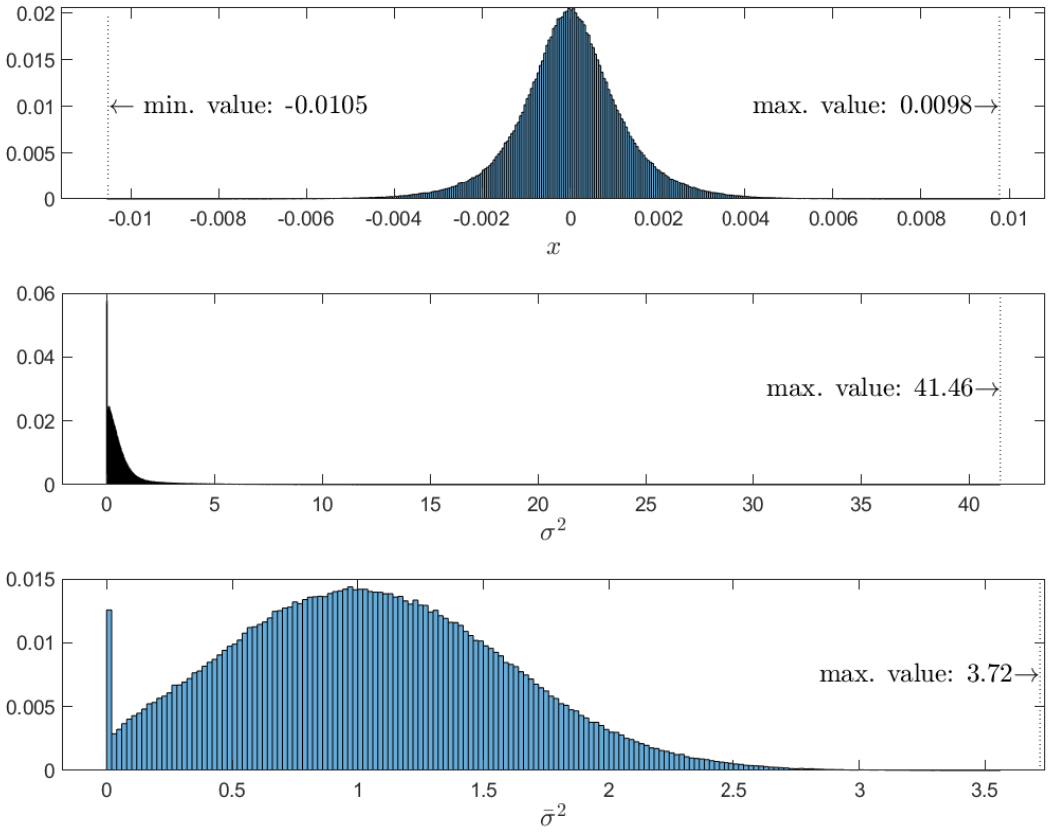
C Additional Plots and Tables

Table 8: Summary Statistics of the Stylized Model

	Data	**Model**		
Moment		**Log-linear**	**Nonlinear var**	**Global**
$E[r^f]$	0.72	2.27	2.27	2.22
$\sigma[r^f]$	3.11	1.66	1.66	1.63
$E[r^d - r^f]$	5.28	7.95	7.95	7.56
$\sigma[r^d]$	19.46	18.61	18.61	18.30
$E[z^d]$	3.41	2.52	2.52	2.57
$\sigma[z^d]$	45.21	22.70	22.70	21.96
$AC1[z^d]$	0.88	0.69	0.69	0.69
$skew[z^d]$	0.23	-0.01	-0.01	0.03
$E[Var_t^{\mathbb{P}}]$	31.72	28.02	28.42	28.36
$\sigma[Var_t^{\mathbb{P}}]$	50.44	0.00	0.12	0.83
$E[Var_t^{\mathbb{Q}}]$	41.07	31.13	30.03	29.70
$\sigma[Var_t^{\mathbb{Q}}]$	43.76	0.00	0.13	0.89
$E[VRP_t]$	9.35	3.11	1.60	1.34
$\sigma[VRP_t]$	31.62	0.00	0.01	0.06

