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

Walter Pohl
University of Zurich

Karl Schmedders
University of Zurich and
Swiss Finance Institute

Ole Wilms
University of Zurich

December 5, 2016

**Abstract**

This paper presents an analysis of higher-order dynamics in asset pricing models with long-run risk. The numerical errors introduced by the ubiquitous Campbell-Shiller log-linearization approach are economically significant for many plausible choices of parameters and exogenous processes. The resulting errors in the model moments can exceed 75% and may lead to qualitatively wrong model predictions. For example, a common belief about long-run risk models, based on the log-linearization, is that conditional risk premia for long-run consumption risk are constant. The correct solution reveals that, on the contrary, risk premia show considerable time variation and are procyclical.

**Keywords:** Asset pricing, long-run risk, log-linearization, nonlinear dynamics.
**JEL codes:** G11, G12.

*This paper was previously circulated under the title "Higher-Order Dynamics in Asset Pricing Models with Recursive Preferences". We are indebted to Nicole Branger, Lars Hansen, Ken Judd, Martin Lettau, and Malte Schumacher for helpful discussions on the subject. We thank seminar audiences at the University of Zurich, the Becker Friedman Institute, the University of Münster and at various conferences for comments. Karl Schmedders gratefully acknowledges financial support from the Swiss Finance Institute. Ole Wilms gratefully acknowledges financial support from the Zürcher Universitätsverein. Send correspondence to: Karl Schmedders, Department of Business, University of Zurich, Moussonstrasse 15, CH–8044 Zurich, Switzerland; telephone: +41 (0)44 634 3770. E-mail: karl.schmedders@business.uzh.ch.

# 1 Introduction

Bansal and Yaron (2004) introduced an economic mechanism based on long-run risk to generate a high equity premium in asset-pricing models and demonstrated that the Campbell–Shiller log-linearization technique provides a simple method for analyzing such models. This paper presents an analysis of higher-order effects in long-run risk models, which the log-linearization approach disregards by construction. We show that the solutions of models that build on the framework of Bansal and Yaron (2004) are potentially very nonlinear, and that for many plausible choices of parameters and exogenous processes the numerical errors introduced by the ubiquitous log-linearization approach are economically significant. In fact, for very persistent processes, as regularly used in the literature, the approximation errors in model moments can exceed 75%. The linearity assumption can even result in misleading qualitative inferences. For example, conventional wisdom since Bansal and Yaron (2004) is that long-run risk cannot generate time-varying risk premia. We show that once nonlinear effects are accounted for, risk premia can show considerable time variation (and in the wrong direction, compared to the empirical evidence in Fama and French (1989) and Ju and Miao (2012)).

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 evolved and matured over time, researchers added more and more complex elements to their models such as, among others, incomplete markets in form of uninsurable income risks, frictions such as borrowing or collateral constraints, time-varying risk aversion, and heterogeneous expectations. While these additional features had varying degrees of success, recently the new generation of long-run risks models (e.g. Bansal and Yaron (2004) or Hansen, Heaton, and Li (2008)) with their interplay of long-run risks, stochastic volatility, and recursive preferences have had considerably more success in resolving long-standing asset pricing puzzles. The key feature in the model is the combination of a preference for the early resolution of risks, paired 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.

Therefore, a large group of researchers have followed Bansal and Yaron (2004) and used the log-linearization technique to solving asset pricing models with recursive preferences (for example, Segal, Shaliastovich, and Yaron (2015), Bansal, Kiku, and Yaron (2010), Bansal, Kiku, and Yaron (2012), Bollerslev, Tauchen, and Zhou (2009), Kaltenbrunner and Lochstoer (2010), Koijen, Lustig, Van Nieuwerburgh, and Verdelhan (2010), Drechsler and Yaron (2011), Bansal and Shaliastovich (2013), Constantinides and Ghosh (2011), Bansal, Kiku, Shaliastovich, and Yaron (2014) or Beeler and Campbell (2012), among others). Examining this strand of literature, it is difficult to find studies that do not rely on the Campbell-Shiller approach; it has become the standard method for solving asset pricing models with long-run risks.

By its very nature, a log-linear approximation will miss higher-order effects. Can we always safely ignore these higher-order effects? The tendency since the original Bansal and Yaron (2004) model has been towards both higher persistence and greater complexity. This suggests it is time to take stock of whether the Campbell-Shiller approximation is still appropriate. To answer this question, we first determine the drivers that introduce higher-order dynamics to asset pricing models featuring long-run risks and analyze its implications for financial market outcomes. Afterwards, we examine the consequences of ignoring nonlinear dynamics in six recent studies in addition to the Bansal and Yaron (2004) model: the newly calibrated version of Bansal, Kiku, and Yaron (2012), the extensive calibration study of Schorfheide, Song, and Yaron (2016), the volatility-of-volatility model of Bollerslev, Xu, and Zhou (2015), and the works on real and nominal bonds of Koijen, Lustig, Van Nieuwerburgh, and Verdelhan (2010) and Bansal and Shaliastovich (2013).

We show that the higher-order dynamics not captured by the Campbell-Shiller approximation can be large and economically significant. For example, Bansal, Kiku, and Yaron (2012) recalibrate the original Bansal and Yaron (2004) model to have more persistent shocks to stochastic volatility. We find that for this calibration the log-linearization introduces approximation errors as large as 20% for key quantities such as the equity premium or the volatility of price-dividend ratio. Schorfheide, Song, and Yaron (2016) perform a Bayesian estimation of the model using the same approximation, and find evidence for a higher persistence for long-run risk. In this case, we find approximation errors as large as 75% for some key model moments. In general, highly persistent processes lead to solutions that are highly nonlinear, and thus to economically relevant approximation errors. Log-linearization can even introduce errors in qualitative conclusions. It wrongly predicts constant risk premia for long-run consumption risks and under high persistence, log-linearization can actually invert the slope of the yield curve in the nominal bond models of Bansal and Shaliastovich (2013) and Koijen, Lustig, Van Nieuwerburgh, and Verdelhan (2010).

Summarizing, by construction, log-linearizing the model as it is commonly done in the asset pricing literature misses higher-order dynamics. If the driving factors of the economy are of low persistence or the risk aversion of the representative agent is low, these dynamics will have a negligible influence on equilibrium outcomes. However, the combination of highly persistent processes, together with recursive preferences and a risk aversion significantly larger than one—features which are required by the long-run risk model to be consistent with the financial market data—can introduce strong nonlinear dynamics to the model. We show that these errors have a strong impact on key financial statistics in some recent asset pricing studies and introduce a bias to the model parameters when it comes to the estimation or calibration of the model. Therefore, researchers should not rely on log-linearized solutions in the future but use more sophisticated methods that can account for higher-order dynamics.

