The Journal of FINANCE
The Journal of THE AMERICAN FINANCE ASSOCIATION

THE JOURNAL OF FINANCE • VOL. LXXIII, NO. 3 • JUNE 2018

Higher Order Effects in Asset Pricing Models with Long-Run Risks

WALTER POHL, KARL SCHMEDDERS, and OLE WILMS*

ABSTRACT

This paper shows that the latest generation of asset pricing models with long-run risk exhibit economically significant nonlinearities, and thus the ubiquitous Campbell-Shiller log-linearization can generate large numerical errors. These errors translate in turn to considerable errors in the model predictions, for example, for the magnitude of the equity premium or return predictability. We demonstrate that these nonlinearities arise from the presence of multiple highly persistent processes, which cause the exogenous states to attain values far away from their long-run means with nonnegligible probability. These extreme values have a significant impact on asset price dynamics.

BANSAL AND YARON (2004) INTRODUCE TWO important innovations to the asset pricing literature. In particular, they put forward an economic mechanism based on long-run risk to explain many empirical asset pricing facts, and they demonstrate that the Campbell-Shiller (1988) log-linearization technique provides a simple method for analyzing such models. These two contributions together have served as the foundation for a large literature on the ability of long-run risk to solve empirical puzzles. By necessity, log-linearization neglects higher order effects in asset pricing models. In this paper, we ask whether these effects matter for the latest generation of long-run risk models. We show that they do. More specifically, we demonstrate that, for many economically plausible choices of parameters and exogenous processes, the errors introduced by log-linearization can themselves be economically significant. In fact, for highly persistent processes, as regularly used in the literature, the approximation errors in moments can exceed 70%.

*Walter Pohl is with Department of Finance, NHH Norwegian School of Economics. Karl Schmedders is with Department of Business Administration, University of Zurich. Ole Wilms is with Department of Finance, Tilburg University. This paper was previously circulated under the title "Higher-Order Dynamics in Asset Pricing Models with Recursive Preferences." We are grateful to two anonymous referees, an anonymous Associate Editor, and the Editor, Stefan Nagel, for thoughtful reviews of earlier versions. We are indebted to Nicole Branger, Lars Hansen, Ken Judd, Martin Lettau, Malte Schumacher, Che-Lin Su, and particularly our discussant Alexis Akira Toda for helpful discussions on the subject. We thank seminar audiences at the University of Zurich, the Becker Friedman Institute, the University of Münster, and various conferences for comments. We are grateful to the copy editor, Brenda Priebe, for excellent editorial support. Ole Wilms gratefully acknowledges financial support from the Zürcher Universitätsverein. The authors do not have any conflicts of interest as per the *Journal of Finance* disclosure policy.

DOI: 10.1111/jofi.12615

Asset pricing models have become increasingly complex over the last three decades. The first generation of such models, developed in the 1980s (Grossman and Shiller (1981), Hansen and Singleton (1982), Mehra and Prescott (1985)), proved inadequate in explaining key features of financial markets, such as the high equity premium and the low risk-free rate. As the literature on asset pricing has evolved and matured over time, researchers have added more complex elements to their models such as incomplete markets (e.g., uninsurable income risks), frictions (e.g., borrowing or collateral constraints), time-varying risk aversion, and heterogeneous expectations. While these additional features have had varying degrees of success, the new generation of long-run risk models (e.g., Bansal and Yaron (2004) or Hansen, Heaton, and Li (2008)), with their interplay of persistent components in consumption growth and recursive preferences, have had considerably more success in resolving long-standing asset pricing puzzles. The key feature of these models is the combination of a preference for the early resolution of risk together with highly persistent state processes that potentially affect long-run model outcomes.

Complex models generally require numerical solution techniques. Bansal and Yaron (2004) show that a simple linearized solution method based on the Campbell-Shiller (1988) present-value relation works well for their original model because the log price-dividend ratio in the model is approximately a linear function of the underlying shocks. Additionally, the linearization procedure has the attractive property of lending itself to a simple analysis of the economic impact of different shocks. This property is particularly appealing as it allows the researcher to draw conclusions about parameter dependencies and economic mechanisms in the model. Accordingly, a large group of researchers have followed Bansal and Yaron (2004) and used the log-linearization technique to solve asset pricing models with recursive preferences (e.g., Bollerslev, Tauchen, and Zhou (2009), Bansal, Kiku, and Yaron (2010), Koijen et al. (2010), Constantinides and Ghosh (2011), Drechsler and Yaron (2011), Bansal, Kiku, and Yaron (2012), Beeler and Campbell (2012), Bansal and Shaliastovich (2013), Bansal et al. (2014), and Segal, Shaliastovich, and Yaron (2015), among others). Examining this strand of literature, it is difficult to find studies that do not rely on the Campbell-Shiller (1988) approach[1]—it has become the standard method for solving asset pricing models with long-run risk.

To better match features of the data, the tendency in the long-run risk literature has been toward both higher persistence and greater model complexity. This development suggests that it is time to take stock of whether the Campbell-Shiller (1988) approximation is still appropriate. To address this question, we first identify the factors that can make higher order dynamics matter for the benchmark long-run risk model with stochastic volatility. We show that, for

[1] Prominent exceptions that do not rely on log-linearized solutions are the studies by Croce, Lettau, and Ludvigson (2015) and Collin-Dufresne, Johannes, and Lochstoer (2016). Kaltenbrunner and Lochstoer (2010) use log-linearized solutions to infer qualitative model implications but use value function iteration for their quantitative results. They approximate the value function with Chebyshev polynomials.

highly persistent processes, these effects do matter. Economically, when shocks are highly persistent, agents can spend longer periods trapped in undesirable parts of the state space, which can have a large impact on investor behavior. The problem is particularly severe when both long-run risk and stochastic volatility are persistent. In that case, the interaction between long-run risk and stochastic volatility becomes especially significant when the long-run growth rate is low but the volatility of the growth rate is high, which in turn creates significant deviations from linearity. As a side effect, this interaction leads to large errors in predictability regressions using simulated data: the log-linearized solutions overstate the predictability of returns and understate the predictability of dividends by the log price-dividend ratio, relative to a higher precision solution. Interestingly, the persistence is a problem only in the presence of a strong preference for the early resolution of risk. However, without a preference for the early resolution of risk, it is difficult to generate a high equity premium.

To show that this nonlinearity matters for real models, and is not just an intellectual exercise, we examine the consequences of ignoring nonlinear dynamics in the Bansal and Yaron (2004) model as well as five other recent studies, namely, the newly calibrated version of Bansal, Kiku, and Yaron (2012), the extensive calibration study of Schorfheide, Song, and Yaron (2018), the volatility-of-volatility model of Bollerslev, Xu, and Zhou (2015), and the studies on real and nominal bonds of Koijen et al. (2010) and Bansal and Shaliastovich (2013). We show that the higher order dynamics not captured by the Campbell-Shiller (1988) approximation can be large and economically significant. For example, for the calibration of Bansal, Kiku, and Yaron (2012), which is an improved calibration of the original Bansal and Yaron (2004) model (in particular, with greater persistence in shocks to stochastic volatility), we find that the log-linearization introduces economically significant errors in model predictions. The equity premium, for instance, is overestimated by about 100 bps and the relative error in the volatility of the price-dividend ratio exceeds 20%. The true amount of return predictability is approximately half of what log-linearization would suggest.

In the Bansal, Kiku, and Yaron (2012) model, most of the nonlinearities are introduced by the pricing of equities, and using a linearized solution for the stochastic discount factor does not lead to economically significant errors. This approach does not generally work, however; for other model specifications analyzed in this paper, we find significant errors in the stochastic discount factor. Therefore, linearized solutions should not be used for any of the return equations without carefully testing its accuracy for the specific asset pricing model under consideration.

Schorfheide, Song, and Yaron (2018) perform a Bayesian estimation of the model using the same approximation and find evidence of higher persistence for long-run risk. In this case, we find approximation errors as large as 70% for some key model moments for the 95% quantiles of the state persistence parameters (for the median estimates they are significantly lower). When persistence is very high, log-linearization can actually invert the slope of the yield curve in

the nominal bond models of Bansal and Shaliastovich (2013) and Koijen et al. (2010).

Long-run risk has proven to be an illuminating angle on asset pricing puzzles, and the Campbell-Shiller (1988) log-linearization provides a good first approximation to understand shocks' propagation mechanisms. But for complex specifications with highly persistent processes, the dynamics have proven to be richer than they seem at first glance, with significant economic implications. Since the data strongly incline researchers to consider such models, it is important to correctly describe these rich dynamics to fully understand them.

The paper is organized as follows. Section I describes the general model framework that we use throughout the paper. In Section II, we identify the factors that introduce higher order dynamics to asset pricing models with long-run risk. Section III analyzes the implications of higher order effects for six recent asset pricing studies. Section IV concludes.

I. Model Framework

We consider a standard asset pricing model with a representative agent and recursive preferences as in Epstein and Zin (1989) and Weil (1990). Indirect utility at time t, V_t, is given recursively by

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

In this parameterization, C_t is consumption, δ is the time discount factor, γ determines the level of relative risk aversion, and $\theta = \frac{1-\gamma}{1-\frac{1}{\psi}}$, where ψ is the elasticity of intertemporal substitution (EIS). The parameters γ and ψ are required to satisfy $0 < \gamma, \psi$, and $\psi \neq 1$.[2] For $\theta = 1$, the agent has standard constant relative risk aversion (CRRA) preferences. Values of $\gamma > 1/\psi$ indicate a preference for the early resolution of risk and values of $\gamma < 1/\psi$ indicate a preference for late resolution. The general asset pricing equation to price any asset i with ex-dividend price $P_{i,t}$ and dividend $D_{i,t}$ is given by

$$\mathrm{E}_t[M_{t+1}R_{i,t+1}] = 1, \tag{2}$$

where $R_{i,t+1} = \frac{P_{i,t+1}+D_{i,t+1}}{P_{i,t}}$. For recursive preferences, the stochastic discount factor M_{t+1} is given by

$$M_{t+1} = \delta\left(\frac{C_{t+1}}{C_t}\right)^{-\frac{1}{\psi}}\left(\frac{V_{t+1}}{\left[\mathrm{E}_t\left(V_{t+1}^{1-\gamma}\right)\right]^{\frac{1}{1-\gamma}}}\right)^{\frac{1}{\psi}-\gamma}. \tag{3}$$

[2] We consider the special case of $\psi = 1$ in Section C.

Epstein and Zin (1989) show that the (unobserved) value of aggregate wealth, W_t, can be expressed in terms of the value function,

$$W_t = \frac{V_t^{1-1/\psi}}{(1-\delta)C_t^{-1/\psi}}. \tag{4}$$

This expression allows us to express M_{t+1} in terms of the gross return to the claim on aggregate consumption $R_{w,t+1}$,

$$M_{t+1} = \delta^{\theta}\left(\frac{C_{t+1}}{C_t}\right)^{-\frac{\theta}{\psi}} R_{w,t+1}^{\theta-1}, \tag{5}$$

where $R_{w,t+1} = \frac{W_{t+1}}{W_t - C_t}$. Since equation (2) has to hold for all assets, it must also hold for the return of the aggregate consumption claim. Thus, $R_{w,t+1}$ is determined by the wealth-Euler equation

$$\mathrm{E}_t\left[\delta^{\theta}\left(\frac{C_{t+1}}{C_t}\right)^{-\frac{\theta}{\psi}} R_{w,t+1}^{\theta}\right] = 1. \tag{6}$$

We also define $R_{m,t+1}$ as the gross return to the claim on aggregate dividends with log dividend growth, Δd_{t+1}. The gross risk-free return is given by $R_{t+1}^f = \mathrm{E}_t[M_{t+1}]^{-1}$ and the (unconditional) equity premium is defined as $\mathrm{E}(R_{m,t} - R_{f,t})$.

Throughout the paper, we consider different setups for the specification of log consumption growth, Δc_{t+1}. The original specification in the seminal long-run risk model of Bansal and Yaron (2004) uses a single volatility process that drives uncertainty in the economy, σ_t, and a long-run growth process, x_t, that affects both log consumption growth and log dividend growth. Specifically, the four processes are given as follows:

$$\begin{aligned}
\Delta c_{t+1} &= \mu_c + x_t + \phi_c \sigma_t \eta_{c,t+1}, \\
x_{t+1} &= \rho x_t + \phi_x \sigma_t \eta_{x,t+1}, \\
\sigma_{t+1}^2 &= \bar{\sigma}^2(1-\nu) + \nu\sigma_t^2 + \phi_\sigma \omega_{t+1}, \\
\Delta d_{t+1} &= \mu_d + \Phi x_t + \phi_d \sigma_t \eta_{d,t+1} + \phi_{d,c}\sigma_t \eta_{c,t+1} \\
&\quad \times \eta_{c,t+1}, \eta_{x,t+1}, \eta_{d,t+1}, \omega_{t+1} \sim \text{i.i.d. } N(0,1).
\end{aligned} \tag{7}$$

We elaborate on the properties and key features of the model at the beginning of Section II. In the remainder of the paper, we consider variations of this setup that include different specifications for the stochastic volatility processes as well as additional state processes such as volatility-of-volatility or inflation.

The common approach to solving the long-run risk models in the finance literature is to log-linearize the model. For examples of this approach, see Bollerslev, Tauchen, and Zhou (2009), Koijen et al. (2010), Bansal, Kiku, and Yaron (2010), Constantinides and Ghosh (2011), Drechsler and Yaron (2011), Beeler and Campbell (2012), Bansal, Kiku, and Yaron (2012), Bansal and Shaliastovich (2013), Bansal et al. (2014), and Segal, Shaliastovich, and Yaron (2015), among others. By construction, however, log-linearization misses the influence

of higher order dynamics—that is, it does not attempt to approximate nonlinear features of the exact solution. But what if these features matter quantitatively for equilibrium outcomes? Does log-linearization still deliver sufficiently accurate approximations of the exact solution?

We address these critical questions in this paper. To do so, we need an alternative solution method that accurately accounts for higher order dynamics and yields robust solutions. In the body of the paper, we use projection methods (see Judd (1992)), which capture these nonlinear effects and are known to converge to the true solution (see Atkinson (1992)). For a stochastic growth model with Epstein-Zin utility, Caldara et al. (2012, p. 189) note that projection methods "provide a terrific level of accuracy with reasonable computational burden." The choice of projection method is not essential, in that other methods known to converge to the true solution, such as Tauchen and Hussey (1991), get similar results, though projection methods seem to do so with less computational cost. In Appendix A, we discuss the pros and cons of different computational methods for asset pricing models, and we provide a detailed description of the log-linear and projection solution methods.

II. Higher Order Dynamics in Long-Run Risk Models

We begin our analysis by identifying the factors that influence the higher order dynamics in long-run risk models. To do so, we use the standard long-run risk framework of Bansal and Yaron (2004) (see the four processes in equation (7)). Key features of the model are the highly persistent state processes for the long-run growth rate, x_t, and stochastic volatility, σ_t. As we demonstrate later, the model requires that the persistence parameters ρ and ν be very close to one, as otherwise the model predictions do not match the data. Combining these highly persistent processes with a preference for the early resolution of risk ($\gamma > \frac{1}{\psi}$), Bansal and Yaron (2004) are able to explain many puzzling features of financial markets such as the high equity premium together with a low risk-free rate and the volatility of the market return, of the risk-free rate, and of the price-dividend ratio. As Beeler and Campbell (2012) observe, the original calibration of Bansal and Yaron (2004) implies, counterfactually, that consumption and dividend growth are highly predictable from stock prices. In response to this observation, Bansal, Kiku, and Yaron (2012) recalibrate the model to better match the level of predictability in the data.[3] We use this calibration to analyze the impact of higher order dynamics on quantitative equilibrium outcomes in the long-run risk model.

A. *Parameter Values and Numerical Errors*

In Figure 1, we plot the log price-dividend ratio, $p_t - d_t$, in the model of Bansal, Kiku, and Yaron (2012) as a function of the states x_t and σ_t^2. The dark

[3] Bansal, Kiku, and Yaron (2012) increase the influence of the stochastic volatility channel relative to the long-run risk channel to reduce the predictability of consumption and dividend growth.

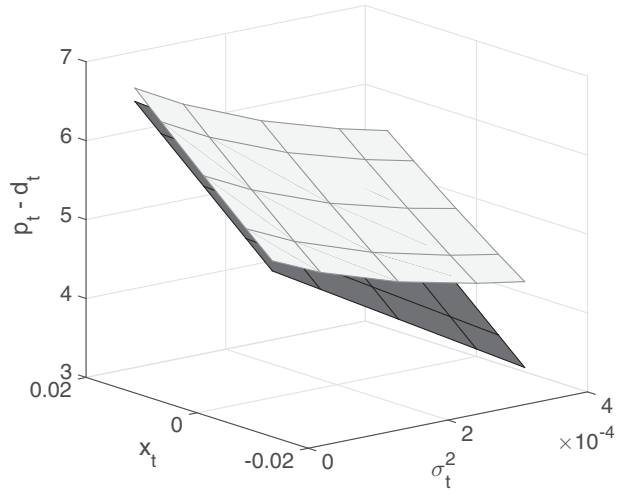

Figure 1. Log price-dividend ratio in the long-run risk model. The figure plots the log-linearized solution for the log price-dividend ratio (dark gray surface) as well as the global solution (transparent gray surface) as a function of the states x_t and σ_t^2. Parameters: $\delta = 0.9989$, $\gamma = 10$, $\psi = 1.5$, $\mu = 0.0015$, $\phi_c = 1$, $\phi_x = 0.038$, $\rho = 0.975$, $\bar{\sigma} = 0.0072$, $\nu = 0.999$, $\phi_\sigma =$ 2.8e-6, $\mu_d = 0.0015$, $\Phi = 2.5$, $\phi_d = 5.96$, and $\phi_{d,c} = 2.6$.

gray surface shows the log-linearized solution as used in Bansal, Kiku, and Yaron (2012) and the transparent gray surface plots the high-precision solution obtained by the projection approach.[4] As can be seen, the log-linearization systematically underestimates the log price-dividend ratio and produces a steeper ratio in both state dimensions. An underestimation of $p_t - d_t$ implies an overestimation of the equity risk premium. Hence, an analysis relying on the log-linearized solution falsely predicts a larger equity premium than does an analysis based on the higher precision solution. Also, the greater steepness in the log price-dividend ratio as a function of the two state variables implies a larger volatility of the price-dividend ratio. As both the equity premium and

[4] Appendix B provides a formal analysis of the accuracy of the projection approach. To compute accurate solutions with the projection method, we increase the approximation interval and the polynomial approximation degree until the solutions no longer change and the polynomial coefficients for the highest degree polynomial are close to zero. This approach ensures that we capture the higher order dynamics introduced by the tails of the state processes. A key issue in specification (7) is that the variance process σ_t^2 can become negative. The log-linearization "ignores" this effect as it linearizes around a reference point that is well above zero. When using a global method—such as the projection approach—the best approach is not obvious. We, therefore, conducted several tests using different setups for the state space and the evaluation of negative realizations. These show that the treatment of the negative range of the variance process has a negligible influence on the equilibrium outcomes of the global method. Figure 1 reveals the reason for this robust property. The difference between the log-linear solution and the global solution is large for large σ_t^2, but for small σ_t^2 there is almost no difference and the log price-dividend ratio becomes almost linear. Hence, the domain subset with small volatility values, where the errors are small, has a negligible influence on equilibrium outcomes.