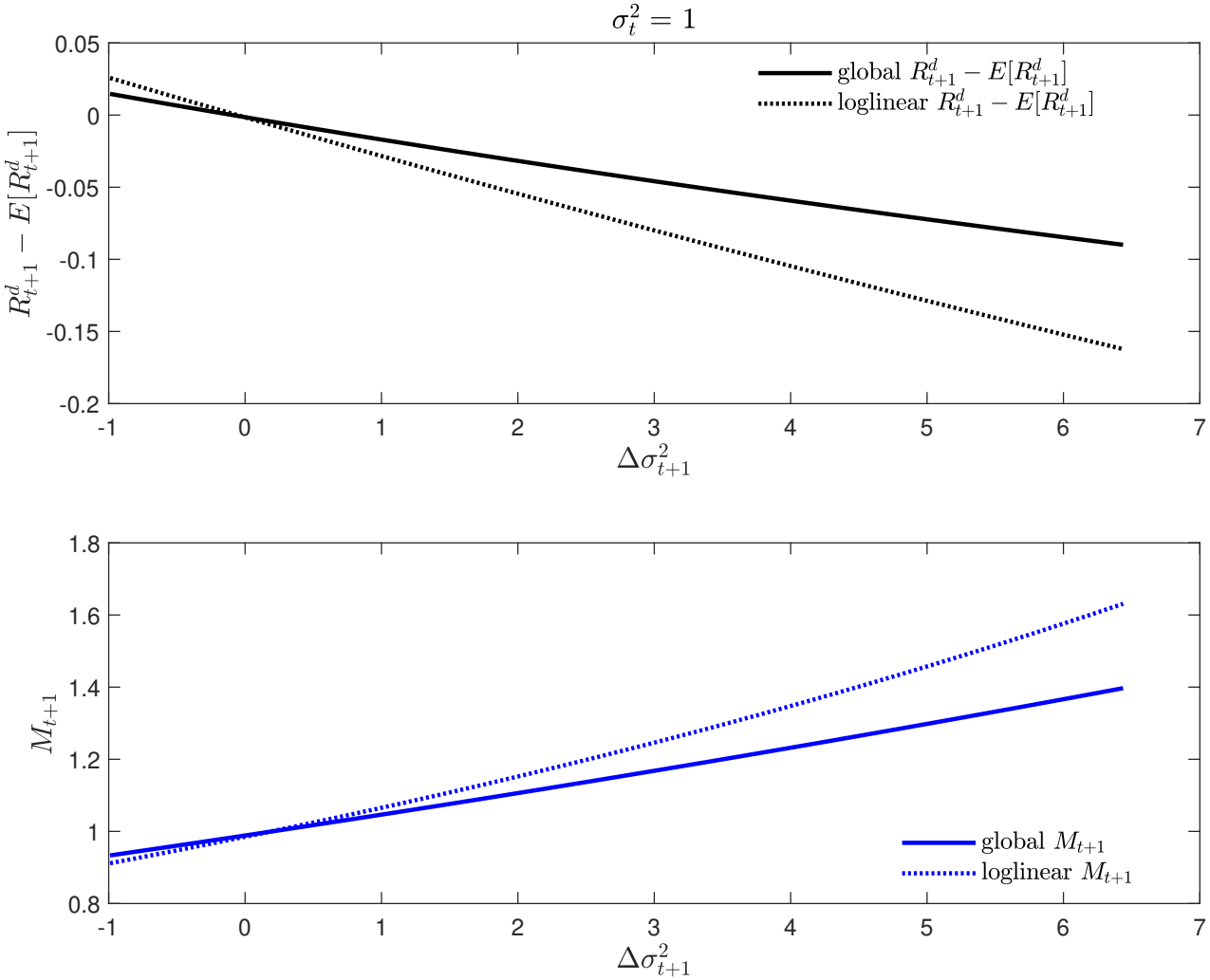
The table reports the summary statistics of the risk-free rate, return on equity, and the log price–dividend ratio for the stylized model. We follow Beeler and Campbell (2012) and Drechsler and Yaron (2011) to calculate and time-aggregate all moments. The second column displays the estimated moment of US data for the sample period 1930 to 2016, the next column the moments implied by the model solved with log-linearization and log-linear conditional variance, the fourth column the results if we use the log-linear wealth–consumption ratio and price–dividend ratio but a nonlinear conditional variance, and the last column the moments implied by the global solution. The model-implied moments are the 50% quantile values from 1,000 simulations with the same length as the data sample, where we use the same seed for each calibration to ensure comparability. All parameters except for J_{t+1}^x, which we calibrate following Drechsler and Yaron (2011), are from Bansal and Yaron (2004) and are given by $\delta = 0.998$, $\gamma = 10$, $\psi = 1.5$, $\mu_c = 0.0016$, $\rho = 0.979$, $\mu_d = 0.0016$, $\Phi = 3$, $\sigma_t = 0.0078 \; \forall \; t$, $\omega_{cd} = 0$, $\omega_c = 1$, $\omega_d = 1$, $\phi_c = 1$, $\phi_d = 4.5$, $l_{0,x} = 0.8/12$, $l_{1,x} = 0$, and $\phi_{\xi_x} = 3.645\phi_x$.