The paper is organized as follows. Section 2 describes the general model framework that is used throughout the paper. In Section 3, we determine the drivers that introduce higher-order dynamics to asset pricing models with long-run risks. Subsequently, we show how they qualitatively and quantitatively influence model outcomes. Section 4 analyzes the implications of higher-order effects for six recent asset pricing studies. Section 5 concludes.

# 2 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 as

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

In this parametrization, $C_t$ is consumption, $\delta$ is the time discount factor, $\gamma$ determines the level of relative risk aversion, and $\theta = \frac{1-\gamma}{1-\frac{1}{\psi}}$, where $\psi$ is the elasticity of intertemporal substitution (EIS). The parameters $\gamma$ and $\psi$ are required to satisfy $0 < \gamma, \psi$, and $\psi \neq 1$. For $\theta = 1$ the agent has standard 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

$$E_t\left[M_{t+1}R_{i,t+1}\right] = 1, \tag{2}$$

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

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

Epstein and Zin (1989) show that the (unobserved) value of the 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 in turn permits expressing $M_{t+1}$ 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 $R_{w,t+1} = \frac{W_{t+1}}{W_t - C_t}$. As 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

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

Throughout the paper we consider different setups for the specification of log consumption growth, $\Delta c_{t+1}$. 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, $\sigma_t$, and a long-run growth process, $x_t$, affecting both log consumption growth and log dividend growth, $\Delta d_{t+1}$. Specifically, the four processes are given as follows:

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

We elaborate in more detail on the properties and key features of the model in the beginning of Section 3. 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 solve long-run risk models used in the finance literature is to log-linearize the model, see Segal, Shaliastovich, and Yaron (2015), Bansal, Kiku, and Yaron (2010), Bansal, Kiku, and Yaron (2012), Bollerslev, Tauchen, and Zhou (2009), Kaltenbrunner and Lochstoer (2010), Koijen, Lustig, Van Nieuwerburgh, and Verdelhan (2010), Drechsler and Yaron (2011), Bansal and Shaliastovich (2013), Constantinides and Ghosh (2011), Bansal, Kiku, Shaliastovich, and Yaron (2014) or Beeler and Campbell (2012), among others. However, log-linearization misses, by construction, the influence of higher order dynamics — that is, the approach does not attempt to approximate nonlinear features of the exact solution. But what if these features matter qualitatively or quantitatively for equilibrium outcomes? Does log-linearization still deliver sufficiently accurate approximations of the exact solution?

We address these critical issues in this paper. For this task we need an alternate solution method that accurately accounts for higher-order dynamics and yields robust solutions. In the body of the paper, we use projection methods (Judd (1992)), which capture these nonlinear effects, and are known to converge to the true solution (Atkinson (1992)). For a stochastic growth model with Epstein-Zin utility, Caldara, Fernandez-Villaverde, Rubio-Ramirez, and Yao (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 provide a discussion of the pros and cons of different computational methods for asset pricing models, as well as a detailed description of the log-linear and projection solution methods.

# 3 Higher-Order Dynamics in Long-Run Risk Models

We begin our analysis by determining the factors that influence the higher-order dynamics in long-run risk models. For this, we use the standard long-run risk framework of Bansal and Yaron (2004) (see the four processes in (7)). Key features of the model are the highly persistent state processes for the long-run growth rate, $x_t$, and the stochastic volatility $\sigma_t^2$. As we demonstrate below, the model requires the persistence parameters $\rho$ and $\nu$ to be very close to 1, as otherwise the model predictions do not match the data. Combining these highly persistent processes with a preference for the early resolution of risks ($\gamma > \frac{1}{\psi}$), Bansal and Yaron (2004) are able to explain many puzzling features on financial markets like the high equity premium together with a low risk-free rate or the volatility of the market return, the risk-free rate, and the price-dividend ratio. As pointed out by Beeler and Campbell (2012), the original calibration of Bansal and Yaron (2004) implies, counterfactually, that consumption and dividend growth are highly predictable from stock prices. In reply to this criticism, Bansal, Kiku, and Yaron (2012) recalibrate the model to be more consistent with the data.[1] We use this calibration for a first analysis of the impact of higher-order dynamics on quantitative

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

and qualitative equilibrium outcomes in the long-run risk model, see Sections 3.1 and 3.2. In Section 4 we analyze the influence of higher-order dynamics on model predictions in five additional studies.

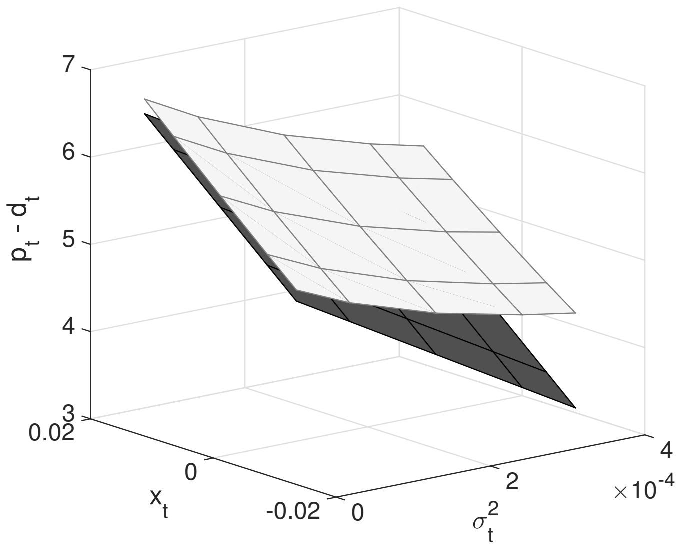
## 3.1 Quantitative Model Predictions

In Figure 1 we show 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 $\sigma_t^2$. The dark gray area shows the log-linearized solution as used in Bansal, Kiku, and Yaron (2012) and the transparent gray area shows the correct solution obtained by the projection approach.[2] We observe that 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 an analysis based on the correctly solved model does. Also, the larger derivatives of log price-dividend ratio with respect to the two state variables implies a larger volatility of the price-dividend ratio. As both, the equity premium and the volatility of the price-dividend ratio, are key quantities that the long-run risk model tries to explain, the systematical underestimation of the log price-dividend ratio is an unpleasant property of the solution method. In practice, of course, it is only relevant whether the difference matters quantitatively and qualitatively for the empirical predictions of the model. We document in this paper that they do. Moreover, the more persistent the state processes of the economy, the larger are the errors introduced by missing higher-order effects. As highly persistent state processes are required by the long-run risk model to produce model outcomes to be consistent with the data, using log-linearized solutions introduces large errors to the model solution that in turn lead to incorrect model predictions.