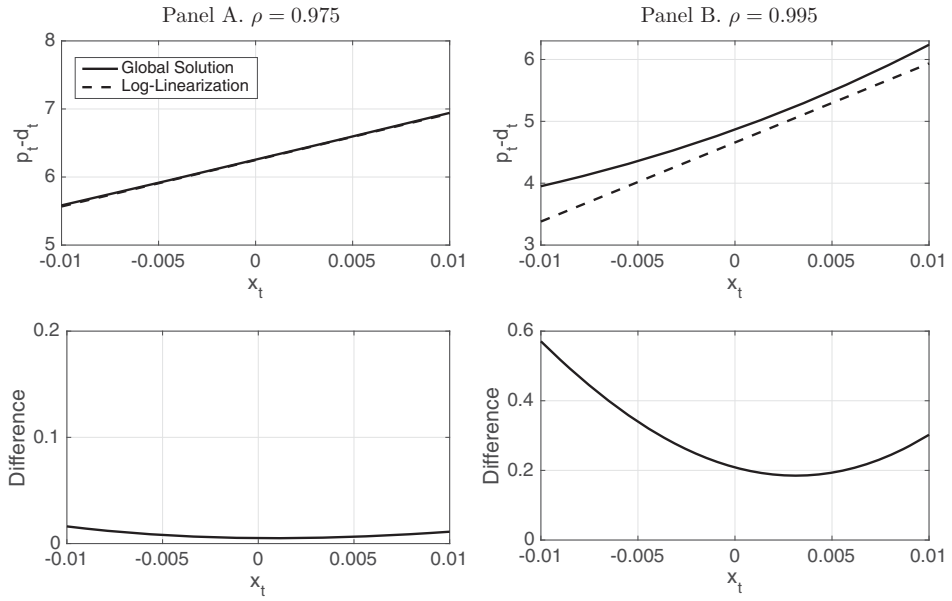

Figure 2. State persistence and errors in the log price-dividend ratio. The figure plots the log price-dividend ratio in the long-run risk model (7) with constant volatility ($\sigma_t = \bar{\sigma}\ \forall\ t$). The left panel shows the results for $\rho = 0.975$ and the right panel for $\rho = 0.995$. In the top row, the solid line shows the global solution and the dashed line shows the solution from the Campbell-Shiller (1988) log-linearization. The figures in the bottom row show the differences between the two solutions. Parameters: $\delta = 0.9989$, $\gamma = 10$, $\psi = 1.5$, $\mu = 0.0015$, $\phi_c = 1$, $\phi_x = 0.038$, $\bar{\sigma} = 0.0072$, $\mu_d = 0.0015$, $\Phi = 2.5$, $\phi_d = 5.96$, and $\phi_{d,c} = 2.6$.

the volatility of the price-dividend ratio are key quantities that the long-run risk model tries to explain, the systematic overestimation of these features is an unpleasant property of the solution method. In practice, of course, it is only relevant if the differences are numerically large enough to matter. We document that they are, and that these errors are particularly severe when the underlying state processes are highly persistent.

To understand the influence of the persistence of the state processes on the nonlinearities of the model, consider first the simplified example of a long-run risk model with constant volatility and hence only one state process (specification (7) with $\sigma_t = \bar{\sigma}\ \forall\ t$). In Figure 2, we plot the log price-dividend ratio obtained by the Campbell-Shiller (1988) log-linearization, the global solution, as well as the difference between the two as a function of the state x_t. The left panel shows the results for $\rho = 0.975$ and the right panel for $\rho = 0.995$. We find that for $\rho = 0.975$ the log price-dividend ratio is almost linear and the difference between the two solutions is close to zero, while for $\rho = 0.995$ the nonlinearity in the global solution of the log price-dividend ratio becomes more pronounced, and thus the difference between the two solutions increases strongly. The difference is especially large for small and large x_t, which

highlights the nonlinear nature of the log price-dividend ratio for highly persistent state processes. Since the long-run risk model requires that persistent processes be consistent with the data, this feature suggests that the limitations of log-linearization are of considerable practical significance.

Figure 3 plots errors in the log price-dividend ratio introduced by the log-linearization for different values for the persistence of long-run risk ρ and stochastic volatility ν for the full long-run risk model with stochastic volatility. The left panel shows relative errors in the unconditional mean of the log price-dividend ratio and the right panel shows the relative errors in the unconditional volatility. We find that the errors increase dramatically with the persistence of the state processes ρ and ν, with the errors in the volatility of the log price-dividend ratio becoming as large as 50% for large values of ρ and ν. To demonstrate that these calibrations are not artificially constructed to obtain large errors, we also plot the point estimates for ρ and ν used in the studies of Bansal and Yaron (2004) (BY), Bansal, Kiku, and Yaron (2012) (BKY), and Schorfheide, Song, and Yaron (2018). For the study of Schorfheide, Song, and Yaron (2018), we plot the median estimates (SSY1) as well as the 95% estimates (SSY2) to demonstrate the range of parameters, and hence errors, that are included within the estimation procedure. We find that the parameter values used in the studies—except for those in the study of Bansal and Yaron (2004)—are in the area where approximation errors are large and significant. For example, in Bansal, Kiku, and Yaron (2012), the volatility of the log price-dividend ratio is overestimated by more than 20%.[5]

While Figure 3 plots errors in the moments of the monthly log price-dividend ratio, asset pricing models such as Bansal and Yaron (2004) and Bansal, Kiku, and Yaron (2012) are usually calibrated to match annualized market outcomes. Therefore, in Figures 4 and 5, we plot the annualized equity premium $E(R_{m,t} - R_{f,t})$ and the annual volatility of the log price-dividend ratio, respectively, obtained by the log-linearized solution as well as the global solution as a function of the risk aversion, γ, and the EIS, ψ, in the first row and the serial correlations in the long-run risk channel, ρ, and the stochastic volatility channel, ν, in the second row. We find that for this particular calibration, for a risk aversion of approximately five, the log-linearized solution basically coincides with the solution from the projection approach, which suggests that a linear solution gives a reasonable approximation of the model. However, for this small a risk aversion the implied model moments collapse, with an equity premium slightly above 2% and a sharp decrease in the volatility of the log price-dividend ratio. When we increase the risk aversion, the errors in the equity premium and the volatility of the log price-dividend ratio increase significantly, with a large overestimation of both quantities. Furthermore, in line with the previous results, the accuracy depends highly on the persistence of the processes

[5] Note that, except for the study of Bansal, Kiku, and Yaron (2012), the values reported for the errors do *not* correspond to the errors in the studies, as the authors use different calibrations for the other model parameters. The exercise serves to demonstrate the potential errors introduced when using log-linearization to solve highly persistent models. A full evaluation of the different studies is conducted in Section III.

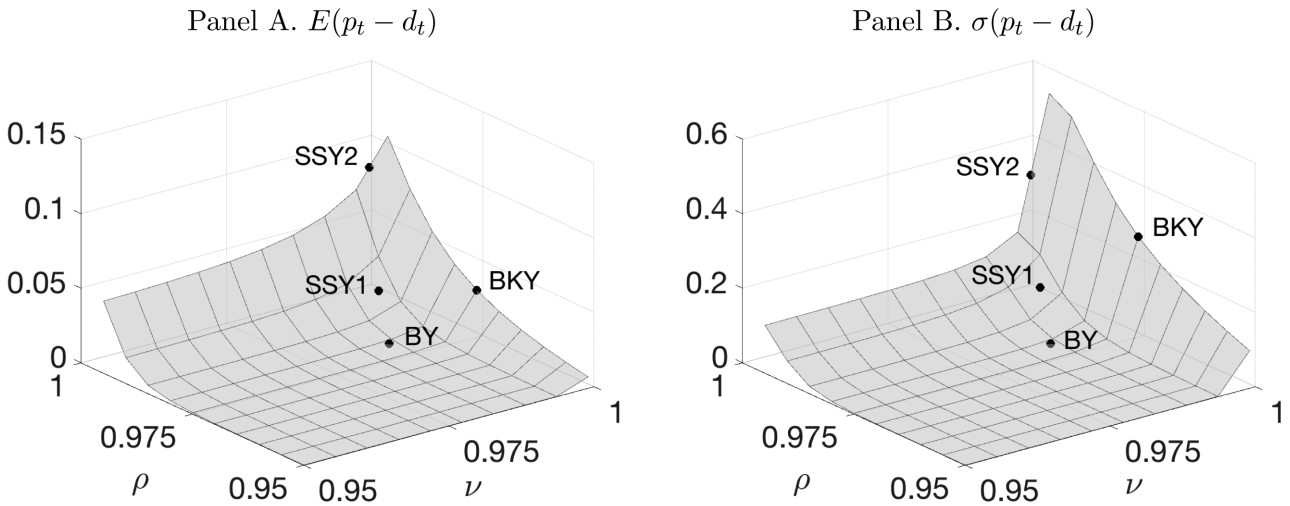

Figure 3. Relative errors in the monthly log price-dividend ratio. The figure plots the numerical errors in the monthly log price-dividend ratio introduced by the log-linearization. The left panel shows the relative errors in the unconditional mean of the log price-dividend ratio for different values of the persistence of long-run risk ρ and that of stochastic volatility ν. The right panel shows the corresponding relative errors in the unconditional standard deviation of the log price-dividend ratio. Parameters: $\delta = 0.9989$, $\gamma = 10$, $\psi = 1.5$, $\mu = 0.0015$, $\phi_c = 1$, $\phi_x = 0.038$, $\bar{\sigma} = 0.0072$, $\phi_\sigma = 2.8\text{e-}6$, $\mu_d = 0.0015$, $\Phi = 2.5$, $\phi_d = 5.96$, and $\phi_{d,c} = 2.6$. The figure also plots the errors for the persistence parameters used in the studies of Bansal and Yaron (2004) (BY), Bansal, Kiku, and Yaron (2012) (BKY), and Schorfheide, Song, and Yaron (2018) for their median parameter estimates (SSY1) and their 95% estimates (SSY2). (See Table IV for all the parameters used in the studies.)

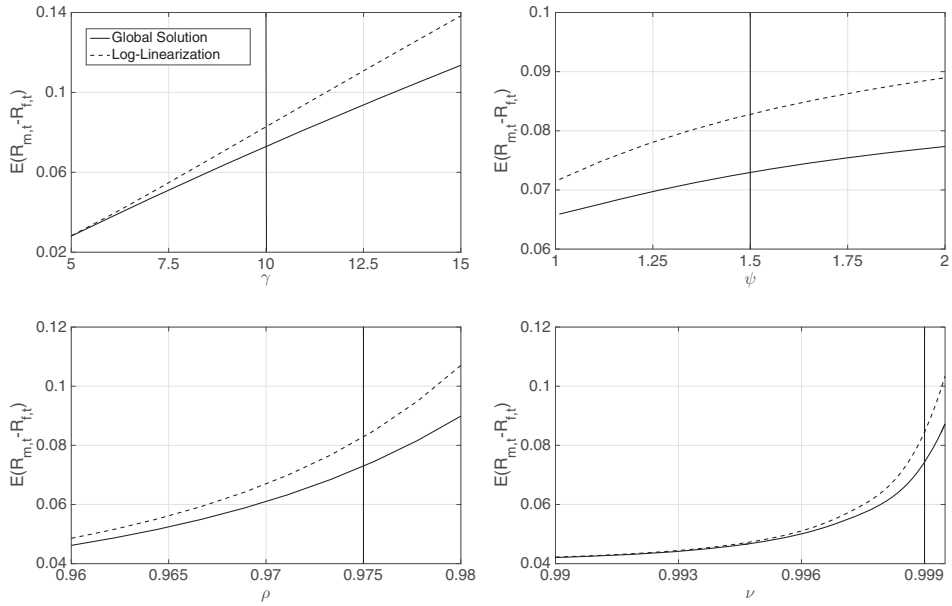

Figure 4. Sensitivity of the approximation errors for the annualized equity premium in the long-run risk model. The figure plots the annual equity premium obtained by the log-linearization (dashed line) as well as the global solution (solid line) as a function of the model parameters γ, ψ, ρ, and ν, assuming that the other parameters are kept constant. In each panel, the vertical line denotes the estimate used in the calibration of Bansal, Kiku, and Yaron (2012). Parameters: $\delta = 0.9989$, $\gamma = 10$, $\psi = 1.5$, $\mu = 0.0015$, $\phi_c = 1$, $\phi_x = 0.038$, $\rho = 0.975$, $\bar{\sigma} = 0.0072$, $\nu = 0.999$, $\phi_\sigma = 2.8\text{e-}6$, $\mu_d = 0.0015$, $\Phi = 2.5$, $\phi_d = 5.96$, and $\phi_{d,c} = 2.6$.

for both long-run risk and stochastic volatility. We find that even very small changes can dramatically increase approximation errors. For example, in the original calibration with a persistence in the long-run risk of $\rho = 0.975$, the overestimation of the equity premium is about 100 bps (vertical black line). By slightly increasing ρ to 0.98, however, the difference almost doubles, with an overestimation of 180 bps. For the persistence in the stochastic volatility, ν, even a change of 0.0005 (from 0.999 to 0.9995) increases the overestimation to 170 bps. The figures also show that lowering the persistence parameters significantly decreases approximation errors. For example, for $\nu = 0.99$ the approximation error is close to zero. Again, however, for this calibration the implied model moments collapse. Given the model requires highly persistent state processes, a large degree of risk aversion, and an EIS exceeding one, it follows that nonlinear dynamics in the model are present and strong. This model feature renders the use of log-linearization a problematic method for solving the model since it implies large approximation errors and even tiny changes in the model parameters can strongly affect these errors.

REMARK: While most readers will agree with our qualitative insights concerning the presence of numerical errors in the log-linearized solution, some researchers in the asset pricing literature may disagree with our describing

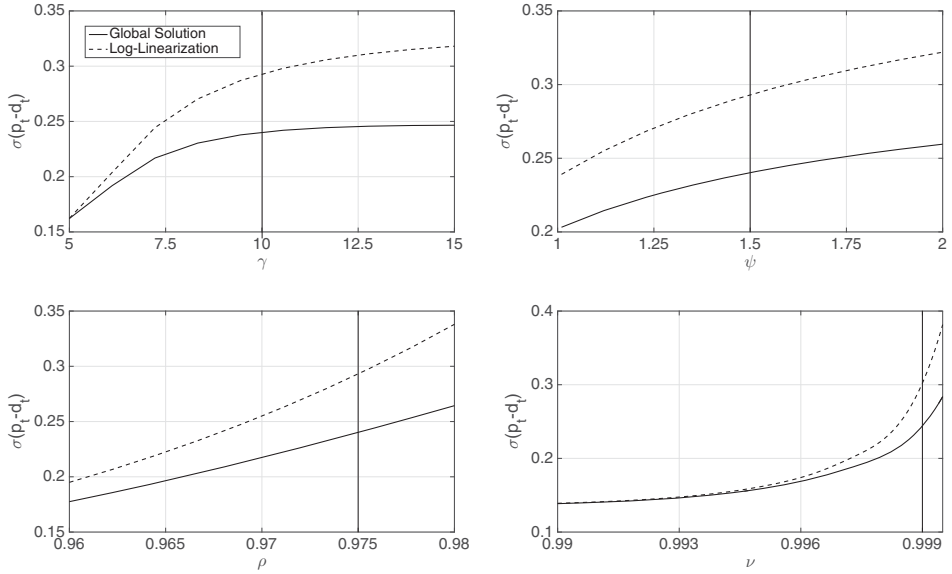

Figure 5. Sensitivity of the approximation errors for the annual volatility of the price-dividend ratio in the long-run risk model. The figure plots the annual volatility of the log price-dividend ratio obtained by the log-linearization (dashed line) as well as the global solution (solid line) as a function of the model parameters γ, ψ, ρ, and ν, assuming that the other parameters are kept constant. In each panel, the vertical line denotes the estimate used in the calibration of Bansal, Kiku, and Yaron (2012). Parameters: $\delta = 0.9989$, $\gamma = 10$, $\psi = 1.5$, $\mu = 0.0015$, $\phi_c = 1$, $\phi_x = 0.038$, $\rho = 0.975$, $\bar{\sigma} = 0.0072$, $\nu = 0.999$, $\phi_\sigma = 2.8\text{e-}6$, $\mu_d = 0.0015$, $\Phi = 2.5$, $\phi_d = 5.96$, and $\phi_{d,c} = 2.6$.

several numerical errors as "large." We therefore briefly discuss the quantitative relevance of our results.

We consider an overestimation of the annualized equity premium of 100 bps (see Figure 4) to be a serious error deserving our attention. However, Bansal and Yaron (2004, Table IV) report a point estimate of 6.33% with a standard error of 2.15% for the equity premium in U.S. market data. In relation to these data, the approximation error is less than one-half of the standard error, and thus, statistically speaking, the difference between the log-linearized solution and the projection solution is rather small. Obviously, a reader with this purely statistical viewpoint will disregard our analysis. (Of course, the question then arises as to how many asset pricing "puzzles" survive such a statistical standard, when a model equity premium of 2% is considered sufficient because it lies in the 95% confidence interval.)

We recognize this criticism of our interpretation of the results, but we also propose two counterarguments. First, judging from a good number of papers in the asset pricing literature, many authors appear to interpret the idea of "matching the data" as not only a qualitative but also as a quantitative objective. From this perspective, a numerical error of 100 bps relative to a point estimate of 6.33% is economically relevant and deserves our attention. Second,

Table I

Nonlinearities with and without Interaction of Stochastic Volatility and Long-Run Risk

The table reports the annualized mean and standard deviation of the log wealth-consumption ratio, the log price-dividend ratio, the equity premium, as well as the relative errors of the log-linear solution. Results are shown for the standard long-run risk model (specification (7)) as well as the case without stochastic volatility in x_t (equation (7) with equation (8) for the long-run risk process). Parameters: $\delta = 0.9989$, $\gamma = 10$, $\psi = 1.5$, $\mu = 0.0015$, $\phi_c = 1$, $\phi_x = 0.038$, $\rho = 0.975$, $\bar{\sigma} = 0.0072$, $\nu = 0.999$, $\phi_\sigma = 2.8\text{e}-6$, $\mu_d = 0.0015$, $\Phi = 2.5$, $\phi_d = 5.96$, and $\phi_{d,c} = 2.6$.

	$\mathrm{E}(w_t - c_t)$	$\sigma(w_t - c_t)$	$\mathrm{E}(p_t - d_t)$	$\sigma(p_t - d_t)$	$\mathrm{E}(R_{m,t} - R_{f,t})$
Panel A: With Stochastic Volatility in x_t					
Log-Linear	4.14	0.10	3.04	0.29	8.28
Global	4.21	0.09	3.24	0.24	7.30
Error	1.74%	12.10%	6.04%	22.02%	13.49%
Panel B: With Constant Volatility in x_t					
Log-Linear	4.51	0.04	3.94	0.15	4.48
Global	4.51	0.04	3.94	0.15	4.48
Error	0.11%	2.16%	0.03%	0.11%	0.10%

we solve an economic model not only to compare forecasts with data, but also to understand the economic mechanisms within the model. To gain such an understanding, we want to know the exact forecasts of the model and how these react to parameter changes. From this model analysis perspective, solutions of high accuracy are desirable even in the presence of large standard errors in the data.

B. Sources of Numerical Errors

In this section, we analyze in more detail the source of nonlinearity in the model solution. We found in the previous section that the two persistence parameters ρ and ν play a central role in the model: the larger their values, the larger the errors in the log-linear approximation become. However, as we show below, it is the interaction between long-run risk and stochastic volatility that makes the errors particularly severe. Put differently, the possibility of a large shock to volatility also increasing the probability of a large (negative) shock to the consumption growth rate affects both the model solution and the numerical error of the log-linear approximation.