Figure 8: Ergodic Sets of the State Variables



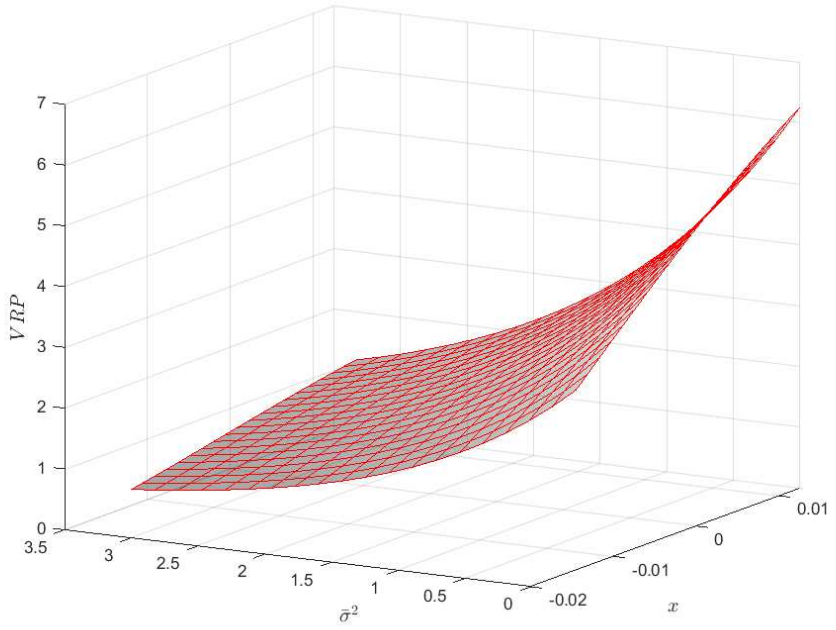
This figure shows the distributions of the state variables in the model of Drechsler and Yaron (2011). We follow Drechsler and Yaron (2011) and simulate 1,000 samples of 912 months and replace negative realizations of the volatility-related state variables by a small positive number.

Figure 9: VRP Constituents



This plot shows the pricing kernel M_{t+1} and the unexpected return $R^d_{t+1} - E_t(R^d_{t+1})$ for $\sigma_t^2 = 1$. The horizontal axis denotes the change in σ^2_{t+1} ($\Delta\sigma^2_{t+1} = \sigma^2_{t+1} - \sigma^2_t$). We set all other state variables as well as both consumption and dividend growth to their unconditional means. Then we keep all innovations except $e_{\sigma,t+1}$ and J^{σ}_{t+1} at 0, impose the shocks on σ^2_t, and calculate the pricing kernel and the unexpected return for the respective $\Delta\sigma^2_{t+1}$.

Figure 10: Sensitivity of the VRP w.r.t. $\bar{\sigma}^2$ and x



This plot shows the VRP as a function of $\bar{\sigma}^2$ and x for the calibration $\rho = 0.981$, keeping $\sigma_t^2 = 1$. We do not plot the log-linear VRP as it would be a plane parallel to the xy plane.

References

Abel, A. B. (1999): "Risk premia and term premia in general equilibrium," *Journal of Monetary Economics*, 43, 3–33.

Andersen, T. G., L. Benzoni, and J. Lund (2002): "An empirical investigation of continuous-time equity return models," *The Journal of Finance*, 57, 1239–1284.

Andersen, T. G., T. Bollerslev, F. X. Diebold, and P. Labys (2003): "Modeling and Forecasting Realized Volatility," *Econometrica*, 71, 579–625.

Andersen, T. G., N. Fusari, and V. Todorov (2015): "The Risk Premia Embedded in Index Options," *Journal of Financial Economics*, 3, 558–584.