Figure 2 shows errors in the log price-dividend ratio introduced by the log-linearization for different values for the persistences of long-run risk $\rho$ and stochastic volatility $\nu$. 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 observe that the errors increase dramatically with the persistence of the state processes $\rho$ and $\nu$ and the errors in the volatility of the log price-dividend ratio become as large as 50% for large values of $\rho$ and $\nu$. To demonstrate that these are not artificially constructed calibrations to obtain large errors, we also show the point estimates for $\rho$ and $\nu$ used in the studies of Bansal and

[2] A formal analysis of the accuracy of the projection approach is conducted in Appendix B. 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. By this approach we make sure, that we capture the higher-order dynamics introduced by the tails of the state processes.

Figure 1: Log Price-Dividend Ratio in the Long-run Risk Model




The graph shows the log-linearized solution for the log price-dividend ratio (dark grey area) as well as the correct solution (transparent gray area) as a function of the states $x_t$ and $\sigma_t^2$. Parameters: $\delta = 0.9989, \mu = 0.0015, \bar{\sigma} = 0.0072, \phi_x = 0.038, \gamma = 10, \psi = 1.5, \mu_d = 0.0015, \Phi = 2.5, \phi_d = 5.96, \phi_{d,c} = 2.6$.

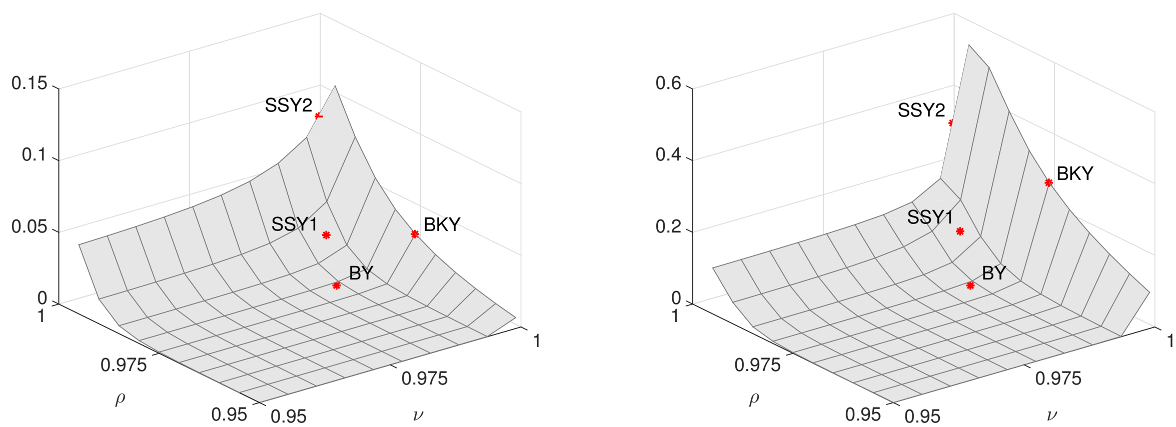
Yaron (2004) (BY), Bansal, Kiku, and Yaron (2012) (BKY), and Schorfheide, Song, and Yaron (2016). For the estimation study of Schorfheide, Song, and Yaron (2016), we show the median estimates (SSY1) as well as the 95% estimates (SSY2) to demonstrates the range of parameters and hence errors, that are included within the estimation procedure. We observe that the parameters used in the study—except for the study of BY—are in the area where approximation errors are large and significant. For example in the study of Bansal, Kiku, and Yaron (2012) the volatility of the log price-dividend ratio is overestimated by more than 25%.[3]

While Figure 2 shows errors in the moments of the monthly log price-dividend ratio, asset pricing models like the kinds of Bansal and Yaron (2004) or Bansal, Kiku, and Yaron (2012) usually try to match annualized market outcomes. Therefore, in Figures 3 and 4 we show the annualized equity premium and the annual volatility of the log price-dividend ratio, respectively, obtained by the log-linearized solution as well as the correct solution as a function of the risk aversion, $\gamma$, and the intertemporal elasticity of substitution, $\psi$, in the first row and the serial correlations in the long-run risk channel, $\rho$, and the stochastic volatility channel,

[3]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 rather serves to demonstrate the potential errors introduces when using log-linearization to solve highly persistent models. A full evaluation of the different studies is conducted in Section 4.

Figure 2: Errors in the Monthly Log Price-Dividend Ratio in the Long-Run Risk Model

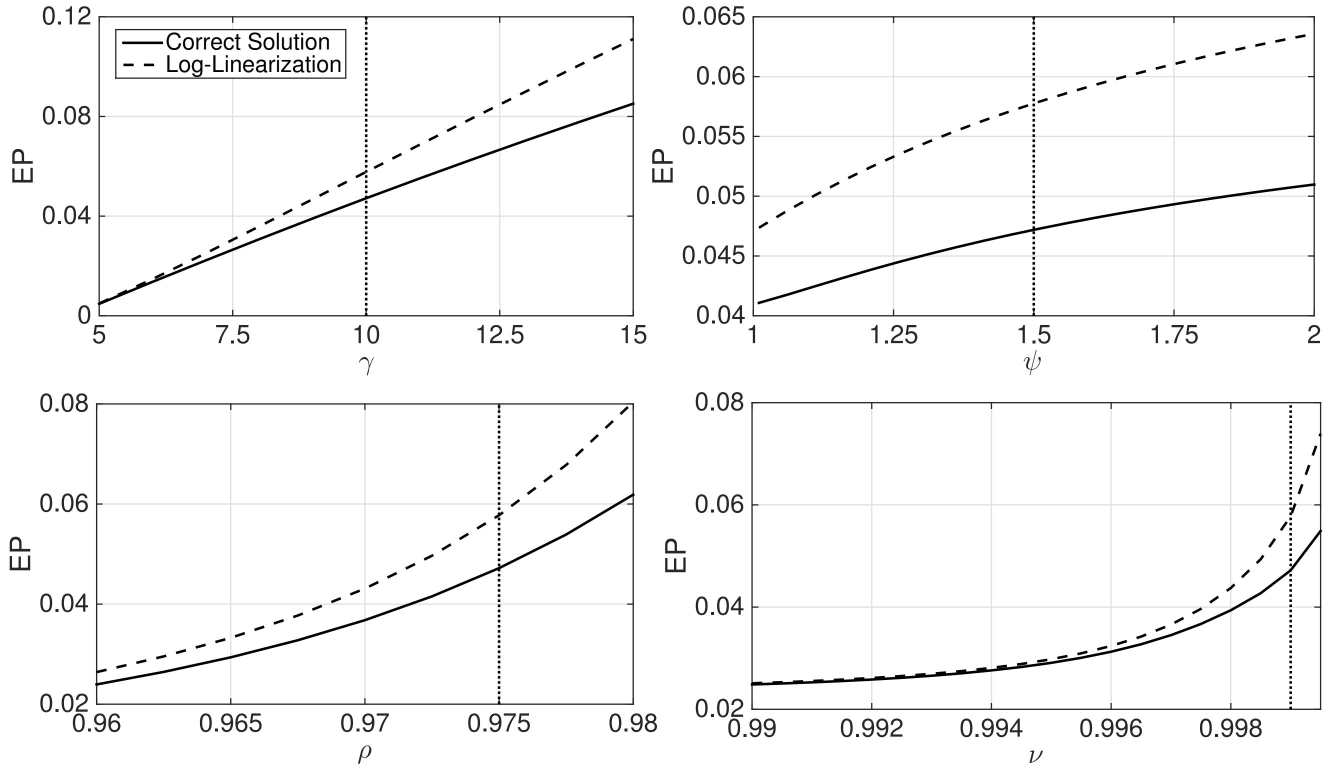
(a) $E(p_t - d_t)$ (b) $\sigma(p_t - d_t)$