Consider a model setup in which we turn off the influence of stochastic volatility on long-run risk. That is, we set the volatility of x_t from the cash flow equations (7) to its constant long-run mean, $\bar{\sigma}$, which results in the specification

$$x_{t+1} = \rho x_t + \phi_x \bar{\sigma} \eta_{x,t+1}. \tag{8}$$

Note that the stochastic volatility σ_t still influences log consumption and log dividend growth. Table I reports the annualized mean and standard deviation

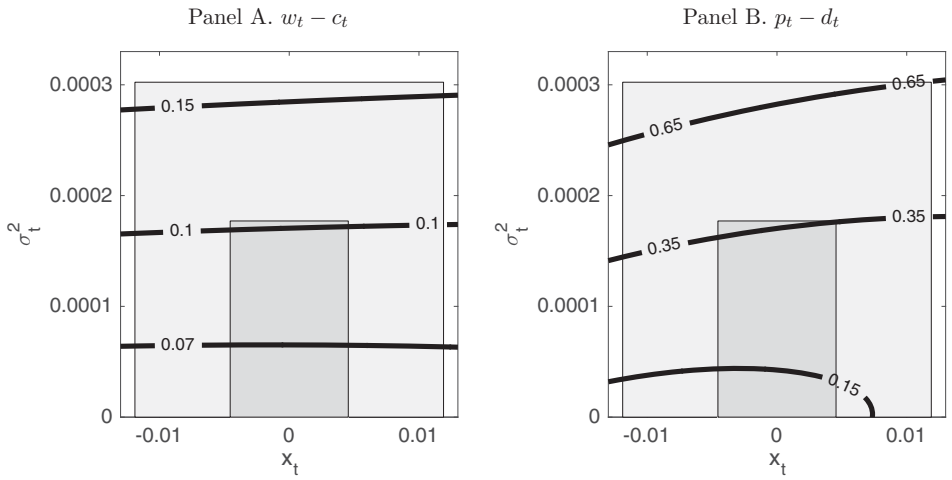

Figure 6. Absolute approximation errors in the log wealth-consumption and the log price-dividend ratio in the long-run risk model with stochastic volatility in x_t. The figure plots isolines (black solid lines) for the absolute errors in the log wealth-consumption ratio (left panel) and the log price-dividend ratio (right panel) of the log-linearization as a function of the states x_t and σ_t^2. The dark gray area shows the range of two unconditional standard deviations of the state processes. For x_t the area is computed using a given shock of two unconditional standard deviations in σ_t^2. The light gray area shows the corresponding range for four unconditional standard deviations. Results are shown for the standard long-run risk model with stochastic volatility in x_t (specification (7)). Parameters: $\delta = 0.9989$, $\gamma = 10$, $\psi = 1.5$, $\mu = 0.0015$, $\phi_c = 1$, $\phi_x = 0.038$, $\rho = 0.975$, $\bar{\sigma} = 0.0072$, $\nu = 0.999$, $\phi_\sigma = 2.8\text{e-}6$, $\mu_d = 0.0015$, $\Phi = 2.5$, $\phi_d = 5.96$, and $\phi_{d,c} = 2.6$.

of the log wealth-consumption ratio, the log price-dividend ratio, the equity premium, as well as the relative errors of the log-linear solution.

In line with our previous results, the errors for the standard long-run risk model are large and significant, with the largest error in the volatility of the log price-dividend ratio of 22%. For the case with no stochastic volatility in x_t, the errors decrease to almost zero and the log-linear approximation provides an accurate solution. On the other hand, the equity premium also decreases from 7.3% to 4.48%. These results show that the interaction between long-run consumption risk and stochastic volatility is a crucial risk channel in the long-run risk model. However, the standard Campbell-Shiller (1988) log-linear approximation misses any interaction effects between x_t and σ_t and only captures the linear dependencies.

Why does the interaction between the two risk channels matter so much? In Figure 6, we plot isolines (black solid lines) for the absolute errors in the log wealth-consumption ratio (left panel) and the log price-dividend ratio (right panel) of the log-linearization as a function of the states x_t and σ_t^2 for the model with stochastic volatility in x_t. For example, along the line marked with "0.1," the absolute error of the log-linearization is 0.1. The figures also depict the regions that the model actually visits and the regions in which it "spends most of its time" over long simulations. We display the ranges of two and four

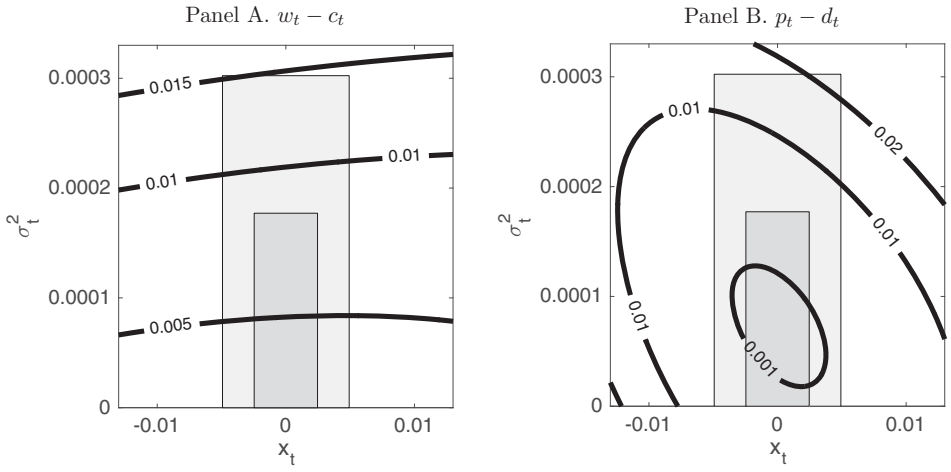

Figure 7. Absolute approximation errors in the log wealth-consumption ratio and the log price-dividend ratio in the long-run risk model with constant volatility in x_t. The figure plots isolines (black solid lines) for the absolute errors in the log wealth-consumption ratio (left panel) and the log price-dividend ratio (right panel) of the log-linearization as a function of the states x_t and σ_t^2. The dark gray area shows the range of two unconditional standard deviations of the state processes. For x_t the area is computed using a given shock of two unconditional standard deviations in σ_t^2. The light gray area shows the corresponding range for four unconditional standard deviations. Results are shown for the long-run risk model without stochastic volatility in x_t (specification (7) with equation (8) for the long-run risk process). Parameters: $\delta = 0.9989$, $\gamma = 10$, $\psi = 1.5$, $\mu = 0.0015$, $\phi_c = 1$, $\phi_x = 0.038$, $\rho = 0.975$, $\bar{\sigma} = 0.0072$, $\nu = 0.999$, $\phi_\sigma = 2.8\text{e-}6$, $\mu_d = 0.0015$, $\Phi = 2.5$, $\phi_d = 5.96$, and $\phi_{d,c} = 2.6$.

unconditional standard deviations for the stochastic volatility state around its unconditional mean. For the long-run growth rate, we also show the ranges of two and four unconditional standard deviations for a given level of σ_t^2 for which we again use the values of two and four unconditional standard deviations. This region nicely demonstrates the interaction between the two risk sources as will become clear below.

We find that the absolute errors in the log wealth-consumption ratio are not too severe, with errors in the wealth-consumption ratio reaching about 16% ($e^{0.15} \approx 1.16$) for extreme values of the stochastic volatility process. The absolute errors in the log price-dividend ratio are considerably larger—they reach values of up to 0.35 in the two-standard-deviation and 0.65 in the four-standard-deviation range. In other words, the price-dividend ratio obtained by the log-linearization is off by a factor of $e^{0.35} \approx 1.42$ in the two-standard-deviation range and can be off by a factor of about two for extreme values of the state processes.[6]

Figure 7 plots the corresponding results for the case in which the volatility of x_t is constant. We find that the range of x_t decreases significantly, as the shocks

[6] In Figures E1 and E2 in Appendix E, we show the corresponding errors in the first and second derivatives of the log wealth-consumption and price-dividend ratio.

to volatility have no influence on x_t anymore. To better understand this effect, consider the unconditional distribution of x_{t+1} for a given level of σ_t^2. For the case in which the volatility of x_t is stochastic, we find that

$$x_{t+1}|\sigma_t^2 \sim N\left(0, \frac{\phi_x^2 \sigma_t^2}{(1-\rho^2)}\right).$$

This distribution implies that a negative two-standard-deviation shock to x_t, assuming that σ_t^2 is at its two-standard-deviation level, gives a value of $x_{t+1} = -0.0046$,[7] or a 5.52% reduction in log mean consumption growth annually. For the case in which the volatility of x_{t+1} is constant, and hence the level of σ_t^2 does not influence x_{t+1}, we find that

$$x_{t+1}|\sigma_t^2 \sim N\left(0, \frac{\phi_x^2 \bar{\sigma}^2}{(1-\rho^2)}\right).$$

This distribution implies that a negative two-standard-deviation shock to x_t, assuming that σ_t^2 is at its two-standard-deviation level (which has no influence), gives a value of $x_{t+1} = -0.0025$,[8] or a 2.96% reduction in mean consumption growth annually. So the reduction in mean consumption growth is almost twice as large in the case in which the volatility of x_t is stochastic. Hence, for the case in which the volatility of x_t is constant, the extreme regions of x_t do not matter as much anymore. We therefore find in Figure 7 that the errors in the wealth-consumption as well as the price-dividend ratio decrease significantly compared to the stochastic volatility case and are very close to zero. We conclude, therefore, that, once there is an interaction between the highly persistent risk sources, the nonlinearities in the model increase significantly. Since, by construction, the log-linearization does not include these interacting effects, the approximation errors are potentially large.

C. The Special Case of Unitary EIS

To compute approximate closed-form solutions for the long-run risk model, the Campbell-Shiller (1988) return linearization needs to be applied twice—first for the return on wealth to obtain a linear solution for the log wealth-consumption ratio, and second for the return on equity to obtain a linear solution for the log price-dividend ratio. Therefore, it may be helpful to understand which steps introduce the nonlinearities to the model. Perhaps the linearization can be applied to one of the return equations without introducing significant approximation errors. Or should we avoid linearization for both pricing equations? To shed light on this question, we analyze the special case of a unitary

[7] The value is computed by $-2\frac{\phi_x(\bar{\sigma}^2+2\frac{\phi_\sigma}{(1-\nu^2)^{0.5}})^{0.5}}{(1-\rho^2)^{0.5}}$.

[8] The value is computed by $-2\frac{\phi_x(\bar{\sigma}^2)^{0.5}}{(1-\rho^2)^{0.5}}$.

Table II

Euler Errors in the Long-Run Risk Model with Stochastic Volatility

The table reports the maximum absolute Euler error (MAE) as well as the root mean squared Euler error (RMSE) in the pricing equations of the wealth-consumption ratio and the price-dividend ratio in the long-run risk model (7). Euler errors are computed using 500 equally spaced nodes for each state. The approximation interval is taken from using the minimum and maximum values from 1,000,000 years of simulated data. Results are shown for $\psi = 1$ and $\psi = 1.5$. Parameters: $\delta = 0.9989$, $\gamma = 10$, $\mu = 0.0015$, $\phi_c = 1$, $\phi_x = 0.038$, $\rho = 0.975$, $\bar{\sigma} = 0.0072$, $\nu = 0.999$, $\phi_\sigma = 2.8\text{e-}6$, $\mu_d = 0.0015$, $\Phi = 2.5$, $\phi_d = 5.96$, and $\phi_{d,c} = 2.6$.

		Log-Linearization		Global Solution	
		$w_t - c_t$	$p_t - d_t$	$w_t - c_t$	$p_t - d_t$
$\psi = 1.5$	MAE	0.0102	0.0060	1.5e-9	4.1e-8
	RMSE	0.0027	0.0009	7.5e-10	7.4e-9
$\psi = 1$	MAE	–	0.0090	–	1.5e-6
	RMSE	–	0.0018	–	2.5e-7

EIS, $\psi = 1$. In this case, the log wealth-consumption ratio is constant,

$$w_t - c_t = \log\left(\frac{1}{1-\delta}\right) \tag{9}$$

(see Tallarini (2000)). The stochastic discount factor (3) for the long-run risk model can be derived in closed form and is given by

$$M_{t+1} = \delta\left(\frac{C_{t+1}}{C_t}\right)^{-\gamma} \frac{e^{(1-\gamma)vc_{t+1}}}{e^{\frac{1-\gamma}{\delta}vc_t}}, \tag{10}$$

where $vc_t = \log\frac{V_t}{C_t}$ is a linear function of the state variables given by $vc_t = a + bx_t + c\sigma_t^2$ (see Appendix C for the derivations of the stochastic discount factor and the closed-form expressions for a, b, and c). This result implies that the log stochastic discount factor is a linear function of the state variables. So, using the model specification with $\psi = 1$ implies that the only source of errors in the linearized model is the linearization of equity returns, which allows us to separately analyze this source of errors.

In Table II, we report Euler errors in the pricing equations for the return on wealth and the return on equity for the model of Bansal, Kiku, and Yaron (2012) with an EIS of $\psi = 1.5$ and an EIS of $\psi = 1$.

We find that there are significant Euler errors for the case with $\psi = 1.5$, with a maximum error of 0.0102 in the pricing equation for the return on wealth and 0.0060 in the pricing equation for the return on equity. For the case with $\psi = 1$, the linear solution for the return on wealth is exact and hence the Euler errors are zero. However, we still find that there are significant errors in the pricing equation for the return on equity, with a maximum Euler error of 0.0090. Table III reports the corresponding errors in the annualized mean and standard deviation of the log price-dividend ratio as well as in the equity premium for $\psi = 1$. If we compare the reported relative errors to the errors for the model

Table III

Nonlinearities in the Long-Run Risk Model with an EIS of $\psi = 1$

The table reports the annualized mean and standard deviation of the log price-dividend ratio, the equity premium, as well as the relative errors of the log-linear solution. Results are shown for the standard long-run risk model (equation (7)) with an EIS of $\psi = 1$. Parameters: $\delta = 0.998$, $\gamma = 10$, $\mu = 0.0015$, $\phi_c = 1$, $\phi_x = 0.044$, $\rho = 0.979$, $\bar{\sigma} = 0.0078$, $\nu = 0.987$, $\phi_\sigma = 2.3\text{e}-6$, $\mu_d = 0.0015$, $\Phi = 3$, $\phi_d = 4.5$, and $\phi_{d,c} = 0$.

	$\text{E}(p_t - d_t)$	$\sigma(p_t - d_t)$	$\text{E}(R_{m,t} - R_{f,t})$
Log-Linear	2.98	0.24	7.15
Global	3.10	0.20	6.58
Error	3.78%	17.55%	8.64%

with $\psi = 1.5$ (see Panel A in Table I), we observe that the relative errors are slightly smaller but still considerable. For example, the error in the volatility of the price-dividend ratio decreases from 22.02% for $\psi = 1.5$ to 17.55% for $\psi = 1$, and the error in the equity premium decreases from 13.49% to 8.64%. So, even if the pricing equation for wealth is truly linear as in the model with $\psi = 1$, linearizing the return on equity introduces significant approximation errors to the asset pricing model.

These findings suggest that nonlinearities come primarily from the pricing equation of the return on equity, in which case it might be reasonable to use a linear solution for the stochastic discount factor. To verify this claim in the general model with $\psi = 1.5$, we solve the model again using the linearized version for the return on wealth, but using the global method to price equities. If the approximation errors are of the same order of magnitude as for the case in which we use the global solution for the pricing of both wealth and equity, this would suggest that it might be reasonable to use log-linearized solutions to derive the stochastic discount factor. Indeed, we find that the equity premium is close to the value of the fully global solution with a value of 7.51% (compared to a value of 7.30% for the global solution; see Table I). This implies an error of 2.97%, which is significantly smaller than the error of the fully linearized model (the model has an error of 13.49%). This result is not true in general, however, but rather depends highly on the model parameterization. In the next section, we present results for different model specifications that also show significant errors in the return on wealth for the linearized solutions. We therefore conclude that the Campbell-Shiller (1988) return linearization should not be applied to any return equations without carefully testing its accuracy for the specific asset pricing model under consideration.

III. Higher Order Dynamics in Six Asset Pricing Models

In this section, we compare the implications of using log-linearized solutions for the prediction of economically relevant quantities in asset pricing models. Specifically, we perform this comparison for six different studies from the

recent asset pricing literature on long-run risk. The six models are the original long-run risk model of Bansal and Yaron (2004), the recalibrated version of the model by Bansal, Kiku, and Yaron (2012), the extensive estimation study of Schorfheide, Song, and Yaron (2018), the volatility-of-volatility model of Bollerslev, Xu, and Zhou (2015), and the two studies of real and nominal bonds of Koijen et al. (2010) and Bansal and Shaliastovich (2013). Common to all of these studies is the methodological attempt to match several key statistics on financial markets such as the high equity premium, a low risk-free rate, volatile stock prices, real and nominal bond prices, the volatility premium, or patterns in return predictability. Obviously, to obtain a reasonable calibration of the model, it is essential to solve the model without significant errors in the approximation of these key statistics since such errors could potentially bias the calibration or estimation.

In the previous section, we see that depending on the persistence of the state processes, the log-linearization approach produces sizable approximation errors in the calibration of the long-run risk model of Bansal, Kiku, and Yaron (2012). We now demonstrate that using log-linearized solutions has a strong impact on the predictions of these six asset pricing models.

A. *Six Model Specifications*

The six studies share the same basic model setup for log consumption and dividend growth as the model of Bansal and Yaron (2004):

$$\begin{aligned} \Delta c_{t+1} &= \mu_c + x_t + \phi_c \sigma_{c,t} \eta_{c,t+1}, \\ x_{t+1} &= \rho x_t + \phi_x \sigma_{x,t} \eta_{x,t+1}, \\ \Delta d_{t+1} &= \mu_d + \Phi x_t + \phi_d \sigma_{d,t} \eta_{d,t+1} + \phi_{d,c} \sigma_{c,t} \eta_{c,t+1} \\ &\quad \times \eta_{c,t+1}, \eta_{x,t+1}, \eta_{d,t+1} \sim \text{i.i.d. } N(0,1). \end{aligned} \tag{11}$$

In the following, we describe how the models differ in regard to the detailed specifications.

B. *The Studies of Bansal and Yaron (2004) and Bansal, Kiku, and Yaron (2012)*

Bansal and Yaron (2004) and Bansal, Kiku, and Yaron (2012) assume that a single volatility process drives uncertainty in the economy, $\sigma_{c,t} = \sigma_{x,t} = \sigma_{d,t} = \sigma_t$, with

$$\sigma_{t+1}^2 = \bar{\sigma}^2(1 - \nu_c) + \nu_c \sigma_t^2 + \phi_\sigma \omega_{t+1}, \qquad \omega_{t+1} \sim \text{i.i.d. } N(0,1). \tag{12}$$

Recall that this is the model setup (with $\nu = \nu_c$) that we employ above to examine the approximation errors of the log-linearized solution (see also equation (7)).

C. *The Estimation Study of Schorfheide, Song, and Yaron (2018)*

Schorfheide, Song, and Yaron (2018) relax the assumption of a single volatility process, allowing for three separate volatility processes for consumption,

dividends, and long-run risk.[9] The volatility processes for consumption growth and the long-run risk factor are required to account for the weak correlation between the risk-free rate and consumption growth. As shown in their estimation study, the volatility dynamics of dividends differ significantly from the other two processes, and thus a third process is required to model the stochastic volatility of dividends. Schorfheide, Song, and Yaron (2018) assume that the logarithm of each volatility process is normal to ensure that the standard deviation of the shocks remains positive,

$$\begin{aligned} \sigma_{i,t} &= \varphi_i \bar{\sigma} \exp(h_{i,t}), \\ h_{i,t+1} &= \nu_i h_{i,t} + \sigma_{h_i} \sqrt{1-\nu_i^2}\, \omega_{i,t+1}, \qquad i \in \{c, x, d\} \\ &\quad \times \omega_{i,t+1} \sim \text{i.i.d. } N(0,1). \end{aligned} \tag{13}$$

To derive analytical solutions for the log-linearization coefficients needed in their estimation study, Schorfheide, Song, and Yaron (2018) use a linear approximation of the volatility dynamics that follows Gaussian dynamics,

$$\sigma_{i,t}^2 \approx 2(\varphi_i \bar{\sigma})^2 h_{i,t} + (\varphi_i \bar{\sigma})^2, \tag{14}$$

which in turn yields

$$\sigma_{i,t+1}^2 = \bar{\sigma}_i^2 (1-\nu_i) + \nu_i \sigma_{i,t}^2 + \phi_{\sigma_i} \omega_{i,t+1}$$

with $\phi_{\sigma_i} = 2\bar{\sigma}_i^2 \sigma_{h_i} \sqrt{1-\nu_i^2}$ and $\bar{\sigma}_i = \varphi_i \bar{\sigma}$. To obtain consistent results with the original study, we proceed in the same way as Schorfheide, Song, and Yaron (2018) and solve the model using the linearized version of the volatility dynamics for both the linearized as well as the global solution obtained by the projection approach.