Baele, L., J. Driessen, S. Ebert, J. M. Londono, and O. G. Spalt (2018): "Cumulative prospect theory, option returns, and the variance premium," *The Review of Financial Studies*, 32, 3667–3723.

Bakshi, G., C. Cao, and Z. Chen (1997): "Empirical performance of alternative option pricing models," *The Journal of Finance*, 52, 2003–2049.

Bakshi, G., N. Kapadia, and D. Madan (2003): "Stock Return Characteristics, Skew Laws, and the Differential Pricing of Individual Equity Options," *The Review of Financial Studies*, 16, 101–143.

Bansal, R., D. Kiku, and A. Yaron (2012): "An Empirical Evaluation of the Long-Run Risks Model for Asset Prices," *Critical Finance Review*, 1, 183–221.

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

Beeler, J. and J. Y. Campbell (2012): "The Long-Run Risks Model and Aggregate Asset Prices: An Empirical Assessment," *Critical Finance Review*, 1, 141–182.

Bekaert, G. and E. Engstrom (2017): "Asset Return Dynamics under Habits and Bad Environment–Good Environment Fundamentals," *Journal of Political Economy*, 125, 713–760.

Bekaert, G., E. C. Engstrom, and N. R. Xu (2019): "The Time Variation in Risk Appetite and Uncertainty," Working Paper.

Bekaert, G. and M. Hoerova (2016): "What do Asset Prices have to say about Risk Appetite and Uncertainty?" *Journal of Banking & Finance*, 67, 103–118.

Benzoni, L., P. Collin-Dufresne, and R. S. Goldstein (2011): "Explaining asset pricing puzzles associated with the 1987 market crash," *Journal of Financial Economics*, 101, 552–573.

Bollerslev, T., G. Tauchen, and H. Zhou (2009): "Expected Stock Returns and Variance Risk Premia," *The Review of Financial Studies*, 22, 4463–4492.

Bollerslev, T. and V. Todorov (2011): "Tails, Fears, and Risk Premia," *The Journal of Finance*, 66, 2165–2211.

Bollerslev, T., L. Xu, and H. Zhou (2015): "Stock Return and Cash Flow Predictability: The Role of Volatility Risk," *Journal of Econometrics*, 187, 458–471.

Bonomo, M., R. Garcia, N. Meddahi, and R. Tédongap (2015): "The Long and the Short of the Risk-Return Trade-Off," *Journal of Econometrics*, 187, 580–592.

Branger, N., P. Rodrigues, and C. Schlag (2018): "Level and Slope of Volatility Smiles in Long-Run Risk Models," *Journal of Economic Dynamics and Control*, 86, 95–122.

Britten-Jones, M. and A. Neuberger (2000): "Option Prices, Implied Price Processes, and Stochastic Volatility," *The Journal of Finance*, 55, 839–866.

Campbell, J. Y. and R. J. Shiller (1988): "Stock prices, earnings, and expected dividends," *The Journal of Finance*, 43, 661–676.

Carr, P. and L. Wu (2009): "Variance risk premiums," *The Review of Financial Studies*, 22, 1311–1341.

Christoffersen, P., M. Fournier, K. Jacobs, and M. Karoui (2017): "Option-Based Estimation of the Price of Co-Skewness and Co-Kurtosis Risk," Rotman School of Management Working Paper No. 2656412.

Corsi, F. (2009): "A simple approximate long-memory model of realized volatility," *Journal of Financial Econometrics*, 7, 174–196.

Coval, J. D. and T. Shumway (2001): "Expected Option Returns," *The Journal of Finance*, 56, 983–1009.

Demeterfi, K., E. Derman, M. Kamal, and J. Zou (1999): "A Guide to Volatility and Variance Swaps," *The Journal of Derivatives*, 6, 9–32.

Dew-Becker, I., S. Giglio, A. Le, and M. Rodriguez (2017): "The price of variance risk," *Journal of Financial Economics*, 123, 225–250.

Drechsler, I. (2013): "Uncertainty, Time-Varying Fear, and Asset Prices," *The Journal of Finance*, 68, 1843–1889.

Drechsler, I., A. Moreira, and A. Savov (2019): "Liquidity Creation as Volatility Risk," Working Paper.

Drechsler, I. and A. Yaron (2011): "What's Vol Got to Do with It," *The Review of Financial Studies*, 24, 1–45.

Duffie, D., J. Pan, and K. Singleton (2000): "Transform analysis and asset pricing for affine jump-diffusions," *Econometrica*, 68, 1343–1376.