The graph shows 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 for the persistences of long-run risk $\rho$ and stochastic volatility $\nu$. The right panel shows the corresponding errors in the unconditional standard deviation of the log price-dividend ratio. Parameters: $\delta = 0.9989, \mu = 0.0015, \bar{\sigma} = 0.0072, \phi_x = 0.038, \gamma = 10, \psi = 1.5, \mu_d = 0.0015, \Phi = 2.5, \phi_d = 5.96, \phi_{d,c} = 2.6$. It also shows 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 (2016) for their median parameter estimates (SSY1) and their 95% estimates (SSY2). (See Table 1 for all the parameters used in the studies.)

$\nu$, in the second row. As mentioned before, a key property of the long-run risk model is the preference for the early resolution of risks, which is obtained by setting $\gamma > \frac{1}{\psi}$. Hence, either increasing $\gamma$ or $\psi$ implies a stronger preference for the early resolution of risks and hence amplifies the model predictions.

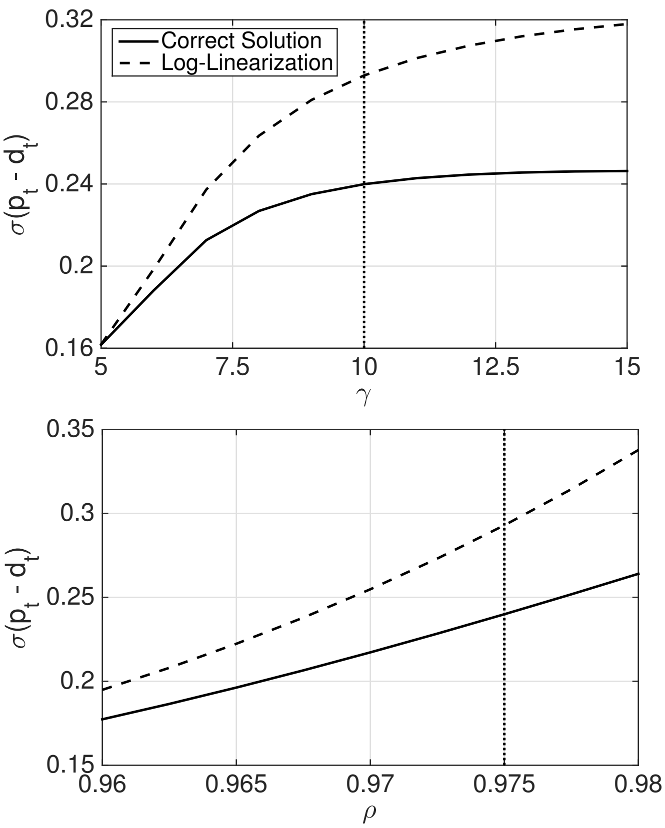
Figure 3: Sensitivity of the Approximation Errors for the Annualized Equity Premium in the Long-Run Risks Model




The figure shows the annual equity premium obtained by the log-linearization (dashed line) as well as the correct solution (solid line) as a function of the model parameters $\gamma, \psi, \rho$ and $\nu$, respectively, assuming that the other parameters are kept constant. In each panel, the dotted vertical line denotes the estimate used in original calibration. Parameters: $\delta = 0.9989, \mu = 0.0015, \bar{\sigma} = 0.0072, \phi_x = 0.038, \gamma = 10, \psi = 1.5, \mu_d = 0.0015, \Phi = 2.5, \phi_d = 5.96, \phi_{d,c} = 2.6$.

We find that, for this particular calibration, for a risk aversion of approximately 5, the log-linearized solution basically coincides with the solution from the projection approach, which suggests that a linear solution gives a reasonable approximation to the model. However, for this calibration also the implied model moments collapse with an equity premium below 1% 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 for both,

Figure 4: Sensitivity of the Approximation Errors for the Annual Volatility of the Price-Dividend Ratio in the Long-Run Risks Model




The figure shows the annual volatility of the log price-dividend ratio obtained by the log-linearization (dashed line) as well as the correct solution (solid line) as a function of the model parameters $\gamma, \psi, \rho$ and $\nu$, respectively, assuming that the other parameters are kept constant. In each panel, the dotted vertical line denotes the estimate used in original calibration. Parameters: $\delta = 0.9989, \mu = 0.0015, \bar{\sigma} = 0.0072, \phi_x = 0.038, \gamma = 10, \psi = 1.5, \mu_d = 0.0015, \Phi = 2.5, \phi_d = 5.96, \phi_{d,c} = 2.6$.

the long-run risk and the stochastic volatility. We observe 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 basis points (vertical dotted line). By slightly increasing $\rho$ to 0.98, however, the difference doubles with an overestimation of 200 basis points. For the persistence in the stochastic volatility, $\nu$, even a change of 0.0005, from 0.999 to 0.9995, increases the overestimation to 200 basis points. The figures also show that lowering the persistence parameters significantly decreases approximation errors. For example for $\nu = 0.99$ the approximation error becomes close to zero. However, for this calibration also the implied model moments collapse. Hence, as the model requires highly persistent state processes as well as a large degree of risk aversion and an intertemporal elasticity of substitution exceeding one, nonlinear dynamics in the model are present and strong. This model feature, in turn, renders the use of log-linearization a poor solution method for solving the model since it implies large approximation errors and even tiny changes in the model parameters can strongly affect these errors.

The log-linearization approach is also appealing for its apparent ability to draw qualitative economic conclusions. For example, in the context of the long-run risk model, the approximate closed-form solutions serve as a powerful tool to analyze qualitative dependencies of model inputs and outputs (see Bansal and Yaron (2004)). But, as we demonstrate next, these clear-cut model predictions can be highly misleading. Bansal and Yaron (2004) write that "because of our assumption of a constant $\sigma$, the conditional risk premium on the market portfolio is constant, and so is its conditional volatility. Hence, the ratio of the two, namely the Sharpe ratio, is also constant. In order to address issues that pertain to time-varying risk premia and predictability of risk premia, we augment our model in the next section and introduce time-varying economic uncertainty." This conclusion that risk premia for the long-run consumption risk channel are constant is a pure artifact of the log-linearized solution. The correct solution reveals that, on the contrary, risk premia show considerable time variation and move, compared to the empirical results in Fama and French (1989) and Ju and Miao (2012), in the wrong direction.