D. The Estimation Study of Bollerslev, Xu, and Zhou (2015)

The fourth model stems from the estimation study of Bollerslev, Xu, and Zhou (2015). In a standard long-run risk model with stochastic volatility, many long-standing puzzling behaviors of financial markets such as a high equity risk premium together with a low risk-free rate, volatile price dynamics, and the predictability of stock returns can be explained. However, recent research goes a step further by showing that the standard model is not able to generate a time-varying variance risk premium that has predictive power for stock returns. Fortunately, the literature has suggested a possible solution to this puzzle by adding time-varying volatility of volatility (vol-of-vol) to the model (see, e.g., Bollerslev, Tauchen, and Zhou (2009), Drechsler and Yaron (2011), Tauchen (2011), Bollerslev, Xu, and Zhou (2015), and Dew-Becker et al. (2017)).

[9] Schorfheide, Song, and Yaron (2018) also introduce a shock to the time rate of preferences. Since in this study we are interested in the influence of higher order effects introduced by the highly persistent state processes, we omit the preference shock. For this purpose we set $\rho_\lambda = \sigma_\lambda = 0$ in their model specification.

Bollerslev, Xu, and Zhou (2015) consider a slight variation of the long-run risk factor compared to the baseline model (7) in which the vol-of-vol factor q_t drives the volatility,[10]

$$\begin{aligned} \sigma_{t+1}^2 &= \bar{\sigma}^2(1-\nu) + \nu\sigma_t^2 + \phi_\sigma\sqrt{q_t}\omega_{\sigma,t+1}, \\ q_{t+1} &= \mu_q(1-\rho_q) + \rho_q q_t + \phi_q\sqrt{q_t}\omega_{q,t+1}, \\ x_{t+1} &= \rho x_t + \phi_x\sqrt{q_t}\eta_{x,t+1} \\ &\quad \times \eta_{x,t+1}, \omega_{\sigma,t+1}, \omega_{q,t+1} \sim \text{i.i.d. } N(0,1). \end{aligned} \tag{15}$$

The vol-of-vol factor q_t follows a square root process. This process specification has also been used, for example, in Tauchen (2011) and in the seminal work on vol-of-vol in this model class by Bollerslev, Tauchen, and Zhou (2009). However, a square root process poses a new challenge to the model, as the process can become complex-valued when q_t is negative. This problem is usually circumvented either by assuming a reflecting boundary at zero or by truncation to ensure positivity. However, for a simple computation of model solutions, the assumption of a nontruncated distribution for the log-linearization is commonly used. (For example, Bansal and Yaron (2004) use the nontruncated distribution to compute the log-linearized solutions but replace negative realizations in the simulations of the stochastic volatility process with small positive numbers. This approach has been used by many subsequent papers in the long-run risk literature.) In Appendix D, we analyze in more detail how the square root process specification and the issue of complexity affects the log-linearized solution. We find that, for the calibration in Bollerslev, Tauchen, and Zhou (2009), equilibrium model solutions are not real numbers but rather are complex numbers. For the parameters in Bollerslev, Xu, and Zhou (2015), the process is centered well above zero and the standard log-linearization technique yields a real solution. We therefore concentrate on this calibration in the main text.

E. The Model of Koijen et al. (2010)

The fifth study under consideration is the study of real and nominal bonds and the size of the martingale component in the stochastic discount factor by Koijen et al. (2010). They add inflation, π_t, with a stochastic growth rate, $x_{\pi,t}$, to the standard model (11) and price nominal bonds:[11]

[10] Drechsler and Yaron (2011) use a similar model where the volatility of x_t is driven by σ_t instead of q_t; see their 2007 working paper version. However, Bollerslev, Xu, and Zhou (2015) provide evidence of a better empirical match for their model specification. The estimation study of Bollerslev, Xu, and Zhou (2015) also models cross-correlations between the shocks of the state processes. For the analysis of the nonlinear dynamics of the model, we keep the model as parsimonious as possible and drop the cross-correlations.

[11] The model setup is the same as in the 2008 version of Bansal and Shaliastovich (2013). In the paper, they write $\bar{\pi}_t$ for $x_{\pi,t}$.

$$\begin{aligned}\pi_{t+1} &= \mu_\pi + x_{\pi,t} + \phi_{\pi,c}\sigma_{c,t}\eta_{c,t+1} + \phi_{\pi,x}\sigma_{x,t}\eta_{x,t+1} + \sigma_\pi\eta_{\pi,t+1},\\ x_{\pi,t+1} &= \mu_{x_\pi}(1-\rho_\pi) + \rho_\pi x_{\pi,t} + \rho_{\pi,x}x_t\\ &\quad + \phi_{x_\pi,c}\sigma_{c,t}\eta_{c,t+1} + \phi_{x_\pi,x}\sigma_{x,t}\eta_{x,t+1} + \sigma_{x_\pi}\eta_{\pi,t+1}\\ &\quad \times\eta_{\pi,t+1} \sim \text{i.i.d. } N(0,1).\end{aligned} \tag{16}$$

Koijen et al. (2010) assume that there are two stochastic volatility processes for consumption growth and the long-run risk component ($\sigma_{d,t} = \sigma_{c,t}$),

$$\sigma^2_{i,t+1} = \bar{\sigma}^2_i(1-\nu_i) + \nu_i\sigma^2_{i,t} + \phi_{\sigma_i}\omega_{i,t+1}, \qquad i \in \{c, x\},$$

and that inflation, the stochastic growth rate of inflation, and dividends have loadings on these two volatility channels.

F. The Model of Bansal and Shaliastovich (2013)

The sixth and last study under consideration is the work on nominal and real bonds of Bansal and Shaliastovich (2013). The setup is very similar to Koijen et al. (2010), but they assume that $x_{\pi,t}$ enters the real stochastic growth rate of consumption, x_t, to model the nonneutral effect of expected inflation on future expected growth,

$$\begin{aligned}\pi_{t+1} &= \mu_\pi + x_{\pi,t} + \sigma_\pi\eta_{\pi,t+1},\\ x_{\pi,t+1} &= \rho_\pi x_{\pi,t} + \sigma_{\pi,t}e_{\pi,t+1},\\ x_{t+1} &= \rho x_t + \rho_{x\pi}x_{\pi,t} + \sigma^x_t e_{x,t+1}\\ &\quad \times\eta_{\pi,t+1}, e_{\pi,t+1}, e_{x,t+1} \sim \text{i.i.d. } N(0,1).\end{aligned} \tag{17}$$

Also, they assume that there is a separate AR(1) process for the volatility of the stochastic growth rate of inflation, $\sigma_{\pi,t}$, and that the volatility of consumption growth is constant ($\sigma_{c,t} = \bar{\sigma}_c$). The process for $\sigma^2_{i,t+1}$ is

$$\sigma^2_{i,t+1} = \bar{\sigma}^2_i(1-\nu_i) + \nu_i\sigma^2_{i,t} + \phi_{\sigma_i}\omega_{i,t+1}, \qquad i \in \{x, \pi\}.$$

As the focus of Bansal and Shaliastovich (2013) is the bond market, they do not include a process for dividends.

Table IV reports all of the parameter values used in the six studies.[12]

While the parameters in Bansal and Yaron (2004) and Bansal, Kiku, and Yaron (2012) are calibrated, Schorfheide, Song, and Yaron (2018), Bollerslev, Xu, and Zhou (2015), and Bansal and Shaliastovich (2013) estimate the model parameters to match annual financial market characteristics. In the first five models, the investor has a monthly decision interval, while Bansal and Shaliastovich (2013) use quarterly intervals. This distinction explains, for example, the considerable difference in the level parameters. The main difference between the sets of parameters in the original Bansal and Yaron (2004) calibration and the updated calibration of Bansal, Kiku, and Yaron (2012) is that in the latter

[12] For the model of Bollerslev, Xu, and Zhou (2015), we use the parameter estimates in the study for ρ, ν, and ρ_q. As they do not report values for the remaining parameters, we use the calibration as reported in the 2007 working paper version of Drechsler and Yaron (2011).

Table IV

Model Parameters

The table reports parameter values for the studies of Bansal and Yaron (2004) (from Table 4 on p. 1495), Bansal, Kiku, and Yaron (2012) (from Table 1 on p. 193), Schorfheide, Song, and Yaron (2018) (the median estimates from Table 5 on p. 25 are shown, as well as the 95% estimates for the persistence parameters in parentheses), Bansal and Shaliastovich (2013) (from Table 4 on p. 22), Koijen et al. (2010) (from Table 1 on p. 19 in the online appendix), and Bollerslev, Xu, and Zhou (2015) (from Section 3.3 starting on p. 464).

		BY (2004)	BKY (2012)	SSY (2016)	BS (2013)	KLVV (2010)	BXZ (2015)
Pref.	γ	10	10	8.598	20.90	8	10
	ψ	1.5	1.5	1.935	1.81	1.5	1.5
	δ	0.998	0.9989	0.999	0.994	0.9987	0.999
Consumption	μ_c	0.0015	0.0015	0.0016	0.0049	0.0016	0.0015
	ϕ_c	1	1	1	1	1	0.00546
	ρ	0.979	0.975	0.9872 (0.9995)	0.81	0.991	0.988
	ϕ_x	0.044	0.038	1	1	1	3.12e–4
	$\rho_{x\pi}$	–	–	–	–0.047	0	–
Volatility	ν_c	0.987	0.999	0.9914 (0.9958)	0	0.85	0.64
	ν_x	–	–	0.9943 (0.9988)	0.994	0.996	–
	ν_d	–	–	0.9665 (0.9841)	–	–	–
	ν_π	–	–	–	0.979	–	–
	ϕ_{σ_c}	2.3e–6	2.8e–6	1.9e–6	0	1.15e–6	1
	ϕ_{σ_x}	–	–	6.9e–11	1.85e–7	4.19e–9	–
	ϕ_{σ_d}	–	–	1.0e–4	–	–	–
	ϕ_{σ_π}	–	–	–	1.81e–7	–	–
	$\bar{\sigma}_c$	0.0078	0.0072	0.0032	4.6e–3	0.004	1
	$\bar{\sigma}_x$	–	–	2.34e–5	1.09e–3	1.60e–5	–
	$\bar{\sigma}_d$	–	–	0.0161	–	–	–
	$\bar{\sigma}_\pi$	–	–	–	1.11e–3	–	–
Dividends	μ_d	0.0015	0.0015	0.001	–	0.0015	0.0015
	Φ	3.0	2.5	4.147	–	1.5	3.0
	ϕ_d	4.5	5.96	1	–	6	0.0246
	$\phi_{d,c}$	0	2.6	1.544	–	0.6	0
Inflation	μ_π	–	–	–	0.0090	0	–
	μ_{x_π}	–	–	–	0	0.0032	–
	σ_π	–	–	–	0.0055	0.0035	–
	σ_{x_π}	–	–	–	0	4e–6	–
	$\phi_{\pi,c}$	–	–	–	0	0	–
	$\phi_{\pi,x}$	–	–	–	0	–2	–
	$\phi_{x_\pi,c}$	–	–	–	0	0	–
	$\phi_{x_\pi,x}$	–	–	–	0	–1	–
	ρ_π	–	–	–	0.988	0.83	–
	$\rho_{\pi,x}$	–	–	–	0	–0.35	–
V. of V.	μ_q	–	–	–	–		0.211
	ϕ_q	–	–	–	–		0.632
	ρ_q	–	–	–	–		0.46

calibration the persistence of the volatility shock, ν_c, is higher and shocks to dividends are correlated with short-run shocks to consumption growth ($\phi_{d,c} = 2.6$ in the updated calibration compared to $\phi_{d,c} = 0$ in the original calibration). These changes increase the influence of the volatility channel compared to the long-run risk channel of the model. The adjustment is needed to get rid of some implications of the original calibration that are inconsistent with the data. In particular, as, for example, Zhou and Zhu (2015) or Beeler and Campbell (2012) point out with respect to the original 2004 calibration, the log price-dividend ratio has predictive power for future consumption growth, while this relationship is not present in the data. By increasing the influence of the volatility channel, this predictability vanishes.

The extensive estimation study of Schorfheide, Song, and Yaron (2018) provides further evidence of highly persistent state processes and hence potentially large nonlinear dynamics. In particular, they report a median estimate for the persistence of long-run risk of 0.9872, and median values for the persistence parameters of the stochastic volatility of consumption, long-run risk, and dividends of 0.9914, 0.9943, and 0.9665, respectively. These four numbers are the median estimates from Bayesian estimation. For the 95% estimates, the authors report values as large as 0.9995, 0.9958, 0.9988, and 0.9841 (values provided in parentheses in Table IV). As equilibrium outcomes for those parameters are evaluated within the estimation procedure, we also provide the results for the 95% quantile estimates of the persistence parameters. Unfortunately, we were not able to compute results for the full set of 95% quantile parameters using also the higher estimates for the other cash flow and preference parameters. For extreme model parameters, it can be the case that there exists no solution for the asset pricing model (see Pohl, Schmedders, and Wilms (2015)). More surprisingly, in those cases the log-linearized solution may still deliver a well-behaved, though apparently nonsensical, solution due to its systematic underestimation of the price-dividend ratio (see Section A). Hence, the existence of solutions for the full range of parameters used in Schorfheide, Song, and Yaron (2018) is not necessarily satisfied. Since in the present study we are interested in the influence of higher order dynamics instead of the existence of solutions, we focus on the parameter range over which the model solutions are still well behaved. This is a rather conservative approach, as approximation errors increase the more extreme the values in the calibration (recall Figures 4 and 5).

G. Moments and Errors

Table V reports annualized summary statistics and numerical errors for the five models that include a dividend process. The reported financial statistics are the mean and standard deviation of the price-dividend ratio, the averages of the market excess return and the risk-free rate, and the volatilities of the excess return and the risk-free rate.[13] The table reports these statistics for

[13] We solve the model for the return of the wealth portfolio, z_w, the market portfolio, z_m, and the risk-free rate, z_{rf}. To compute the annualized moments, we simulate 1,000,000 years of artificial

Table V

Annualized Moments and Errors

The table reports the mean and the standard deviation of the annualized log price-dividend ratio, the annualized market over the risk-free return, and the risk-free rate. Results obtained by the log-linearization and the global solution as well as the relative error of the log-linearization are shown for the models of Bansal and Yaron (2004), Bansal, Kiku, and Yaron (2012), Schorfheide, Song, and Yaron (2018) (set (1) shows the results for the median parameter estimates reported and set (2) shows the results for the 95% estimates for the persistence of the state processes), Bollerslev, Xu, and Zhou (2015), and Koijen et al. (2010) (see Table IV). All returns and volatilities are shown in percent, so a value of 1.5 is a 1.5% annualized figure.

	$\mathrm{E}(p_t - d_t)$	$\sigma(p_t - d_t)$	$\mathrm{E}(R_{m,t} - R_{f,t})$	$\mathrm{E}(R_{f,t})$	$\sigma(R_{m,t})$	$\sigma(R_{f,t})$
			Bansal and Yaron (2004)			
Log-Lin	3.0105	0.1969	5.88	2.63	18.45	1.35
Global	3.0379	0.1946	5.79	2.63	18.35	1.34
Error	0.90 %	1.17%	1.67%	0.06%	0.54%	0.08%
			Bansal, Kiku, and Yaron (2012)			
Log-Lin	3.0414	0.2931	8.28	0.99	24.35	1.30
Global	3.2370	0.2402	7.30	1.11	23.76	1.28
Error	6.04%	22.02%	13.49%	10.34%	2.49%	1.32%
			Schorfheide, Song, and Yaron (2018) (1)			
Log-Lin	3.2853	0.2704	4.52	1.74	16.21	0.57
Global	3.3580	0.2557	4.27	1.75	15.59	0.57
Error	2.17%	5.78 %	5.87%	0.13%	4.01%	0.17%
			Schorfheide, Song, and Yaron (2018) (2)			
Log-Lin	2.5943	0.8748	10.08	−0.02	16.12	0.91
Global	3.3365	0.7841	5.93	0.92	14.29	0.54
Error	22.25%	11.57 %	70.07%	102.49%	12.76%	67.29%
			Bollerslev, Xu, and Zhou (2015)			
Log-Lin	2.7479	0.2737	8.88	1.18	18.29	1.56
Global	2.8222	0.2835	8.49	1.19	17.92	1.55
Error	2.63%	3.56%	4.52%	0.71%	2.09%	0.36%
			Koijen et al. (2010)			
Log-Lin	3.1137	0.1808	5.85	1.40	12.53	1.18
Global	3.3514	0.1479	4.37	1.72	11.33	1.15
Error	7.09%	22.31%	33.99%	18.77%	10.53%	2.27%

data. Beeler and Campbell (2012) provide a detailed description of how to compute the annual moments from monthly observations. A significant issue in the model is that the variance process σ_t^2 can, in fact, become negative. To overcome this problem, Bansal and Yaron (2004) replace all negative realizations with very small but positive values. We proceed in the same way for both methods to achieve consistent results. For the approximation interval of the projection methods, we set the interval to be slightly larger than the maximum observation range for the long simulations. As in the previous section, we increase the polynomial degree until the coefficients of the highest

the log-linearized solutions and the global solutions obtained by the projection approach, as well as the relative errors induced by the linearization.

We find that the log-linearization does a good job for the parameters in Bansal and Yaron (2004), with a maximal error of 1.67% for the equity premium. For the parameter set of Bansal, Kiku, and Yaron (2012), however, the results are considerably worse. The log-linearization overstates the equity premium by almost 100 bps. Further, it predicts a volatility of the log price-dividend ratio of 0.2931 instead of 0.2402, which corresponds to a relative error of about 22%. Simply put, the log-linearization produces a large equity premium and volatile log price-dividend ratio even though the true model predictions seem to be significantly smaller.

Beeler and Campbell (2012) argues that the predictability of excess returns in the original long-run risk calibration of Bansal and Yaron (2004) is too low compared to data, while at the same time it implies a counterfactually large predictability of consumption and dividend growth. In response to this criticism, Bansal, Kiku, and Yaron (2012) recalibrate the model to align it with the predictability evidence in the data.[14] Therefore, we next investigate the influence of numerical errors on the predictability regressions in Bansal, Kiku, and Yaron (2012). Table VI displays the R^2 statistics and regression coefficients from regressing cumulative log excess returns, consumption growth, and dividend growth on the lagged log price-dividend ratio. Statistics are shown for the annualized time series over horizons of one, three, and five years. (Table EI in Appendix E reports the corresponding results for the original calibration in Bansal and Yaron (2004).)

We find that substantial portions of the increased predictability of excess returns and the decreased predictability of dividend and consumption growth stem solely from the approximation errors of the log-linearized solution for the new calibration. For example, the log-linearization produces an R^2 of 0.74% for one-year excess returns and 3.38% for five-year excess returns. But the correct statistics are 0.39% and 1.83%, respectively. So the log-linearization overstates the predictability of returns in the model. For log consumption (dividend) growth the log-linearization produces values of 8.36% (10.47%) for one-year predictions. The correct values are significantly larger, with one-year predictability of 12.15% for consumption and 15.44% for dividends. Hence, using the global solutions amplifies the predictability concerns in the long-run risk model emphasized by Beeler and Campbell (2012).

For the model of Schorfheide, Song, and Yaron (2018), we find that approximation errors are in a reasonable range for the median parameter estimates (results (1) in Table V). However, using the 95% quantile estimates for the persistence parameters (results (2)), approximation errors increase dramatically,

order polynomial are close to zero. We double-check the accuracy of the solution by increasing the approximation interval until the solutions do not change.