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

Eraker, B. (2008): "Affine General Equilibrium Models," *Management Science*, 54, 2068–2080.

Eraker, B. and I. Shaliastovich (2008): "An Equilibrium Guide to designing Affine Pricing Models," *Mathematical Finance*, 18, 519–543.

Fama, E. F. and K. R. French (1989): "Business conditions and expected returns on stocks and bonds," *Journal of Financial Economics*, 25, 23–49.

Feller, W. (1968): *An Introduction to Probability Theory and its Applications*, vol. 1, NY: Wiley, 3 ed.

Fernández-Villaverde, J. and O. Levintal (2018): "Solution methods for models with rare disasters," *Quantitative Economics*, 9, 903–944.

Garleanu, N., L. H. Pedersen, and A. M. Poteshman (2008): "Demand-based option pricing," *The Review of Financial Studies*, 22, 4259–4299.

He, Z., B. Kelly, and A. Manela (2017): "Intermediary asset pricing: New evidence from many asset classes," *Journal of Financial Economics*, 126, 1–35.

Jiang, G. J. and Y. S. Tian (2005): "The Model-Free Implied Volatility and Its Information Content," *The Review of Financial Studies*, 18, 1305–1342.

Judd, K. L. (1992): "Projection Methods for Solving Aggregate Growth Models," *Journal of Economic Theory*, 58, 410–452.

Judd, K. L., L. Maliar, and S. Maliar (2011): "Numerically stable and accurate stochastic simulation approaches for solving dynamic economic models," *Quantitative*

Economics, 2, 173–210.

Kelly, B. and H. Jiang (2014): "Tail risk and asset prices," *The Review of Financial Studies*, 27, 2841–2871.

Kilic, M. and I. Shaliastovich (2018): "Good and bad variance premia and expected returns," *Management Science*, 65, 2522–2544.

Lettau, M. and S. C. Ludvigson (2010): "Measuring and Modeling Variation in the Risk-Return Trade-off," in *Handbook of Financial Econometrics: Tools and Techniques*, ed. by Y. Aït-Sahalia and L. P. Hansen, San Diego: North-Holland, vol. 1 of *Handbooks in Finance*, 617–690.

Liu, J., J. Pan, and T. Wang (2005): "An Equilibrium Model of Rare-Event Premia and Its Implication for Option Smirks," *The Review of Financial Studies*, 18, 131–164.

Londono, J. M. and N. R. Xu (2019): "Variance Risk Premium Components and International Stock Return Predictability," Working Paper.

Lorenz, F. and M. Schumacher (2018): "Downside Risks and the Price of Variance Uncertainty," Working Paper.

Merton, R. C. (1976): "Option pricing when underlying stock returns are discontinuous," *Journal of Financial Economics*, 3, 125–144.

Nagel, S. and Z. Xu (2019): "Asset Pricing with Fading Memory," National Bureau of Economic Research Working.

Piazzesi, M. (2001): "An econometric model of the yield curve with macroeconomic jump effects," NBER Working paper no. 8246.

Pohl, W., K. Schmedders, and O. Wilms (2018): "Higher-Order Effects in Asset-Pricing Models with Long-Run Risks," *The Journal of Finance*, 73, 1061–1111.

——— (2019): "Relative Existence for Recursive Utility," Working Paper.

Pyun, S. (2019): "Variance risk in aggregate stock returns and time-varying return predictability," *Journal of Financial Economics*, 132, 150–174.

Schreindorfer, D. (forthcoming): "Macroeconomic Tail Risks and Asset Prices," *The Review of Financial Studies*.

Tsai, J. and J. A. Wachter (2018): "Pricing long-lived securities in dynamic endowment economies," *Journal of Economic Theory*, 177, 848–878.

Wachter, J. A. (2013): "Can Time-Varying Risk of Rare Disasters Explain Aggregate Stock Market Volatility?" *The Journal of Finance*, 68, 987–1035.

Wachter, J. A. and Y. Zhu (2019): "Learning with Rare Disasters," Working Paper.

Weil, P. (1990): "Nonexpected Utility in Macroeconomics." *Quarterly Journal of Economics*, 105, 29–42.

Zhou, H. (2018): "Variance risk premia, asset predictability puzzles, and macroeconomic uncertainty," *Annual Review of Financial Economics*, 10, 481–497.