## 3.2 Qualitative Model Predictions

Part of the appeal of log-linearization is that the approximate closed form seemingly lends itself to tractable analysis of the economic implications of a model. For example, Bansal and Yaron (2004) show that the time-$t$ expected risk premium $E_t(r^m_{t+1} - r^f_{t+1})$ in the long-run risk model (7) with constant volatility ($\sigma_t = \bar{\sigma} \ \forall \ t$) is given by

$$E_t(r^m_{t+1} - r^f_{t+1}) = (\beta_{m,e}\lambda_{m,e} + \beta_{m,\nu}\lambda_{m,\nu})\bar{\sigma}^2 - 0.5\,\mathrm{var}_t(r_{m,t}), \tag{8}$$

with[4]

$$\beta_{m,e} = [\kappa_{m,1}(\Phi - \frac{1}{\psi})\frac{\phi_x}{1-\kappa_{m,1}\rho}], \quad \beta_{m,\nu} = \phi_{d,c}, \quad \lambda_{m,e} = (1-\theta)(\kappa_{w,1}(1-\frac{1}{\psi})\frac{\phi_x}{1-\kappa_{w,1}\rho})$$

$$\lambda_{m,\nu} = \gamma, \quad \mathrm{var}_t(r_{m,t}) = [\beta_{m,e}^2 + \phi_d^2 + \phi_{d,c}^2]\bar{\sigma}^2.$$

This expression has a simple economic consequence: the risk premium is constant across the exogenous states of nature, since all expressions are independent of $x_t$. This fact, in turn, implies that long-run risk alone cannot generate a time-varying risk premia.

We show that this conclusion arises solely from neglecting nonlinear effects. Figure 5 displays the expected risk premium as a function of the state $x_t$. Results are shown for three different values of the persistence parameter $\rho$. The dashed lines show the (constant) premia predicted by the log-linearization. Solid lines show the correct solutions obtained by the projection approach. We observe that once nonlinear effects are included, the risk premia does indeed depend on the state, $x_t$, and the dependence becomes stronger as $\rho$ becomes larger. Hence, once the nonlinear dynamics of the model are correctly accounted for, risk premia can be seen to vary over time. The conclusion of a constant premium is thus an artifact of the log-linearization. Even more surprisingly, we observe that expected risk premia are high for large values of $x_t$ and low for small values of $x_t$, that is, they are procyclical. This result is exactly the opposite of the empirical results documented by Fama and French (1989)[5] Ju and Miao (2012, Table I.C.) confirm several stylized facts mentioned by Fama and French (1989) and find a negative correlation between consumption growth and 1-year excess returns.

In sum, the log-linearization not only introduces significant quantitative errors to the model outcomes, but it even leads to qualitatively wrong model predictions that potentially mislead researchers. For example, Bansal and Yaron (2004) explain that they introduce time-varying exogenous uncertainty to circumvent the otherwise zero time variation and perfect predictability of risk premia. The correct solution reveals that risk premia vary over time even in a model with only long-run risk.

# 4 Higher-Order Dynamics in Six Asset Pricing Models

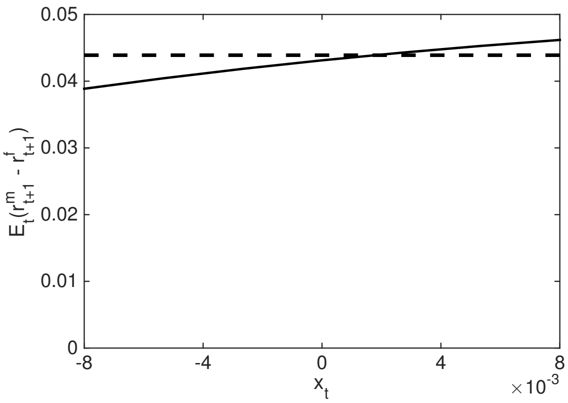
In this section we compare the implications of using log-linearized solutions for the prediction of economically relevant quantities in asset pricing models. Specifically, we perform this

[4]Here $\kappa_{m,1}$ and $\kappa_{w,1}$ are the linearization coefficients, see Appendix A.1.

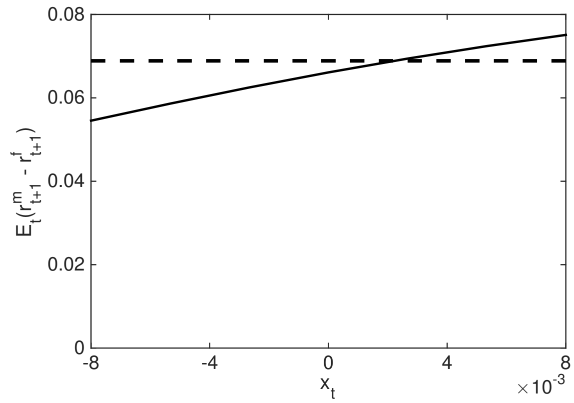
[5]Fama and French (1989) find that expected excess returns are "inversely related to business conditions." They offer consumption smoothing as the main reason in favor of their findings. When incomes are high, households both consume more and save more; and higher savings lead to lower expected returns. When incomes are low, households reduce both consumption and savings; and lower savings are expected to result in higher returns.

Figure 5: Expected Risk Premia in the Long-Run Risk Model

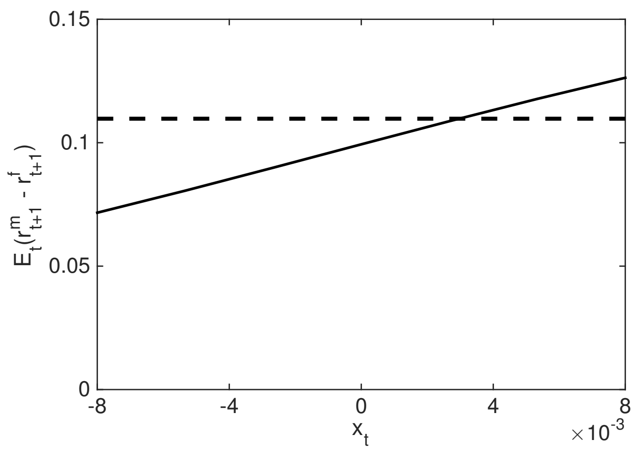
(a) $\rho = 0.98$




(b) $\rho = 0.985$




(c) $\rho = 0.99$




The graphs show the conditional expected risk premium $E_t(r^m_{t+1} - r^f_{t+1})$ as a function of the state $x_t$ in the model of Bansal and Yaron (2004). The dashed line shows the constant log-linearized solution and the solid line shows the correct solution. Parameters are from Bansal and Yaron (2004) and given by $\delta = 0.998, \mu = 0.0015, \sigma_t = \bar{\sigma} = 0.0078 \ \forall \ t, \phi_x = 0.044, \gamma = 10, \psi = 1.5, \mu_d = 0.0015, \Phi = 3, \phi_d = 4.5, \phi_{d,c} = 0$.