[14] As the influence of the stochastic volatility channel compared to the long-run growth channel—which directly affects consumption growth—is increased in Bansal, Kiku, and Yaron (2012), the predictability of dividends and consumption is less pronounced in their calibration.

Table VI

Predictability of Excess Returns, Consumption, and Dividends in the Long-Run Risk Model of Bansal, Kiku, and Yaron (2012)

The table reports R^2 statistics and regression coefficients from regressing cumulative log excess returns, consumption growth, and dividend growth on the lagged log price-dividend ratio. Statistics are shown for the annualized time series with one-, three-, and five-year horizons for the model of Bansal, Kiku, and Yaron (2012). The table reports results for the log-linearized solution and the global solution as well as the relative error of the log-linearization.

	R^2			β		
	1Y	3Y	5Y	1Y	3Y	5Y
$\sum_h^H (r_{m,t+h} - r_{f,t+h}) = \alpha + \beta(p_t - d_t) + \epsilon_{t+H}$						
Log-Lin	0.0074	0.0212	0.0338	−0.0625	−0.1848	−0.3029
Global	0.0039	0.0113	0.0183	−0.0547	−0.1627	−0.2673
Error	88.62%	86.29%	84.85%	14.12%	13.57%	13.30%
$\sum_h^H (\Delta c_{t+h}) = \alpha + \beta(p_t - d_t) + \epsilon_{t+H}$						
Log-Lin	0.0836	0.0669	0.0531	0.0290	0.0589	0.0749
Global	0.1215	0.0967	0.0768	0.0427	0.0864	0.1099
Error	31.19%	30.87%	30.86%	32.02%	31.86%	31.85%
$\sum_h^H (\Delta d_{t+h}) = \alpha + \beta(p_t - d_t) + \epsilon_{t+H}$						
Log-Lin	0.1047	0.0496	0.0364	0.1808	0.2562	0.2959
Global	0.1544	0.0727	0.0532	0.2679	0.3785	0.4368
Error	32.19%	31.77%	31.70%	32.51%	32.30%	32.27%

with an overestimation of the equity premium of more than 70%. Table EII in Appendix E presents the corresponding results when only one of the 95% estimates of the persistence parameters is used for the persistence parameters ρ, ν_x, ν_c, and ν_d (instead of increasing all of the persistence parameters together). We find that it is sufficient to increase either the persistence in x_t, ρ, or the persistence of the volatility of x_t, ν_x, to obtain approximation errors larger than 37%. For the persistence of the volatility of consumption and dividend growth, ν_c and ν_d, the errors are significantly smaller. This result is in line with our findings from Section A, where we show that it is the combination of long-run risk and a persistent volatility that makes the model highly nonlinear and not the stochastic volatility in consumption or dividends itself. Schorfheide, Song, and Yaron (2018) use a Markov chain Monte Carlo approach to estimate model parameters and make Bayesian inference about model predictions. This requires that the approximation be accurate not only for the posterior median estimates, but also for other points of the parameter space at which the likelihood is evaluated. Otherwise, the posterior median estimates may also be biased. We find that using the 95% estimates for only one of the parameters is sufficient to introduce significant approximation errors. Figure E3 in Appendix E plots the errors in the equity premium as a function of the persistence of long-run growth ρ. An increase in ρ of even 1% from 0.9872 to

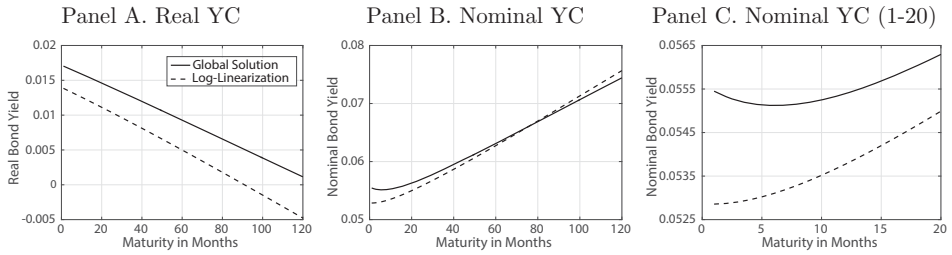

Figure 8. Real and nominal yield curve in the model of Koijen et al. (2010). The figure plots the average annualized yield curves for real and nominal bonds in the model of Koijen et al. (2010). Panel C shows the nominal yield curve only for 1- to 20-month bonds (extract of Panel B).

0.9971 increases the error in the equity premium from about 5% to almost 35%. This result suggests that the posterior median estimates are also biased. We leave analyzing the quantitative impact of these errors on the parameter estimates to future research.

In contrast, the model of Bollerslev, Xu, and Zhou (2015) only features a persistent long-run risk process $\rho = 0.988$, while the persistence parameters of the stochastic volatility and vol-of-vol factors are considerably lower ($\nu = 0.64$ and $\rho_q = 0.46$). The resulting approximation errors are rather small, with a maximum error of 4.52% for the equity premium. This result is not surprising as the authors mention in their estimation that the stochastic volatility and vol-of-vol factors influence the variance premium but have a negligible effect on the price and return dynamics. As expected, we obtain almost the same results when setting the volatility of the two factors to zero ($\phi_\sigma = \phi_q = 0$).

For the study of Koijen et al. (2010), we also find large errors, with a maximum error in the equity premium of 33.99% and an overestimation of the premium of about 150 bps. Their calibration features a highly persistent long-run risk process, $\rho = 0.991$, and highly persistent stochastic volatility of long-run risk, $\nu_x = 0.996$, which introduce large nonlinearities to the model. Koijen et al. (2010) not only analyze equity markets but also price real and nominal bonds to analyze the martingale component of the stochastic discount factor. In Figure 8, we plot the real and nominal average annualized yield curves for their model.

We find that the differences between the yield curves obtained by linearizing the model and solving it accurately using the projection approach are small in absolute value. However, the nominal yield curve from the linearized model differs in its shape. While the true nominal yield curve is downward sloping in the short run and upward sloping in the long run (see Panel C), this pattern does not obtain when using log-linearization. So linearizing the model potentially affects the shape of the real curve. The work of Bansal and Shaliastovich (2013) sheds further light on this finding. In Figure 9, we plot the nominal yield curve in their model.

Panel A shows the yield curve for the parameters in the original study. We find that the difference between the log-linearized solution and the

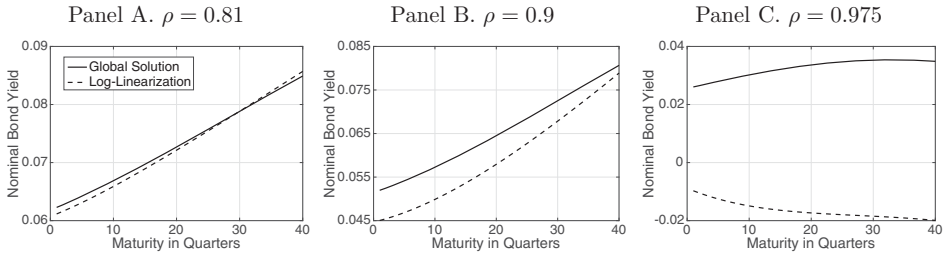

Figure 9. Nominal yield curve in the model of Bansal and Shaliastovich (2013). The figure plots the average annualized yield curve for nominal bonds in the model of Bansal and Shaliastovich (2013).

projection solution is negligible with very small errors, and the shape of the yield curve is also correct. As Bansal and Shaliastovich (2013) use bond data to estimate the model, they find very low persistence in the long-run risk component, with $\rho = 0.81$. This comparably low amount of persistence makes it difficult to match key moments for equity markets. For example, the annualized equity premium for their parameter estimates is only 1.85%.[15] We therefore increase ρ in Panels B and C to 0.9 and 0.975 to increase the premium paid for long-run consumption risk.[16] We find that the errors in the yield curve grow significantly as ρ approaches one. In fact, for $\rho = 0.975$, the log-linearization predicts a downward-sloping nominal yield curve (dashed line) even though the model actually produces an upward-sloping curve (solid line). Hence, relying on the log-linearization to solve the model can lead to incorrect conclusions not only about the magnitude of bond yields but also about the shape of the yield curve.

In sum, we find that, while the log-linearization approach produces satisfactory solutions for an analysis of the models in Bansal and Yaron (2004) and Bollerslev, Xu, and Zhou (2015), the method performs rather poorly for the models in Bansal, Kiku, and Yaron (2012), Schorfheide, Song, and Yaron (2018), and Koijen et al. (2010). For these latter models, the poor approximations have a strong effect on the model predictions for key financial statistics.

IV. Conclusion

In this paper, we analyze higher order effects in asset pricing models with long-run risk. We show that, as the underlying exogenous processes become

[15] The published version of Bansal and Shaliastovich (2013) does not provide a process for dividend growth. For the purposes of comparison, we consider the specification that appears in their 2007 working paper. The process for Δd_{t+1} is the same as in Koijen et al. (2010) (see equation 16). As the 2007 working paper assumes a monthly decision interval and the published version from 2013 employs a quarterly interval, we adjust the volatility of dividends ϕ_d to match the volatility of dividend growth in the data of approximately 11% annualized.

[16] For $\rho = 0.9$ we obtain an equity premium of 4.72%, and for $\rho = 0.975$ we obtain a premium of 11.14%.

increasingly persistent, nonlinear dynamics become increasingly important. Since persistent processes are required to match key properties of asset prices, such as the high equity premium and return predictability, these dynamics prove to be important for the current leading long-run risk models. For these models, use of the Campbell-Shiller (1988) log-linearization introduces significant approximation errors in model moments. For example, for the calibration of Bansal, Kiku, and Yaron (2012), log-linearization introduces an error of 22% in the volatility of the log price-dividend ratio and overstates the amount of return predictability by a factor of two. For highly persistent processes, as regularly used in the literature, the approximation errors in moments can exceed 70%. (Models with lower persistence, such as the original Bansal and Yaron (2004) model or the vol-of-vol model of Bollerslev, Xu, and Zhou (2015), have much smaller approximation errors.) The results for nominal bonds models (such as Koijen et al. (2010) and Bansal and Shaliastovich (2013)) are particularly interesting: for the high level of persistence necessary to explain the equity premium, log-linear approximation can actually produce a downward-sloping yield curve, when the true yield curve is upward sloping.

For many parameter combinations, the numerical errors introduced by the Campbell-Shiller (1988) log-linearization are small. If the driving factors of the economy are not highly persistent, or if the risk aversion of the representative agent is low, these dynamics will have a negligible influence on equilibrium outcomes. In these cases, log-linearization is a very attractive solution procedure. However, the combination of highly persistent processes together with recursive preferences and a risk aversion significantly larger than one—features that the long-run risk model requires to be consistent with financial market data—can introduce strong nonlinear dynamics to the model. We show that these errors have a strong effect on key financial statistics in some recent asset pricing studies, which in turn introduce a bias to the model parameters when it comes to the estimation or calibration of the model. Therefore, when considering highly persistent processes, researchers should not rely on log-linearized solutions alone, but rather should use more sophisticated methods that can account for higher order dynamics.

Initial submission: December 5, 2016; Accepted: May 6, 2017
Editors: Stefan Nagel, Philip Bond, Amit Seru, and Wei Xiong

Appendix A: Computational Methods for Asset Pricing Models with Recursive Preferences

One of the common approaches to solving asset pricing models is to log-linearize the model around its steady state. A discussion of log-linearization methods requires careful attention to several important differences among some well-known approaches. Standard log-linearization methods as in Judd (1996) or Collard and Juillard (2001) linearize around the deterministic steady state of the model. In a deterministic model, recursive preferences collapse to the case of CRRA preferences and hence risk aversion has no influence (as there

is no risk). But if risk aversion has significant influence in the stochastic model, linearizing around the deterministic steady state might not be the best choice. Therefore, new techniques have been developed that linearize around the risky steady state of the model (see, e.g., Juillard (2011), de Groot (2013), and Meyer-Gohde (2014)).[17] Another drawback of the standard log-linearization is that the policies are independent of the volatility of the model (see Caldara et al. (2012)). But as Bansal and Yaron (2004) point out, stochastic volatility is one of the key features of the long-run risk model and essential for asset pricing dynamics. Hence, a log-linear approximation for asset pricing models with recursive preferences and stochastic volatility must account for both the risk adjustment of the steady state and the effects of volatility. Bansal and Yaron (2004) use a linearization technique based on the Campbell and Shiller (1988) return approximation that meets these requirements and therefore has been used extensively in solving asset pricing models with recursive preferences (see Bollerslev, Tauchen, and Zhou (2009), Bansal, Kiku, and Yaron (2010), Koijen et al. (2010), Constantinides and Ghosh (2011), Drechsler and Yaron (2011), Bansal, Kiku, and Yaron (2012), Beeler and Campbell (2012), Bansal and Shaliastovich (2013), Bansal et al. (2014), and Segal, Shaliastovich, and Yaron (2015), among others).[18] One reason for the popularity of this approach is that it allows for approximate closed-form solutions for many different model specifications, for example, when shocks to the economy are normal. The log-linearization technique to solve asset pricing models with recursive preferences is described below.

This study analyzes the log-linearized model solution with regard to the influence of higher order dynamics on equilibrium outcomes that, by construction, cannot be captured by the log-linear approximation described below. For CRRA preferences, closed-form solutions for various model specifications can be computed. Unfortunately, for the general case of recursive preferences, to the best of our knowledge there are no such solutions. We therefore need a highly accurate solution method that is capable of correctly capturing higher order features of asset returns. A convenient choice is projection methods, which allow us to choose the approximation degree as well as the size of the approximation interval to capture higher order dynamics driven by the tails of the distribution.[19] Projection methods are a general-purpose tool for solving functional equations. They were first introduced by physicists and engineers

[17] These authors define the risky steady state as the state in which, in absence of shocks in the current period, the agent decides to stay in the current state while expecting shocks in the future and knowing their probability distribution.

[18] Another approach, proposed by Kogan and Uppal (2001) and used in Hansen et al. (2007) and Hansen, Heaton, and Li (2008), is to linearize around the special case of unit elasticity of substitution $\psi = 1$, where the wealth-consumption ratio is constant. However most of the subsequent work in the long-run risk literature focuses on the log-linearization used in Bansal and Yaron (2004), so we concentrate on this particular approximation.

[19] Similar results could potentially be obtained by using a perturbation method with a sufficiently high order. But as we do not seek to find the best or most efficient solution method, but rather we are interested in analyzing higher order dynamics, we choose to use the projection approach.

to solve partial differential equations, but they can be used to solve the types of fixed-point equations that arise in economics. (See Judd (1992) for an introduction or Chen, Cosimano, and Himonas (2014) for a brief overview of how to apply projection methods to asset pricing models.) We now provide a detailed description of projection methods and how they can be applied to solve the equilibrium conditions (2) and (6).

A. Log-Linearization Applied to Asset Pricing Models with Recursive Preferences

Here, we provide a short sketch of the linearization method, as in Bansal and Yaron (2004). For a detailed description of the method, see Eraker (2008) and Eraker and Shaliastovich (2008). Assume that the log price-dividend ratio of asset i, $z_{i,t}$, is a linear function of the state variables

$$z_{i,t} = A_{0,i} + A_i y_t, \tag{A1}$$

where $y_t \in \mathbb{R}^l$ is the state vector describing the economy and $A_{0,i} \in \mathbb{R}^1$ and $A_i \in \mathbb{R}^l$ are the unknown linearization coefficients. The log return of the asset i, $r_{i,t+1}$, is then defined as

$$r_{i,t+1} = \log(e^{z_{i,t+1}} + 1) - z_{i,t} + \Delta d_{i,t+1}, \tag{A2}$$

where $\Delta d_{i,t+1}$ is the log dividend growth rate. Making use of the Campbell and Shiller (1988) return approximation, one gets

$$r_{i,t+1} \approx \kappa_{i,0} + \kappa_{i,1} z_{i,t+1} - z_{i,t} + \Delta d_{i,t+1}, \tag{A3}$$

with linearization constants

$$\kappa_{i,1} = \frac{e^{\bar{z}_i}}{1 + e^{\bar{z}_i}}, \tag{A4}$$

$$\kappa_{i,0} = -\log\left((1-\kappa_{i,1})^{1-\kappa_{i,1}} \kappa_{i,1}^{\kappa_{i,1}}\right) \tag{A5}$$

that depend only on the model-implied mean price-dividend ratio $\bar{z}_i = A_{0,i} + A_i E(y_t)$. Plugging the return approximation for the return on wealth (A3) into equilibrium condition (6) yields

$$\mathrm{E}_t\left[e^{\theta \log \delta + \left(\theta - \frac{\theta}{\psi}\right)\Delta c_{t+1} + \theta\left(\kappa_{w,0} + \kappa_{w,1} z_{w,t+1} - z_{w,t}\right)}\right] = 1. \tag{A6}$$

The equilibrium condition now depends only on the state of the economy and the linearization coefficients $A_{0,i}$ and A_i. As the equilibrium equation has to hold for any realization of the state of the economy, one can collect the terms for each state to obtain a square system of $l+1$ equations. Once we have solved for the return on wealth, one can apply the linearization approach to the general pricing equation (2) to solve for the log price-dividend ratio of any asset i. For

certain state processes, such as for processes with normal innovations as in Bansal and Yaron (2004) or Bollerslev, Tauchen, and Zhou (2009), the expectation can be evaluated analytically. This allows for approximate closed-form solutions for the linearization coefficients that depend only on the linearization constants $\kappa_{i,0}$ and $\kappa_{i,1}$. Eraker (2008), Eraker and Shaliastovich (2008), and Drechsler and Yaron (2011) show how to generalize the approach to include general affine processes and jumps.

B. Projection Methods for Functional Equations

Projection methods (see Judd (1992) for an introduction or Chen, Cosimano, and Himonas (2014) for a brief overview) are a general tool to solve functional equations of the form

$$(\mathcal{G}z)(x) = 0, \tag{A7}$$

where x resides in a (state) space $X \subset \mathbb{R}^l$, $l \geq 1$, and z is an unknown solution function with domain X, so $z : X \to \mathbb{R}^m$. The operator $\mathcal{G}$ is a continuous mapping between two function spaces. Note that solving equation (A7) requires finding an element z in a function space, that is, in an infinite-dimensional vector space.

The first central step of a projection method is to approximate the unknown function z on its domain X using a linear combination of basis functions. For the applications in this paper, it suffices to assume that the domain X is bounded and that the basis functions are polynomials.[20] For a set $\{\Lambda_k\}_{k\in\{0,1,\ldots,n\}}$ of chosen basis functions, the approximation $\hat{z}$ of z is

$$\hat{z}(x;\boldsymbol{\alpha}) = \sum_{k=0}^{n} \alpha_k \Lambda_k(x), \tag{A8}$$

where $\boldsymbol{\alpha} = [\alpha_0, \alpha_1, \ldots, \alpha_n]$ are unknown coefficients. Replacing the function z in equation (A7) by its approximation $\hat{z}$, we can define the residual function $\hat{F}(x;\boldsymbol{\alpha})$ as the error in the original equation,

$$\hat{F}(x;\boldsymbol{\alpha}) = (\mathcal{G}\hat{z})(x;\boldsymbol{\alpha}). \tag{A9}$$

Instead of solving equation (A7) for the unknown function z, we now attempt to choose coefficients $\boldsymbol{\alpha}$ to make the residual $\hat{F}(x;\boldsymbol{\alpha})$ zero. Note that instead of finding an element in an infinite-dimensional vector space, we are now looking for a vector in $\mathbb{R}^{n+1}$. Obviously, this approximation step greatly simplifies the mathematical problem.

[20] For ease of notation, we demonstrate the projection approach using polynomials for the basis functions. Alternative specifications include, for example, rational basis functions or piecewise polynomial approximations. To solve models with high-dimensional state spaces like the model of Schorfheide, Song, and Yaron (2018), we use cubic splines as there is an efficient and fast implementation in Matlab.

This problem is unlikely to have an exact solution, so the second central step of a projection method is to impose certain conditions on the residual function, the so-called "projection" conditions, to make the problem solvable. In other words, the purpose of the projection conditions is to establish a set of requirements that the coefficients $\boldsymbol{\alpha}$ must satisfy. For a formulation of the projection conditions, define a "weight function" (term) $w(x)$ and a set of "test" functions $\{g_k(x)\}_{k=0}^n$. We can then define an inner product between the residual function $\hat{F}$ and the test function g_k,

$$\int_X \hat{F}(x;\boldsymbol{\alpha})g_k(x)w(x)dx.$$

This inner product induces a norm on the function space X. Natural restrictions for the coefficient vector α are now the projection conditions,

$$\int_X \hat{F}(x;\boldsymbol{\alpha})g_k(x)w(x)dx = 0,\ k = 0, 1, \ldots, n. \tag{A10}$$

Observe that this system of equations imposes $n+1$ conditions on the $(n+1)$-dimensional vector α. Different projection methods vary in the choice of the weight function and the set of test functions. In this paper, we describe two different projections, the collocation method and the Galerkin method.

The collocation method chooses $n+1$ distinct nodes in the domain, $\{x_k\}_{k=0}^n$, and uses Dirac delta functions as the test functions, $g_k = \delta(x - x_k)$. With a weight term $w(x) \equiv 1$, the projection conditions (A10) simplify to

$$\hat{F}(x_k;\boldsymbol{\alpha}) = 0,\ k = 0, 1, \ldots, n. \tag{A11}$$

Simply put, the collocation method determines the coefficients in the approximation (A8) by solving the square system (A11) of nonlinear equations.

The Galerkin method uses the fact that Chebyshev polynomials are orthogonal on $[-1, 1]$ with respect to the inner product using the weight function $w(x) \equiv \frac{1}{\sqrt{1-x^2}}$. Hence, the Galerkin method uses the basis functions as the test functions, $g_k(x) = \Lambda_k(x)$, and the projection conditions (A10) become

$$\int_X \hat{F}(x;\boldsymbol{\alpha})\Lambda_k(x)\frac{1}{\sqrt{1-x^2}}dx = 0,\ k = 0, 1, \ldots, n. \tag{A12}$$

Next, we show how to apply the general projection approach to solve the equilibrium pricing equations (6) and (2).

C. *Projection Methods Applied to Asset Pricing Models*

To apply a projection method to the asset pricing model, we express the equilibrium conditions as a functional equation of the type in (A7). For this purpose, we need to choose an appropriate state space and perform the usual transformation from an equilibrium described by infinite sequences (with time index t) to the equilibrium being described by functions of some state variable(s) x

on a state space X. We denote the current state of the economy by x and the state in the next period by x'. (For example, in the original model of Mehra and Prescott (1985), the state x is log consumption growth and $X \subset \mathbb{R}^1$; in the model of Bansal and Yaron (2004), the state x consists of the long-run mean of consumption growth (denoted by x_t in that paper) and the variance of consumption growth (denoted by σ_t^2), so $X \subset \mathbb{R}^2$.) We assume that the probability distribution of next period's state x' conditional on the current state x is defined by density f_x.

We solve the model in two steps. In the first step, we use the projection method to solve the wealth-Euler equation (6) to obtain the return on wealth. Once the return on wealth is known, then we proceed to the second step and can solve for any asset return by applying the projection approach to equation (2). In the first step, write equation (6) in state-space representation as

$$\mathrm{E}\left[\exp\left(\theta \log\delta - \frac{\theta}{\psi}\Delta c(x'|x) + \theta r_w(x'|x)\right)\middle| x\right] = 1, \quad \forall x, \tag{A13}$$

where lower case letters denote logs of variables and $\Delta c(x'|x) = c(x') - c(x)$. We write the model in logs because the function we solve for is the log wealth-consumption ratio $z_w(x) = \log(\frac{W(x)}{C(x)})$. Next, we write the state-dependent log return of the aggregate consumption claim as

$$\begin{aligned} r_w(x'|x) &= \log\left(\frac{W(x')}{W(x) - C(x)}\right) = \log\left(\frac{\frac{W(x')}{C(x')}}{\frac{W(x)}{C(x)} - 1} \times \frac{C(x')}{C(x)}\right) \\ &= z_w(x') - \log(e^{z_w(x)} - 1) + \Delta c(x'|x). \end{aligned} \tag{A14}$$

Inserting the last term in equation (A13) yields

$$\mathrm{E}\left[\exp\left(\theta\left(\log\delta + \left(1 - \frac{1}{\psi}\right)\Delta c(x'|x) + z_w(x') - \log(e^{z_w(x)} - 1)\right)\right) - 1\middle| x\right] = 0, \quad \forall x. \tag{A15}$$

Equivalently,

$$0 = \int_X \left[\exp\left(\theta\left(\log\delta + \left(1 - \frac{1}{\psi}\right)\Delta c(x'|x) + z_w(x') - \log(e^{z_w(x)} - 1)\right)\right) - 1\right] \mathrm{d}f_x, \tag{A16}$$

which is a functional equation of the form (A7) that allows us to apply the projection approach.

The unknown solution function to this equilibrium condition, z_w, is an element of a function space that is an infinite-dimensional vector space. A key feature of every projection method is that one approximates the solution function z_w by an element from a finite-dimensional space. We use the approximation $\hat{z}_w(x; \boldsymbol{\alpha_w}) = \sum_{k=0}^{n} \alpha_{w,k}\Lambda_k(x)$, where $\{\Lambda_k\}_{k\in\{0,1,\ldots,n\}}$ is a set of chosen (known)

basis functions and $\boldsymbol{\alpha_w} = [\alpha_{w,0}, \alpha_{w,1}, \ldots, \alpha_{w,n}]$ are unknown coefficients. Replacing the exact solution $z_w(x)$ by the approximation $\hat{z}_w(x;\boldsymbol{\alpha_w})$ leads us to the residual function $\hat{F}_w$ for the rearranged wealth-Euler equation (A16), which is defined by

$$\hat{F}_w(x;\boldsymbol{\alpha_w}) \tag{A17}$$
$$= \int_X \left[\exp\left(\theta \left(\log\delta + \left(1 - \frac{1}{\psi}\right) \Delta c(x'|x) + \hat{z}_w(x') - \log(e^{\hat{z}_w(x)} - 1) \right) \right) - 1 \right] \mathrm{d}\, f_x.$$

We can determine values for the unknown solution coefficients $\boldsymbol{\alpha_w}$ by imposing a projection condition on the residual term $\hat{F}_w(x;\boldsymbol{\alpha_w})$. In this paper, we employ two such projection conditions, the collocation method and the Galerkin method (see earlier). The values for the coefficients $\boldsymbol{\alpha_w}$ determine the state-dependent wealth-consumption ratio $\hat{z}_w(x;\boldsymbol{\alpha_w})$, which in turn leads to the (approximate) return function of the aggregate consumption claim, $\hat{r}_w(x'|x;\boldsymbol{\alpha_w}) = \hat{z}_w(x';\boldsymbol{\alpha_w}) - \log(e^{\hat{z}_w(x;\boldsymbol{\alpha_w})} - 1) + \Delta c(x'|x)$.

With $\hat{r}_w(x'|x;\boldsymbol{\alpha_w})$ at hand, we can now develop an approach to compute the return of any asset i using equation (2). Analogous to the first step, we solve for the log price-dividend ratio $z_i(x) = \log(\frac{P(x)}{D(x)})$ and rewrite the state-dependent log return of asset i as

$$r_i(x'|x) = \log\left(\frac{P_i(x') + D_i(x')}{P_i(x)}\right) = \log\left(\frac{\frac{P_i(x')}{D_i(x')} + 1}{\frac{P_i(x)}{D_i(x)}} \times \frac{D_i(x')}{D_i(x)}\right)$$
$$= \log\left(e^{z_i(x')} + 1\right) - z_i(x) + \Delta d_i(x'|x). \tag{A18}$$

Writing the Euler equation (2) in state-space representation and formulating it in logs yields

$$\mathrm{E}\left[\exp\left(\theta\log\delta - \frac{\theta}{\psi}\Delta c(x'|x) + (\theta - 1) r_w(x'|x) + r_i(x'|x) \right) \middle| x \right] = 1. \tag{A19}$$

Substituting the return expressions (A14) and (A18) into this equation and replacing the log price-dividend ratio $z_i(x) = p_i(x) - d_i(x)$ by its approximation $\hat{z}_i(x;\boldsymbol{\alpha_i}) = \sum_{k=0}^{n} \alpha_{i,k}\Lambda_k(x)$ leads to the residual function

$$\hat{F}_i(x;\boldsymbol{\alpha_i}) = \int_X \left[\exp\left(\theta\log\delta - \frac{\theta}{\psi}\Delta c(x'|x) + (\theta - 1)\hat{r}_w(x'|x;\boldsymbol{\alpha_w}) \right.\right.$$
$$\left.\left. + \log(e^{\hat{z}_i(x';\boldsymbol{\alpha_i})} + 1) - \hat{z}_i(x;\boldsymbol{\alpha_i}) + \Delta d_i(x'|x) \right) - 1 \right] \mathrm{d}\, f_x. \tag{A20}$$

Recall that the coefficients $\boldsymbol{\alpha_w}$ and thus the function $\hat{r}_w(x'|x;\boldsymbol{\alpha_w})$ have been computed previously. Therefore, we can now apply one of the projection conditions to solve for the unknown vector $\boldsymbol{\alpha_i}$.

In sum, we apply the projection method twice. In the first step, we approximate the log wealth-consumption ratio $\hat{z}_w(x;\boldsymbol{\alpha_w})$ by applying the projections on the residual function of the wealth-Euler equation (A17). Once $\boldsymbol{\alpha_w}$ is

known, the projections can be applied to equation (A20) to solve for the price-dividend ratio $\hat{z}_i(x;\boldsymbol{\alpha_i})$ of any asset i. Formally, the algorithm can be described as follows.

Algorithm. Solving Asset Pricing Models with Recursive Preferences.

Initialization. Define the state space $X \subset \mathbb{R}^l$; choose the functional forms for $\hat{z}_w(x;\boldsymbol{\alpha_w})$ and $\hat{z}_i(x;\boldsymbol{\alpha_i})$ as well as the projection method.

Step 1. Use the wealth–Euler equation (6) together with the approximated log wealth–consumption ratio $\hat{z}_w(x;\boldsymbol{\alpha_w})$ and the definition of the return equation (A14) to derive the residual function for the return on wealth

$$\hat{F}_w(x;\boldsymbol{\alpha_w}) = \int_X \left[\exp\left(\theta\left(\log\delta + \left(1-\frac{1}{\psi}\right)\Delta c(x'|x) + \hat{z}_w(x') - \log(e^{\hat{z}_w(x)}-1)\right)\right) - 1\right] \mathrm{d}f_x.$$

Compute the unknown solution coefficients $\boldsymbol{\alpha_w}$ by imposing the projections on $\hat{F}_w(x;\boldsymbol{\alpha_w})$.

Step 2. Use the solution for the wealth–consumption ratio $\hat{z}_w(x;\boldsymbol{\alpha_w})$ and the Euler equation (2) for asset i together with the approximated log price–dividend ratio $\hat{z}_i(x;\boldsymbol{\alpha_i})$ and the definition of the return equation (A18) to derive the residual function for asset i,

$$\hat{F}_i(x;\boldsymbol{\alpha_i}) = \int_X \left[\exp\left(\theta\log\delta - \frac{\theta}{\psi}\Delta c(x'|x) + (\theta-1)\hat{r}_w(x'|x;\boldsymbol{\alpha_w}) + \log\left(e^{\hat{z}_i(x';\boldsymbol{\alpha_i})}+1\right) - \hat{z}_i(x;\boldsymbol{\alpha_i}) + \Delta d_i(x'|x)\right) - 1\right] \mathrm{d}f_x$$

Compute the unknown solution coefficients $\boldsymbol{\alpha_i}$ by imposing the projections on $\hat{F}_i(x;\boldsymbol{\alpha_i})$.

Evaluation. Choose a set of evaluation nodes $\mathbb{X}^e = \{x_j^e : 1 \le j \le m^e\} \subset X$ and compute approximation errors in the residual function of the wealth portfolio and the residual function of asset i. If the errors do not satisfy a predefined error bound, start over at *Initialization* and change the number of approximation nodes or the degree of the basis functions.

To implement the algorithm, we need to specify additional algorithmic details such as the choice of basis functions and the integration technique.

D. Algorithmic Ingredients

In the *Initialization* step of the algorithm mentioned above, we need to choose a set of basis functions for the polynomial approximation, a projection method, and a set of nodes. To simplify the presentation, we describe the necessary choices for a one-dimensional state space approximated over an interval $X = [x_{min}, x_{max}]$. We approximate the solution functions z_w and z_i using Chebyshev

polynomials (of the first kind) (see Judd (1998)). We obtain the Chebyshev polynomials via the recursive relationship

$$T_0(\xi) = 1, \quad T_1(\xi) = \xi, \quad T_{k+1}(\xi) = 2\xi T_k(\xi) - T_{k-1}(\xi),$$

with $T_k : [-1, 1] \to \mathbb{R}$. Since we need to approximate functions on the domain X and the Chebyshev polynomials are defined on the interval $[-1, 1]$, we need to transform the argument for the polynomials. The basis functions for the approximate solutions $\hat{z}_w(x; \boldsymbol{\alpha_w})$ and $\hat{z}_i(x; \boldsymbol{\alpha_i})$ are given by

$$\Lambda_k(x) = T_k\left(2\left(\frac{x - x_{\min}}{x_{\max} - x_{\min}}\right) - 1\right) \tag{A21}$$

for $k = 0, 1, \ldots, n$.

In this paper, we only show the results using the collocation method but we verified the solutions using the Galerkin approach. The application of a projection method requires a set of nodes, $\mathbb{X} = \{x_j : 0 \le j \le m\} \subset X$; we choose the $m+1$ zeros of the Chebyshev polynomial T_{m+1}. These points are called Chebyshev nodes,

$$\xi_j = \cos\left(\frac{2j+1}{2m+2}\pi\right), \qquad j = 0, 1, \ldots, m.$$

Since all Chebyshev nodes are in the interval $[-1, 1]$, we need to transform them to obtain nodes in the state space X. This transformation is

$$x_j = x_{\min} + \frac{x_{\max} - x_{\min}}{2}(1 + \xi_j), \qquad j = 0, 1, \ldots, m.$$

For the collocation method, the number of basis functions, $n + 1$, must be identical to the number of approximation nodes, $m + 1$, and so $m = n$. In *Step 1* (and *Step 2*, if applicable), we must solve the projection conditions involving the residual function. The residual functions defined in equations (A17) and (A20) contain a conditional expectations operator, which also requires numerical calculations. The underlying exogenous processes in the models we consider are normally distributed, and so we apply Gauss-Hermite quadrature to calculate expectations.

The collocation approach leads to a square system of nonlinear equations, which can be solved with a standard nonlinear equation solver. The Galerkin projection is slightly more complex and uses integral operators as projection conditions; these in turn can be accurately approximated by Gauss-Chebyshev quadrature.

For the *Evaluation* step, we use $m^e >> m$ equally spaced evaluation nodes in X to evaluate the errors in the residual function. In particular, for asset i we compute the root mean squared Euler error (RMSE) and maximum absolute

Euler error (MAE) in the residual function (A20). These errors are

$$\text{RMSE}_i = \sqrt{\frac{1}{m^e}\sum_{j=1}^{m^e}\hat{F}_i\left(x_j^e|\boldsymbol{\alpha_i}\right)^2}, \tag{A22}$$

$$\text{MAE}_i = \max_{j=1,2,\ldots,m^e}\left|\hat{F}_i\left(x_j^e|\boldsymbol{\alpha_i}\right)\right|, \tag{A23}$$

respectively, with

$$x_j^e = x_{\min} + \frac{x_{\max}-x_{\min}}{m^e-1}(j-1), \qquad j = 1,\ldots,m^e. \tag{A24}$$

Appendix B: Computational Details and Accuracy of the Projection Method

All of the results in the paper are computed using Matlab. We use the solver "fmincon" with the SQP algorithm. Since we have to solve a system of nonlinear equations for the projection approach and not an optimization problem, this is implemented by minimizing a constant subject to the nonlinear constraints from the system of equations. This procedure has proven to be far more efficient and robust than the simple use of "fsolve." For the high-dimensional models of Schorfheide, Song, and Yaron (2018) and Koijen et al. (2010), we use cubic splines with not-a-knot end conditions instead of Chebyshev polynomials due to the faster implementation in Matlab.

Table BI demonstrates the accuracy of the projection approach. We consider the long-run risk model of Bansal and Yaron (2004) with constant volatility, where there exist closed-form solutions for the case of CRRA preferences (see de Groot (2015)). In the case of recursive preferences, we determine the correct solution using the projection approach with a very large degree and state space. (We use $n_\sigma = 50$ and increase the degree until the highest order coefficient is close to zero. We double check the solution by using the discretization method of Tauchen and Hussey (1991) with a very large number of discretization nodes.) We report errors in the mean and volatility of the wealth-consumption ratio for the log-linearization, the collocation projection, as well as the discretization method of Tauchen and Hussey (1991) for different numbers of discretization nodes.[21] Another class of methods popular in macroeconomics is perturbation methods (see, e.g., Caldara et al. (2012)). de Groot (2015) compares these methods to the analytical solutions and finds that they perform worse than even the log-linearization for long-run risk models, so we do not consider them further.

We find that, for the calibration with $\rho = 0.95$, a first-order approximation with an approximation interval of $n_\sigma = 1$ standard deviation around the

[21] We report the results when solving the Euler equation for wealth. Alternatively, we could solve the fixed-point equation for utility. The results using this alternative approach are almost identical—the coefficients differ by less than 10^{-12}.

Table BI

Accuracy of the Projection Method: Model Moments

The table reports the mean wealth-consumption ratio for the long-run risk model of Bansal and Yaron (2004) with constant volatility (equations (7) with $\sigma_t = \bar{\sigma}\ \forall\, t$). Results are shown for the log-linearization, the projection, as well as the discretization by Tauchen and Hussey (1991) with the extension of Flodén (2008) that performs better for highly persistent processes. For the projection method, results are shown for three different degrees n, where the approximation interval is set up n_σ standard deviations around the unconditional mean of the long-run risk process x_t. For the discretization method, results are shown for three numbers of approximation nodes n_D. The table also reports the relative error of the solutions; in the case of $\gamma = 1/\psi$ the closed-form solution is taken from de Groot (2015); in the case of $\gamma \neq 1/\psi$ we compute the accurate solution by solving the model using the discretization method with a very large number of discretization nodes or equivalently the projection method with a very large degree and state space. We use the same calibration of Bansal and Yaron (2004) with $\delta = 0.9989$, $\mu_c = 0.0015$, $\bar{\sigma} = 0.0078$, and $\phi_x = 0.044$.