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 (2016), the volatility-of-volatility model of Bollerslev, Xu, and Zhou (2015), and the two studies of real and nominal bonds of Koijen, Lustig, Van Nieuwerburgh, and Verdelhan (2010) and Bansal and Shaliastovich (2013). Common to all 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, in order to determine a reasonable calibration of the model it is essential to solve the model without significant errors in the approximation of those key statistics since such errors could potentially bias the calibration or estimation.

In the previous section we have seen that depending on the persistence of the state processes, the log-linearization approach produces sizable approximation errors in the long-run risk model of Bansal and Yaron (2004). Now we demonstrate that using log-linearized solutions has a strong impact on the predictions of these six asset pricing models.

## 4.1 Six Model Specifications

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

$$\begin{aligned}
\Delta c_{t+1} &= \mu_c + x_t + \phi_c \sigma_{c,t} \eta_{c,t+1} \\
x_{t+1} &= \rho x_t + \phi_x \sigma_{x,t} \eta_{x,t+1} \\
\Delta d_{t+1} &= \mu_d + \Phi x_t + \phi_d \sigma_{d,t} \eta_{d,t+1} + \phi_{d,c} \sigma_{c,t} \eta_{c,t+1} \\
&\eta_{c,t+1}, \eta_{x,t+1}, \eta_{d,t+1} \sim i.i.d.\ N(0,1).
\end{aligned} \tag{9}$$

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

### 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 there is a single volatility process that drives uncertainty in the economy, $\sigma_{c,t} = \sigma_{x,t} = \sigma_{d,t} = \sigma_t$ with

$$\sigma_{t+1}^2 = \bar{\sigma}^2(1-\nu) + \nu \sigma_t^2 + \phi_\sigma \omega_{t+1} \qquad \omega_{t+1} \sim i.i.d.\ N(0,1). \tag{10}$$

Recall that this is the model setup that we have employed above to examine the approximation errors of the log-linearized solution, see also (7).

### The estimation study of Schorfheide, Song, and Yaron (2016)

Schorfheide, Song, and Yaron (2016) relax the assumption of a single volatility process and allow for three separate volatility processes for consumption, dividends, and long-run risks.[6] The two 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. Therefore, a third process is required to model the stochastic volatility of dividends. Schorfheide, Song, and Yaron (2016) assume that the logarithm of the volatility process is normal to ensure that the standard deviation of the shocks remains positive,

$$
\begin{aligned}
\sigma_{i,t} &= \varphi_i \bar{\sigma} \exp(h_{i,t}) \\
h_{i,t+1} &= \nu_i h_{i,t} + \sigma_{h_i} \sqrt{1-\nu_i^2}\,\omega_{i,t+1}, \qquad i \in \{c, x, d\} \\
\omega_{i,t+1} &\sim i.i.d.\ N(0,1).
\end{aligned}
\tag{11}
$$

In order to derive analytical solutions for the log-linearization coefficients that are needed for their estimation study, Schorfheide, Song, and Yaron (2016) use a linear approximation of the volatility dynamics that follows Gaussian dynamics,

$$
\sigma_{i,t}^2 \approx 2(\varphi_i \bar{\sigma})^2 h_{i,t} + (\varphi_i \bar{\sigma})^2 \tag{12}
$$

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 $\bar{\sigma}_i = \varphi_i \bar{\sigma}$.[7]

### 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 on financial markets such as a high equity risk premium together with a low risk-free rate, volatile price dynamics, or the predictability of stock returns can be explained. However, the most recent research has gone one 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 also suggested a possible solution for this puzzle by adding

[6] Schorfheide, Song, and Yaron (2016) 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 use the setting $\rho_\lambda = \sigma_\lambda = 0$ in their model specification.

[7] We proceed in the same way as Schorfheide, Song, and Yaron (2016) by solving the model using the linearized version of the volatility dynamics to obtain quasi-closed-form solutions for the linearization coefficients; for the inference of moments we use the original specification to ensure that the volatility of the model stays positive.

time-varying volatility of volatility (vol-of-vol) to the model, see, for example, Bollerslev, Tauchen, and Zhou (2009), Tauchen (2011), Drechsler and Yaron (2011), Bollerslev, Xu, and Zhou (2015) or Dew-Becker, Giglio, Le, and Rodriguez (2015). Bollerslev, Xu, and Zhou (2015) consider a slight variation of the long-run risk factor compared to the baseline model (7) where the vol-of-vol factor $q_t$ drives the volatility,[8]

$$\begin{aligned} \sigma_{t+1}^2 &= \bar{\sigma}^2(1-\nu) + \nu\sigma_t^2 + \phi_\sigma\sqrt{q_t}\omega_{\sigma,t+1} \\ q_{t+1} &= \mu_q(1-\rho_q) + \rho_q q_t + \phi_q\sqrt{q_t}\omega_{q,t+1} \\ x_{t+1} &= \rho x_t + \phi_x\sqrt{q_t}\eta_{x,t+1} \\ &\eta_{x,t+1}, \omega_{\sigma,t+1}, \omega_{q,t+1} \sim i.i.d.\ N(0,1). \end{aligned} \tag{13}$$

The vol-of-vol factor $q_t$ follows a square root process. This process specification has also been used, for example, in Tauchen (2011) or the seminal work on volatility of volatility 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 $q_t$ becomes negative. This problem is usually circumvented by either assuming a reflecting boundary at zero or by truncation to ensure positivity. However, for a simple computation of model solutions, the assumption of a non-truncated distribution for the log-linearization is commonly used. (For example Bansal and Yaron (2004) use the non-truncated distribution to compute the log-linearized solutions but simply 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 C we analyze in more detail how the square-root process specification and the issue of complexity affects the log-linearized solution. In particular we find that for the calibration in Bollerslev, Tauchen, and Zhou (2009) equilibrium model solutions are not real numbers but instead 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. Therefore, we concentrate on this calibration in the main text.