	Closed-F.	Log-Lin	Projection			Discretization		
			$n = 1$ $n_\sigma = 1$	$n = 4$ $n_\sigma = 4$	$n = 16$ $n_\sigma = 32$	$n_D = 5$	$n_D = 10$	$n_D = 50$
Panel A: $\psi = 1.5, \gamma = 1/\psi$								
$\rho = 0.95$								
E($\frac{W}{C}$)	1,681.20	1,681.16	1,681.18	1,681.20	1,681.20	1,670.59	1,671.43	1,671.71
Error	0	2.11e-5	1.19e-5	2.61e-8	2.60e-8	0.0063	0.0058	0.0056
std($\frac{W}{C}$)	12.1815	12.1812	12.1813	12.1815	12.1815	11.8154	11.8348	12.1120
Error	0	2.20e-5	1.27e-5	2.65e-8	2.64e-8	0.0301	0.0285	0.0057
$\rho = 0.99$								
E($\frac{W}{C}$)	1,868.36	1,862.93	1,865.54	1,868.36	1,868.36	3,492.27	2,134.01	1,856.55
Error	0	0.0029	0.0015	1.21e-7	7.65e-11	0.8692	0.1422	0.0063
std($\frac{W}{C}$)	144.14	143.65	143.86	144.14	144.14	2046.97	388.35	143.17
Error	0	0.0034	0.0020	1.17e-7	8.17e-11	13.2010	1.6942	0.0067
Panel B: $\psi = 1.5, \gamma = 10$								
$\rho = 0.95$								
E($\frac{W}{C}$)	–	1,314.39	1,314.59	1,314.61	1,314.61	1,532.25	1,514.25	1,508.08
Error	–	1.66e-4	1.51e-5	4.37e-11	2.12e-12	0.1655	0.1518	0.1472
std($\frac{W}{C}$)	–	9.4941	9.4956	9.4956	9.4956	10.2019	10.1502	10.3674
Error	0	1.49e-4	2.37e-6	3.10e-11	2.58e-12	0.0744	0.0689	0.0918
$\rho = 0.99$								
E($\frac{W}{C}$)	–	517.13	518.97	529.39	529.39	869.23	653.99	570.65
Error	–	0.0231	0.0196	1.43e-9	4.91e-11	0.6419	0.2353	0.0779
std($\frac{W}{C}$)	–	35.2376	35.3660	35.5695	35.5695	117.2323	59.5289	37.6206
Error	0	0.0093	0.0057	3.49e-10	1.56e-12	2.2959	0.6736	0.0577

unconditional mean of x_t provides a very accurate solution with an approximation error of 1.51e-5 for E($\frac{W}{C}$) and 2.37e-6 for std($\frac{W}{C}$) for the case with recursive utility and $\gamma = 10$. For the high persistence case with $\rho = 0.99$, a larger degree is required, with a degree-four polynomial sufficient to compute a highly accurate solution.

In Table BII we report corresponding Euler errors in the pricing equation of the wealth-consumption ratio (6). In this exercise, we evaluate the pricing

Table BII

Accuracy of the Projection Method: Euler Errors

The table reports the MAE as well as the RMSE for the long-run risk model of Bansal and Yaron (2004) with constant volatility (equations (7) with $\sigma_t = \bar{\sigma}\ \forall\ t$). Results are shown for the log-linearization and the projection approach. For the projection method, results are shown for three different degrees n, where the approximation interval is set up n_σ standard deviations around the unconditional mean of the long-run risk process x_t. Euler errors are computed by evaluating the pricing equation of the wealth-consumption ratio (6) at 10,000 equally spaced nodes on ± 6 standard deviations around the unconditional mean of x_t. We use the same calibration of Bansal and Yaron (2004) with $\delta = 0.9989$, $\mu_c = 0.0015$, $\bar{\sigma} = 0.0078$, and $\phi_x = 0.044$.

	Closed-F.	Log-Lin	Projection		
			$n = 1$ $n_\sigma = 1$	$n = 4$ $n_\sigma = 4$	$n = 16$ $n_\sigma = 32$
Panel A: $\psi = 1.5, \gamma = 1/\psi$					
$\rho = 0.95$					
MAE	0	5.12e-7	5.07e-7	2.53e-11	1.61e-11
RMSE	0	2.25e-7	2.22e-7	1.45e-11	1.57e-11
$\rho = 0.99$					
MAE	0	6.58e-5	6.49e-5	1.28e-10	9.97e-14
RMSE	0	2.57e-5	2.51e-5	6.56e-11	5.61e-14
Panel B: $\psi = 1.5, \gamma = 10$					
$\rho = 0.95$					
MAE	–	2.06e-5	1.74e-5	2.40e-11	9.54e-14
RMSE	–	1.00e-5	7.60e-6	5.71e-12	2.67e-14
$\rho = 0.99$					
MAE	–	0.0051	0.0048	4.16e-9	5.76-12
RMSE	–	0.0020	0.0019	8.82e-10	3.72-12

equation at 10,000 equally spaced nodes on ± 6 standard deviations around the unconditional mean of x_t and compute both the MAE and RMSE. Results are reported for the two continuous methods, that is, for log-linearization and projection, for which these errors can be computed. We observe similar patterns as for the model moments. The Euler errors for the projection method with a degree-four polynomial are consistently below 4e-9. The Euler errors of the log-linearization are rather small for the low persistence cases but increase significantly for high persistence and high risk aversion—the calibration that is most often used in modern asset pricing models as shown in this paper. Overall, we find that the projection method provides highly accurate solutions for all specifications considered in this example.

Appendix C: Closed-Form Solutions for $\psi = 1$

For the special case of an EIS of one, the (log) stochastic discount factor in the long-run risk model is linear. Below we provide the derivations for the

closed-form solutions for the standard long-run risk model with stochastic volatility (7). For $\psi = 1$, indirect utility V_t is given by

$$V_t = C_t^{1-\delta} \left(\mathrm{E}_t V_{t+1}^{1-\gamma} \right)^{\frac{\delta}{1-\gamma}} \tag{C1}$$

(see Tallarini (2000)). Define $vc_t = \log \frac{V_t}{C_t}$ to obtain

$$vc_t = \frac{\delta}{1-\gamma} \log \mathrm{E}_t \left(e^{(1-\gamma)(vc_{t+1}+\Delta c_{t+1})} \right). \tag{C2}$$

Guess and verify that $vc_t = a + bx_t + c\sigma_t^2$:

$$\begin{aligned} a + bx_t + c\sigma_t^2 &= \frac{\delta}{1-\gamma} \log \mathrm{E}_t \left(e^{(1-\gamma)(a+bx_{t+1}+c\sigma_{t+1}^2)} \right) + \delta \left(\mu_c + x_t + 0.5(1-\gamma)\phi_c^2 \sigma_t^2 \right) \\ &= \delta a + \delta b \rho x_t + 0.5\delta(1-\gamma) b^2 \phi_x^2 \sigma_t^2 + \delta c \bar{\sigma}^2 (1-\nu) + \delta c \nu \sigma_t^2 \\ &\quad + 0.5\delta(1-\gamma) c^2 \phi_\sigma^2 + \delta \mu_c + \delta x_t + 0.5\delta(1-\gamma)\phi_c^2 \sigma_t^2. \end{aligned}$$

Solving for a, b, and c yields

$$\begin{aligned} b &= \frac{\delta}{1-\delta\rho}, \\ c &= \frac{0.5\delta(1-\gamma)(\phi_c^2 + b^2\phi_x^2)}{1-\delta\nu}, \\ a &= \frac{\delta\mu_c + \delta c\bar{\sigma}^2(1-\nu) + 0.5\delta(1-\gamma)c^2\phi_\sigma^2}{1-\delta}. \end{aligned}$$

The stochastic discount factor for the special case of $\psi = 1$ is given by

$$M_{t+1} = \delta \left(\frac{C_{t+1}}{C_t} \right)^{-1} \frac{V_{t+1}^{1-\gamma}}{\mathrm{E}_t(V_{t+1}^{1-\gamma})} \tag{C3}$$

(see Tallarini (2000)). Reformulating in terms of vc_t yields

$$\begin{aligned} M_{t+1} &= \delta \left(\frac{C_{t+1}}{C_t} \right)^{-1} \frac{e^{(1-\gamma)(vc_{t+1}+\Delta c_{t+1})}}{\mathrm{E}_t(e^{(1-\gamma)(vc_{t+1}+\Delta c_{t+1})})}, \\ &= \delta \left(\frac{C_{t+1}}{C_t} \right)^{-\gamma} \frac{e^{(1-\gamma)vc_{t+1}}}{e^{\frac{1-\gamma}{\delta}vc_t}}. \end{aligned} \tag{C4}$$

So the log stochastic discount factor $m_{t+1} = \log M_{t+1}$ is a linear function of the state variables given by

$$m_{t+1} = \log\delta - \gamma\Delta c_{t+1} + (1-\gamma)vc_{t+1} - \frac{1-\gamma}{\delta}vc_t. \tag{C5}$$

Appendix D: The Volatility-of-Volatility Factor

In this section, we analyze how log-linearization affects model outcomes when the model dynamics are described by a square root process as, for example, in Bollerslev, Tauchen, and Zhou (2009), Tauchen (2011), and Bollerslev, Xu, and Zhou (2015). To do so, we use the parsimonious model formulation as in Bollerslev, Tauchen, and Zhou (2009), who take the basic model setup (11) without the long-run risk factor, so $\phi_x = 0$, and add vol-of-vol modeled by a square root process q_t:

$$\begin{aligned} \sigma_{t+1}^2 &= \bar{\sigma}^2(1-\nu) + \nu\sigma_t^2 + \sqrt{q_t}\eta_{\sigma,t+1}, \\ q_{t+1} &= \mu_q(1-\rho_q) + \rho_q q_t + \phi_q\sqrt{q_t}\eta_{q,t+1} \\ &\quad \times \eta_{\sigma,t+1}, \eta_{q,t+1} \sim \text{i.i.d. } N(0,1). \end{aligned}$$

As Tauchen (2011) notes, care is needed because q_t can become negative in simulations if the volatility is too large compared to the mean of the process. The common approach in the literature is to assume a reflecting barrier at zero by replacing negative values with very small positive values to ensure positivity of the process (this approach is also used for the stochastic volatility process in the original Bansal and Yaron (2004) study and in many subsequent papers). However, to compute model solutions, the assumption of a nontruncated distribution for the log-linearization is commonly used.

Take, for example, the calibration of Bollerslev, Tauchen, and Zhou (2009) given by $\delta = 0.997$, $\gamma = 10$, $\psi = 1.5$, $\mu_c = 0.0015$, $\nu = 0.978$, $\bar{\sigma}^2 = 0.0078^2$, and $\mu_q =$ 1e-6. Figure D1 shows model outcomes for CRRA preferences with $\psi = 1.5$ (Panel A) and the corresponding Epstein-Zin case with $\gamma = 10$ (Panel B) for various persistence and volatility parameters of the vol-of-vol process ρ_q and ϕ_q. The black numbers show the true mean wealth-consumption ratio under the assumption of a reflecting boundary for q_t at zero. Gray values are the results from log-linearization under the assumption of a standard nontruncated normal distribution. Black circles denote convergence of the projection and the log-linearization approaches. Black diamonds denote cases in which the log-linearization yields a complex solution while the model solution using a truncated normal distribution is real. We find that, depending on the degree of risk aversion, using the standard log-linearization technique can lead to complex solutions. For example, this is the case for the calibration in Bollerslev, Tauchen, and Zhou (2009) with $\rho_q = 0.8$ and $\phi_q =$ 1e-3.[22]

So what factors are the determinants of the complexity of the linearized solution? The square root specification of q_t implies that the coefficient of q_t is determined by a quadratic equation and hence may have more than one

[22] Bollerslev, Tauchen, and Zhou (2009) provide a real solution by assuming a fixed value for the linearization constant $\kappa = 0.9$. However, this approach does not give a solution to the model but rather ex ante fixes the mean value of the price-dividend ratio and hence significantly biases the model outcome.

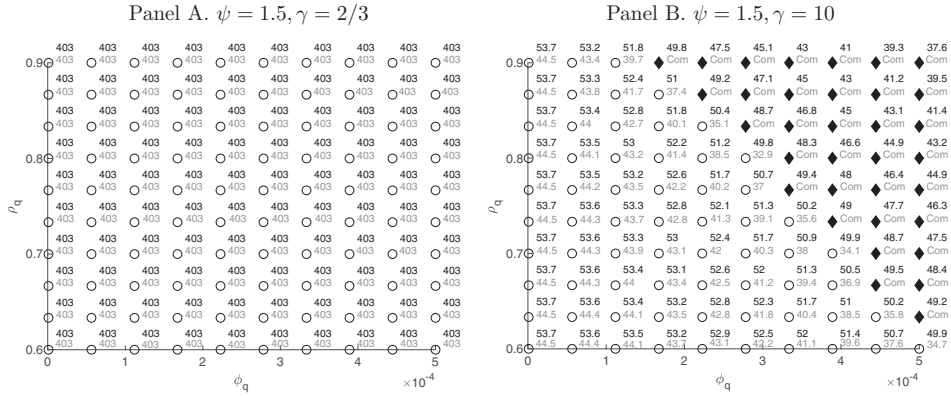

Figure D1. Sensitivity analysis and existence results in the vol-of-vol model. The figure shows the convergence properties as well as the mean wealth-consumption ratio for the vol-of-vol model of Bollerslev, Tauchen, and Zhou (2009). The results are reported for a range of persistence parameters ρ_q and volatility parameters ϕ_q. Panel A depicts the case of CRRA utility with $\psi = 1.5$, while Panel B depicts the corresponding case with recursive utility and $\gamma = 10$. Black numbers show the mean wealth-consumption ratio obtained by the projection approach using a reflecting barrier at zero and gray numbers show the values obtained by the standard log-linearization with normal shocks. Black circles denote convergence of the projection and the log-linearization approaches. Black diamonds denote cases in which the log-linearization yields a complex solution while the model solution using a truncated normal distribution is real. The model parameters are given by $\delta = 0.997$, $\mu_c = 0.0015$, $\nu = 0.978$, $\bar{\sigma}^2 = 0.0078^2$, and $\mu_q =$ 1e-6.

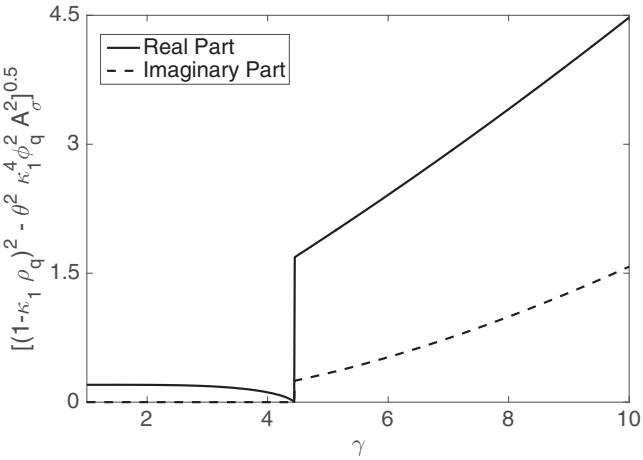

Figure D2. Analysis of square root term in the vol-of-vol model. The figure plots the real and complex parts of the square root term that determines A_q as a function of the risk aversion γ for the vol-of-vol model of Bollerslev, Tauchen, and Zhou (2009). The model parameters are given by $\delta = 0.997$, $\mu_c = 0.0015$, $\nu = 0.978$, $\bar{\sigma}^2 = 0.0078^2$, $\mu_q =$ 1e-6, $\rho_q = 0.8$, and $\phi_q = 1 \times 10^{-3}$.

solution. The log-linear approximation of the log wealth-consumption ratio $z_{w,t}$ has the form

$$z_{w,t} = A_0 + A_\sigma \sigma_t^2 + A_q q_t, \tag{D1}$$

with the linearization coefficients (see Appendix A for the derivation) given by

$$\begin{aligned} A_\sigma &= \frac{(1-\gamma)^2}{2\theta(1-k_1\nu)}, \\ A_0 &= \frac{\log\delta + (1-\frac{1}{\psi})\mu_c + k_0 + k_1\left[A_\sigma\bar{\sigma}^2(1-\nu) + A_q\mu_q(1-\rho_q)\right]}{(1-k_1)}, \\ A_q &= \frac{1-k_1\rho_q \pm \sqrt{(1-k_1\rho_q)^2 - \theta^2 k_1^4\phi_q^2 A_\sigma^2}}{\theta k_1^2\phi_q^2}. \end{aligned} \tag{D2}$$

We find that the coefficient of the vol–of–vol factor A_q does indeed have two solutions. As Bollerslev, Tauchen, and Zhou (2009) show in their paper by the no arbitrage argument, the negative term is the economically meaningful root and the positive solution can be neglected. Complexity of the solution is determined by the term inside the square root in equation (D2) given by $(1-k_1\rho_q)^2 - \theta^2 k_1^4\phi_q^2 A_\sigma^2$. So how does this term depend on the model parameters? Figure D2 plots the values of the square root term as a function of the risk aversion γ. In line with the results above, we find that, for small γ, the solution is well behaved with only a real and no imaginary part. However, if we increase γ, θ becomes significantly larger (it goes from -3 for $\gamma = 2$ to -27 for $\gamma = 10$) and hence the real part of the term decreases. At a certain threshold (about 4.4 in this example), the term hits zero after which the solution contains a significant imaginary part. Panel B of Figure D1 above shows that, for large values of the persistence or large values of the volatility of the vol-of-vol process, solutions are complex. Summarizing, using standard log-linearization with normal shocks to solve models with a large risk aversion and a persistent square root process can yield complex solutions, even if real solutions under the assumption of a reflecting barrier exist. Hence, when solving such models, either log-linearization with the assumption of a truncated normal distribution or more sophisticated methods like the projection approach described in this paper should be used.

Appendix E: Additional Tables and Graphs

Figure 6 in Section B plots isolines (black solid lines) for the absolute errors in the log wealth-consumption ratio (left panel) and the log price-dividend ratio (right panel) of the log-linearization as a function of the states x_t and σ_t^2 for the model with stochastic volatility in x_t. Figures E1 and E2 show the corresponding errors in the first and second derivatives of the log wealth-consumption and price-dividend ratio. Again, the errors in the derivatives of the price-dividend

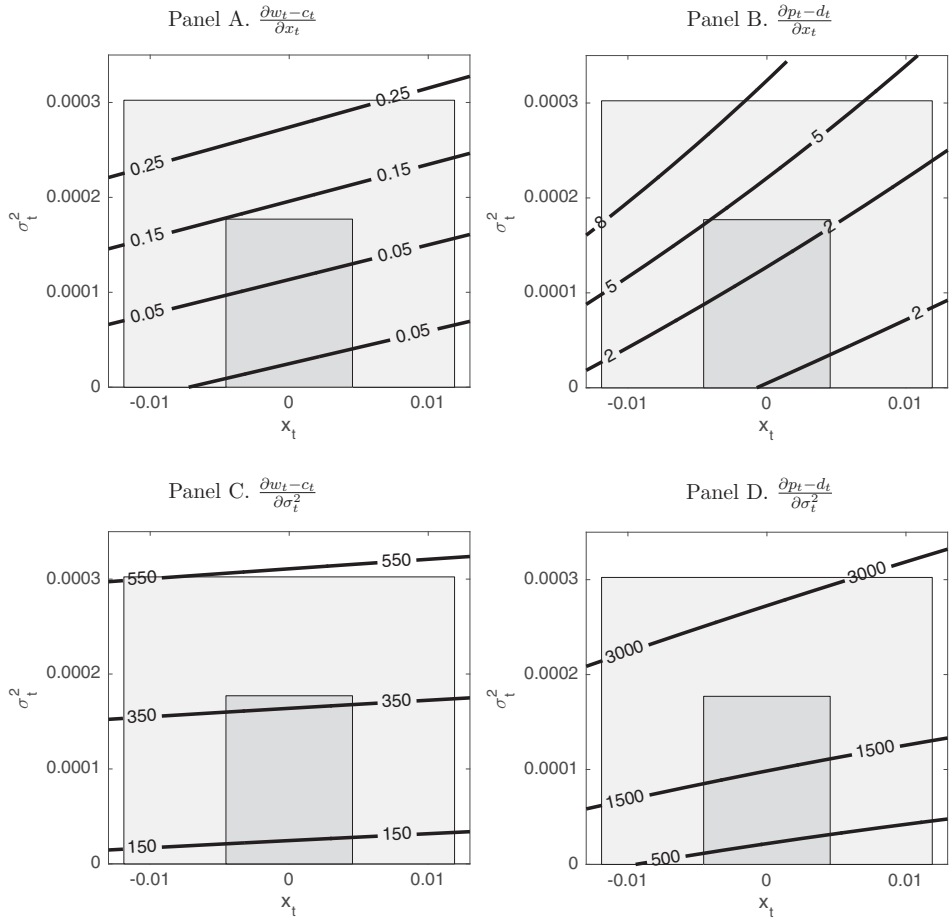

Figure E1. Absolute approximation errors in the first derivatives of the log wealth-consumption and log price-dividend ratio of the log-linearization. The figure plots isolines for the absolute errors in the first derivative of the log wealth-consumption ratio (left panel) and the log price-dividend ratio (right panel) with respect to the states x_t and σ_t^2 of the log-linearization (black solid lines). The dark gray area shows the range of two unconditional standard deviations of the state processes. For x_t the area is computed using a given two-unconditional-standard-deviation shock in σ_t^2. The light gray area shows the corresponding range for four unconditional standard deviations. Results are shown for the standard long-run risk model (specification (7)). Parameters: $\delta = 0.9989$, $\gamma = 10$, $\psi = 1.5$, $\mu = 0.0015$, $\phi_c = 1$, $\phi_x = 0.038$, $\rho = 0.975$, $\bar{\sigma} = 0.0072$, $\nu = 0.999$, $\phi_\sigma = 2.8\text{e-}6$, $\mu_d = 0.0015$, $\Phi = 2.5$, $\phi_d = 5.96$, and $\phi_{d,c} = 2.6$.