### The model of Koijen, Lustig, Van Nieuwerburgh, and Verdelhan (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, Lustig, Van Nieuwerburgh, and Verdelhan (2010). They add inflation, $\pi_t$, with a stochastic growth rate, $x_{\pi,t}$, to the

[8] Drechsler and Yaron (2011) use a similar model where the volatility of $x_t$ is driven by $\sigma_t$ instead of $q_t$; see their 2007 working paper version. However, Bollerslev, Xu, and Zhou (2015) provide evidence for 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.

standard model (7) and price nominal bonds.[9]

$$
\begin{aligned}
\pi_{t+1} &= \mu_\pi + x_{\pi,t} + \phi_{\pi,c}\sigma_{c,t}\eta_{c,t+1} + \phi_{\pi,x}\sigma_{x,t}\eta_{x,t+1} + \sigma_\pi\eta_{\pi,t+1} \\
x_{\pi,t+1} &= \mu_{x_\pi}(1-\rho_\pi) + \rho_\pi x_{\pi,t} + \rho_{\pi,x}x_t \\
&\qquad + \phi_{x_\pi,c}\sigma_{c,t}\eta_{c,t+1} + \phi_{x_\pi,x}\sigma_{x,t}\eta_{x,t+1} + \sigma_{x_\pi}\eta_{\pi,t+1} \\
\eta_{\pi,t+1} &\sim i.i.d.\ N(0,1).
\end{aligned}
\tag{14}
$$

Koijen, Lustig, Van Nieuwerburgh, and Verdelhan (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.

**The model of Bansal and Shaliastovich (2013)**

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

$$
\begin{aligned}
\pi_{t+1} &= \mu_\pi + x_{\pi,t} + \sigma_\pi\eta_{\pi,t+1} \\
x_{\pi,t+1} &= \rho_\pi x_{\pi,t} + \sigma_{\pi,t}e_{\pi,t+1} \\
x_{t+1} &= \rho x_t + \rho_{x\pi}x_{\pi,t} + \sigma^x_t e_{x,t+1} \\
&\eta_{\pi,t+1}, e_{\pi,t+1}, e_{x,t+1} \sim i.i.d.\ N(0,1).
\end{aligned}
\tag{15}
$$

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 1 lists the parameter values of the six studies.[10] While the parameters in Bansal

[9] 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 $x_{\pi,t}$.

[10] For the model of Bollerslev, Xu, and Zhou (2015), we use the parameter estimates in the study for $\rho, \nu$ and

Table 1: Model Parameters

| | | BY (2004) | BKY (2012) | SSY (2014) | BS (2013) | KLVV (2010) | BXZ (2015) |
|---|---|---|---|---|---|---|---|
| Preferences | $\gamma$ | 10 | 10 | 8.598 | 20.90 | 8 | 10 |
| | $\psi$ | 1.5 | 1.5 | 1.935 | 1.81 | 1.5 | 1.5 |
| | $\delta$ | 0.998 | 0.9989 | 0.999 | 0.994 | 0.9987 | 0.999 |
| Consumption | $\mu_c$ | 0.0015 | 0.0015 | 0.0016 | 0.0049 | 0.0016 | 0.0015 |
| | $\phi_c$ | 1 | 1 | 1 | 1 | 1 | 0.00546 |
| | $\rho$ | 0.979 | 0.975 | 0.9872 (0.9995) | 0.81 | 0.991 | 0.988 |
| | $\phi_x$ | 0.044 | 0.038 | 1 | 1 | 1 | 3.12e–4 |
| | $\rho_{x\pi}$ | – | – | – | –0.047 | 0 | – |
| Volatility | $\nu_c$ | 0.987 | 0.999 | 0.9914 (0.9958) | 0 | 0.85 | 0.64 |
| | $\nu_x$ | – | – | 0.9943 (0.9988) | 0.994 | 0.996 | – |
| | $\nu_d$ | – | – | 0.9665 (0.9841) | – | - | – |
| | $\nu_\pi$ | – | – | – | 0.979 | - | – |
| | $\phi_{\sigma_c}$ | 2.3e–6 | 2.8e–6 | 1.9e–6 | 0 | 1.15e–6 | 1 |
| | $\phi_{\sigma_x}$ | – | – | 6.9e–11 | 1.85e–7 | 4.19e–9 | – |
| | $\phi_{\sigma_d}$ | – | – | 1.0e–4 | – | - | – |
| | $\phi_{\sigma_\pi}$ | – | – | – | 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 | $\mu_d$ | 0.0015 | 0.0015 | 0.001 | – | 0.0015 | 0.0015 |
| | $\Phi$ | 3.0 | 2.5 | 4.147 | – | 1.5 | 3.0 |
| | $\phi_d$ | 4.5 | 5.96 | 1 | – | 6 | 0.0246 |
| | $\phi_{d,c}$ | 0 | 2.6 | 1.544 | – | 0.6 | 0 |
| Inflation | $\mu_\pi$ | – | – | – | 0.0090 | 0 | – |
| | $\mu_{x_\pi}$ | – | – | – | 0 | 0.0032 | – |
| | $\sigma_\pi$ | – | – | – | 0.0055 | 0.0035 | – |
| | $\sigma_{x_\pi}$ | – | – | – | 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 | – |
| | $\rho_\pi$ | – | – | – | 0.988 | 0.83 | – |
| | $\rho_{\pi,x}$ | – | – | – | 0 | -0.35 | – |
| Vol–of–Vol | $\mu_q$ | – | – | – | – | | 0.211 |
| | $\phi_q$ | – | – | – | – | | 0.632 |
| | $\rho_q$ | – | – | – | – | | 0.46 |

The table reports parameter values for the studies of Bansal and Yaron (2004) (values taken from Table 4 on page 1495), Bansal, Kiku, and Yaron (2012) (values taken from Table 1 on page 193), Schorfheide, Song, and Yaron (2016) (the median estimates from Table 5 on page 25 are shown as well as the 95% for the persistence parameters in parenthesis), Bansal and Shaliastovich (2013) (values taken from Table 4 on page 22), Koijen, Lustig, Van Nieuwerburgh, and Verdelhan (2010) (values taken from Table 1 on page 19 in the online appendix), and Bollerslev, Xu, and Zhou (2015) (values are provided in Section 3.3 starting on page 464).

and Yaron (2004) and Bansal, Kiku, and Yaron (2012) are calibrated, Schorfheide, Song, and Yaron (2016), 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 new calibration of Bansal, Kiku, and Yaron (2012) is that in the new calibration the persistence of the volatility shock, $\nu_c$, is higher and that shocks to dividends are correlated with short-run shocks to consumption growth ($\phi_{d,c} = 2.6$ in the new 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 risks 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 for 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 (2016) provides further evidence for highly persistent state processes and hence, potentially, large nonlinear dynamics. It reports median estimates for the persistence of long-run risks and for stochastic volatility of consumption, long-run risks, and dividends of 0.9872 and of 0.9914, 0.9943, and 0.9665, respectively. These are the median estimates from the Bayesian estimation study. For the 95% estimates they report values as large as 0.9995, 0.9958, 0.9988, and 0.9841 (values are provided in parentheses in Table 1). 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 3.1. Hence, the existence of solutions for the full range of parameters used in Schorfheide, Song, and Yaron (2016) is not necessarily satisfied. Since in the present study we are rather interested in the influence of higher-order dynamics than the existence of solutions, we focus on the parameter range where the model solutions are still well behaved. This is a rather conservative approach, as approximation errors increase the more extreme the values

$\rho_q$. 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).

in the calibration, recall Figures 3 and 4.