ratio are significantly larger than the errors in the derivatives of the wealth-consumption ratio, and the errors increase monotonically with σ_t^2. We find that the errors in the derivatives with respect to σ_t^2 are especially large, with errors as large as 3,000 for the first derivative of the price-dividend ratio. Also, we find that the second derivatives in the model are substantially different from zero

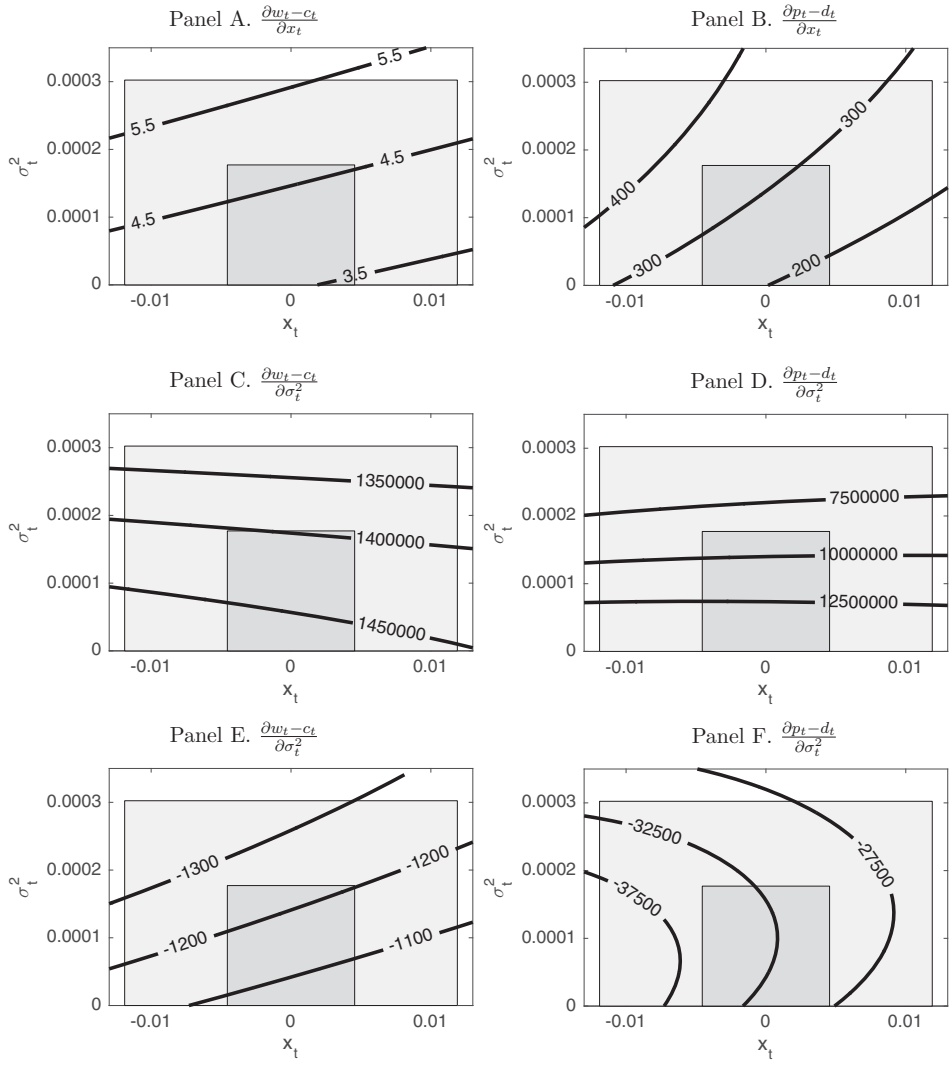

Figure E2. Absolute approximation errors in the second derivatives of the log wealth-consumption and log price-dividend ratio of the log-linearization. The figure plots isolines for the absolute errors in the second derivative of the log wealth-consumption ratio (left panel) and the log price-dividend ratio (right panel) with respect to the states x_t and σ_t^2 of the log-linearization (black solid lines). The dark gray area shows the range of two unconditional standard deviations of the state processes. For x_t the area is computed using a given two-unconditional-standard-deviation shock in σ_t^2. The light gray area shows the corresponding range for four unconditional standard deviations. Results are shown for the standard long-run risk model (specification (7)). Parameters: $\delta = 0.9989$, $\gamma = 10$, $\psi = 1.5$, $\mu = 0.0015$, $\phi_c = 1$, $\phi_x = 0.038$, $\rho = 0.975$, $\bar{\sigma} = 0.0072$, $\nu = 0.999$, $\phi_\sigma = 2.8\text{e-}6$, $\mu_d = 0.0015$, $\Phi = 2.5$, $\phi_d = 5.96$, and $\phi_{d,c} = 2.6$.

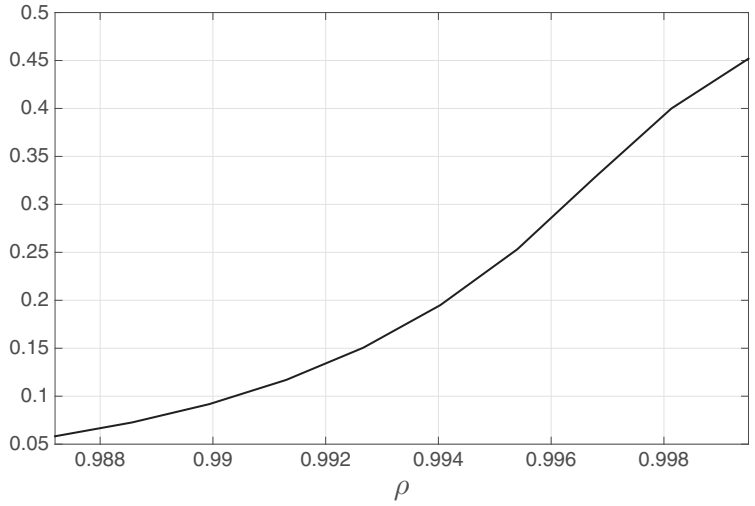

Figure E3. Relative errors in the equity premium in Schorfheide, Song, and Yaron (2018). The figure plots the relative errors in the equity premium as a function of the persistence of the long-run growth rate ρ in the model of Schorfheide, Song, and Yaron (2018). Results are shown for the median parameter estimates (see Table IV).

Table EI

Predictability of Excess Returns, Consumption, and Dividends in the Long-Run Risk Model of Bansal and Yaron (2004)

The table presents R^2 statistics and regression coefficients from regressing cumulative log excess returns, consumption growth, and dividend growth on the lagged log price-dividend ratio. Statistics are shown for the annualized time series with one-, three-, and five-year horizons for the model of Bansal and Yaron (2004). The table reports results for the log-linearized solution and the global solution as well as for the relative error of the log-linearization.

	R^2			β		
	1Y	3Y	5Y	1Y	3Y	5Y
$\sum_h^H (r_{m,t+h} - r_{f,t+h}) = \alpha + \beta(p_t - d_t) + \epsilon_{t+H}$						
Log-Lin	0.0083	0.0205	0.0297	−0.0453	−0.1386	−0.2188
Global	0.0077	0.0197	0.0291	−0.0411	−0.1319	−0.2069
Error	6.86%	4.18%	2.27%	10.14%	5.12%	5.75%
$\sum_h^H (\Delta c_{t+h}) = \alpha + \beta(p_t - d_t) + \epsilon_{t+H}$						
Log-Lin	0.3157	0.3241	0.2596	0.0857	0.2024	0.2507
Global	0.3167	0.3245	0.2582	0.0866	0.2049	0.2530
Error	0.33%	0.11%	0.53%	1.06%	1.23%	0.91%
$\sum_h^H (\Delta d_{t+h}) = \alpha + \beta(p_t - d_t) + \epsilon_{t+H}$						
Log-Lin	0.3990	0.3128	0.2405	0.3834	0.7340	0.8912
Global	0.4017	0.3140	0.2428	0.3888	0.7428	0.9007
Error	0.66%	0.37%	0.96%	1.39%	1.18%	1.05%

Table EII

Annualized Moments and Errors in Schorfheide, Song, and Yaron (2018)

The table reports the mean and the standard deviation of the annualized log price-dividend ratio, the annualized market over the risk-free return, and the risk-free rate. Results obtained by the log-linearized solution and the global solution as well as the relative error of the log-linearization are shown for the median estimates of Schorfheide, Song, and Yaron (2018) (see Table IV) except for the persistence parameters, where the 95% quantile estimates are used (for each set of results, one of the persistence values is increased). All returns and volatilities are shown in percentages, so a value of 1.5 is a 1.5% annualized figure.

	$\mathrm{E}(p_t - d_t)$	$\sigma(p_t - d_t)$	$\mathrm{E}(R_{m,t} - R_{f,t})$	$\mathrm{E}(R_{f,t})$	$\sigma(R_{m,t})$	$\sigma(R_{f,t})$
			Median estimates with $\rho = 0.9995$			
Log-Lin	2.9396	0.7198	6.72	0.83	13.43	0.72
Global	3.4447	0.6758	4.66	1.23	12.49	0.50
Error	14.66 %	6.52%	44.22%	33.64%	7.53%	43.34%
			Median estimates with $\nu_x = 0.9988$			
Log-Lin	2.5670	0.4600	9.75	1.00	20.30	0.77
Global	2.9217	0.3349	7.29	1.17	17.29	0.75
Error	12.14%	37.34%	33.69%	14.56%	17.40%	2.96%
			Median estimates with $\nu_c = 0.9958$			
Log-Lin	3.2775	0.2706	4.60	1.71	16.26	0.57
Global	3.3510	0.2554	4.34	1.71	15.62	0.57
Error	2.19%	5.92 %	5.90%	0.14%	4.10%	0.18%
			Median estimates with $\nu_d = 0.9841$			
Log-Lin	3.2913	0.2715	4.59	1.74	16.80	0.57
Global	3.3641	0.2570	4.34	1.75	16.20	0.57
Error	2.16%	5.66 %	5.76%	0.13%	3.70%	0.17%

(which is the value assumed by the log-linearization), and they are especially large (more than 10^5!) for the second derivative with respect to σ_t^2, which is another reason for the large approximation errors reported in Table I.

Table VI in Section G displays the R^2 statistics and regression coefficients from regressing cumulative log excess returns, consumption growth, and dividend growth on the lagged log price-dividend ratio. Table EI reports the corresponding results for the predictability regressions for the calibration in Bansal and Yaron (2004). We find that the errors of the log-linearized solution are much smaller for the calibration in Bansal and Yaron (2004) than for the more recent calibration in Bansal, Kiku, and Yaron (2012).

Finally, we report results on the model of Schorfheide, Song, and Yaron (2018). Table EII provides additional results on annualized moments and the numerical errors of the log-linearized solutions. These results complement those reported in Table V in Section G. Figure E3 plots the errors in the equity premium as a function of the persistence of long-run growth ρ. Even a small

increase in ρ of 1% from 0.9872 to 0.9971 increases the error in the equity premium from about 5% to almost 35%.

REFERENCES

Atkinson, Kendall E., 1992, A survey of numerical methods for solving nonlinear integral equations, *Journal of Integral Equations and Applications* 4, 15–46.

Bansal, Ravi, Dana Kiku, and Amir Yaron, 2010, Long run risks, the macroeconomy, and asset prices, *American Economic Review* 100, 542–46.

Bansal, Ravi, Dana Kiku, and Amir Yaron, 2012, An empirical evaluation of the long-run risks model for asset prices, *Critical Finance Review* 1, 183–221.

Bansal, Ravi, Dana Kiku, Ivan Shaliastovich, and Amir Yaron, 2014, Volatility, the macroeconomy, and asset prices, *Journal of Finance* 69, 2471–2511.

Bansal, Ravi, and Ivan Shaliastovich, 2013, A long-run risks explanation of predictability puzzles in bond and currency markets, *Review of Financial Studies* 26, 1–33.

Bansal, Ravi, and Amir Yaron, 2004, Risks for the long run: A potential resolution of asset pricing puzzles, *Journal of Finance* 59, 1481–1509.

Beeler, Jason, and John Y. Campbell, 2012, The long-run risks model and aggregate asset prices: An empirical assessment, *Critical Finance Review* 1, 141–182.

Bollerslev, Tim, George Tauchen, and Hao Zhou, 2009, Expected stock returns and variance risk premia, *Review of Financial Studies* 22, 4463–4492.

Bollerslev, Tim, Lai Xu, and Hao Zhou, 2015, Stock return and cash flow predictability: The role of volatility risk, *Journal of Econometrics* 187, 458–471.

Caldara, Dario, Jesús Fernández-Villaverde, Juan F. Rubio-Ramírez, and Wen Yao, 2012, Computing DSGE models with recursive preferences and stochastic volatility, *Review of Economic Dynamics* 15, 188–206.

Campbell, John Y., and Robert J. Shiller, 1988, The dividend-price ratio and expectations of future dividends and discount factors, *Review of Financial Studies* 1, 195–228.

Chen, Yu, Thomas F. Cosimano, and Alex A. Himonas, 2014, On formulating and solving portfolio decision and asset pricing problems, in Karl Schmedders, and Kenneth L. Judd, eds.: *Handbook of Computational Economics, Volume 3* (North-Holland, Elsevier, Amsterdam).

Collard, Fabrice, and Michel Juillard, 2001, Accuracy of stochastic perturbation methods: The case of asset pricing models, *Journal of Economic Dynamics and Control* 25, 979–999.

Collin-Dufresne, Pierre, Michael Johannes, and Lars A. Lochstoer, 2016, Parameter learning in general equilibrium: The asset pricing implications, *American Economic Review* 106, 664–698.

Constantinides, George M., and Anisha Ghosh, 2011, Asset pricing tests with long-run risks in consumption growth, *Review of Asset Pricing Studies* 1, 96–136.

Croce, Mariano M., Martin Lettau, and Sydney C. Ludvigson, 2015, Investor information, long-run risk, and the term structure of equity, *Review of Financial Studies* 28, 706–742.

de Groot, Oliver, 2013, Computing the risky steady state of DSGE models, *Economics Letters* 120, 566–569.

de Groot, Oliver, 2015, Solving asset pricing models with stochastic volatility, *Journal of Economic Dynamics and Control* 52, 308–321.

Dew-Becker, Ian, Stefano Giglio, Anh Le, and Marius Rodriguez, 2017, The price of variance risk, *Journal of Financial Economics* 123, 225–250.

Drechsler, Itamar, and Amir Yaron, 2011, What's vol got to do with it, *Review of Financial Studies* 24, 1–45.

Epstein, Larry G., and Stanley E. Zin, 1989, Substitution, risk aversion and the temporal behavior of consumption and asset returns: A theoretical framework, *Econometrica* 57, 937–969.

Eraker, Bjørn, 2008, Affine general equilibrium models, *Management Science* 54, 2068–2080.

Eraker, Bjørn, and Ivan Shaliastovich, 2008, An equilibrium guide to designing affine pricing models, *Mathematical Finance* 18, 519–543.

Flodén, Martin, 2008, A note on the accuracy of Markov-Chain approximations to highly persistent AR(1) processes, *Economics Letters* 99, 516–520.

Grossman, Sanford J., and Robert J. Shiller, 1981, The determinants of the variability of stock market prices, *American Economic Review* 71, 222–227.

Hansen, Lars P., John Heaton, Junghoon Lee, and Nikolai Roussanov, 2007, Intertemporal substitution and risk aversion, in James J. Heckman, and Edward E. Leamer, eds.: *Handbook of Econometrics, Volume 6* (Elsevier, Amsterdam, London).

Hansen, Lars P., John C. Heaton, and Nan Li, 2008, Consumption strikes back? Measuring long run risk, *Journal of Political Economy* 116, 260–302.

Hansen, Lars P., and Kenneth J. Singleton, 1982, Generalized instrumental variables estimation of nonlinear rational expectations models, *Econometrica* 50, 1269–1286.

Judd, Kenneth L., 1992, Projection methods for solving aggregate growth models, *Journal of Economic Theory* 58, 410–452.

Judd, Kenneth L., 1996, Approximation, perturbation, and projection methods in economic analysis, in Hans M. Amman, David A. Kendrick, and John Rust, eds.: *Handbook of Computational Economics, Volume 1* (Elsevier, Amstedam).

Judd, Kenneth L., 1998, *Numerical Methods in Economics* (The MIT Press: Cambridge, MA).

Juillard, Michel, 2011, Local approximation of DSGE models around the risky steady state, Working paper, IDEAS, available at: http://ideas.repec.org/p/ter/wpaper/0087.html.

Kaltenbrunner, Georg, and Lars A. Lochstoer, 2010, Long-run risk through consumption smoothing, *Review of Financial Studies* 23, 3190–3224.

Kogan, Leonid, and Raman Uppal, 2001, Risk aversion and optimal portfolio policies in partial and general equilibrium economies, Working paper, NBER, available at: http://www.nber.org/papers/w8609.

Koijen, Ralph S. J., Hanno Lustig, Stijn Van Nieuwerburgh, and Adrien Verdelhan, 2010, Long-run risk, the wealth-consumption ratio, and the temporal pricing of risk, *American Economic Review* 100, 552–556.

Mehra, Rajnish, and Edward C. Prescott, 1985, The equity premium: A puzzle, *Journal of Monetary Economics* 15, 145–161.

Meyer-Gohde, Alexander, 2014, Risky linear approximations, Working paper, IDEAS, available at: http://ideas.repec.org/p/hum/wpaper/sfb649dp2014-034.html.

Pohl, Walter, Karl Schmedders, and Ole Wilms, 2015, Higher-order effects in asset-pricing models with long-run risks, Working paper, IDEAS, available at: https://ideas.repec.org/p/red/sed016/306.html.

Schorfheide, Frank, Dongho Song, and Amir Yaron, 2018, Identifying long-run risks: A Bayesian mixed-frequency approach, *Econometrica* (Forthcoming).

Segal, Gill, Ivan Shaliastovich, and Amir Yaron, 2015, Good and bad uncertainty: Macroeconomic and financial market implications, *Journal of Financial Economics* 117, 369–397.

Tallarini, Thomas D., 2000, Risk-sensitive real business cycles, *Journal of Monetary Economics* 45, 507–532.

Tauchen, George, 2011, Stochastic volatility in general equilibrium, *Quarterly Journal of Finance* 1, 707–731.

Tauchen, George, and Robert Hussey, 1991, Quadrature-based methods for obtaining approximate solutions to nonlinear asset pricing models, *Econometrica* 59, 317–396.

Weil, Philippe, 1990, Nonexpected utility in macroeconomics, *Quarterly Journal of Economics* 105, 29–42.

Zhou, Guofu, and Yingzi Zhu, 2015, Macroeconomic volatilities and long-run risks of asset prices, *Management Science* 61, 413–430.

Supporting Information

Additional Supporting Information may be found in the online version of this article at the publisher's website:

Replication code.