## 4.2 Moments and Errors

Table 2 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 return, and the volatilities of the excess return and the risk-free rate.[11] The table reports these statistics for the log-linearization solutions and the correct solutions obtained by the projection approach; in addition, it states the relative errors induced by the linearization.

We observe that the log-linearization does a reasonably good job for the parameters in Bansal and Yaron (2004) with a maximal error of 1.83% for the equity premium. For the parameter set of Bansal, Kiku, and Yaron (2012) the results are considerably worse. The log-linearization overstates the equity premium by more than 100 basis points. Also, it predicts a volatility of the log price-dividend ratio of 0.2931 instead of 0.2402; these values correspond to relative errors 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.

For the model of Schorfheide, Song, and Yaron (2016), we find that approximation errors are in a reasonable range for the median parameter estimates (results (1)). However, using the 95% estimates for the persistence parameters (results (2)), approximation errors increase dramatically with an overestimation of the equity premium of more than 75%. As we have shown in Section 3.1, it is the interplay of the highly persistent state processes that introduces the substantial nonlinearities to the model solutions; as a result, even a slight increase in the persistence parameter of the long-run risk channel can dramatically increase the approximation errors of the log-linearized solution. Hence, using the log-linearized solution to estimate models featuring highly persistent state processes can potentially introduce a large bias to the implied model moments, which may bias estimation results for the model parameters.

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

[11] 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 data. Beeler and Campbell (2012) provide a detailed description of how to compute the annual moments from the monthly observations. A significant issue in the model is that the variance process $\sigma_t^2$ 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 choose the interval to be slightly larger than the maximum observation range of the long simulations. As in the previous section, we increase the polynomial degree until the coefficients of the highest-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.

Table 2: Annualized Moments and Errors

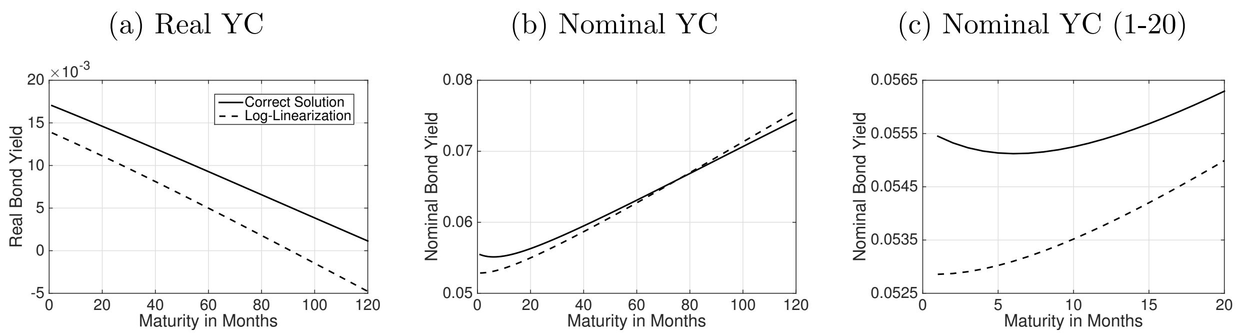
| | $E(p_t - d_t)$ | $\sigma(p_t - d_t)$ | $E\left(r_t^m - r_t^f\right)$ | $E\left(r_t^f\right)$ | $\sigma(r_t^m)$ | $\sigma\left(r_t^f\right)$ |
|---|---|---|---|---|---|---|
| **Bansal and Yaron (2004)** | | | | | | |
| Log-Lin | 3.0105 | 0.1969 | 4.16 | 2.58 | 16.85 | 1.31 |
| Correct | 3.0379 | 0.1946 | 4.09 | 2.58 | 16.76 | 1.31 |
| Error | 0.90 % | 1.17% | 1.83% | 0.06% | 0.54% | 0.08% |
| **Bansal, Kiku, and Yaron (2012)** | | | | | | |
| Log-Lin | 3.0414 | 0.2931 | 5.56 | 0.98 | 21.45 | 1.29 |
| Projection | 3.2370 | 0.2402 | 4.71 | 1.09 | 21.17 | 1.27 |
| Error | 6.04% | 22.02% | 18.00% | 10.39% | 1.32% | 1.46% |
| **Schorfheide, Song, and Yaron (2016) (1)** | | | | | | |
| Log-Lin | 3.2853 | 0.2704 | 3.21 | 1.73 | 15.06 | 0.56 |
| Correct | 3.3580 | 0.2557 | 3.06 | 1.73 | 14.50 | 0.56 |
| Error | 2.17% | 5.78 % | 5.05% | 0.12% | 3.86% | 0.17% |
| **Schorfheide, Song, and Yaron (2016) (2)** | | | | | | |
| Log-Lin | 2.5943 | 0.8748 | 8.56 | -0.00 | 14.42 | 0.91 |
| Correct | 3.3365 | 0.7841 | 4.85 | 0.92 | 12.98 | 0.54 |
| Error | 22.25% | 11.57 % | 76.47% | 102.95% | 11.09% | 69.04% |
| **Bollerslev, Xu, and Zhou (2015)** | | | | | | |
| Log-Lin | 2.7479 | 0.2737 | 7.07 | 1.16 | 16.41 | 1.54 |
| Correct | 2.8222 | 0.2835 | 6.78 | 1.17 | 16.05 | 1.54 |
| Error | 2.63% | 3.56% | 4.36% | 0.72% | 2.28% | 0.37% |
| **Koijen, Lustig, Van Nieuwerburgh, and Verdelhan (2010)** | | | | | | |
| Log-Lin | 3.1137 | 0.1808 | 4.94 | 1.38 | 11.61 | 1.17 |
| Correct | 3.3514 | 0.1479 | 3.64 | 1.38 | 10.63 | 1.14 |
| Error | 7.09% | 22.31% | 35.62% | 18.73% | 9.19% | 2.64% |

The table shows 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 return. Results obtained by the log-linearization and the correct 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 (2016) (Set (1) reports the results for the median parameter estimates reported and set (2) shows the results for the 95% estimates for the persistences of the state processes), Bollerslev, Xu, and Zhou (2015), and Koijen, Lustig, Van Nieuwerburgh, and Verdelhan (2010). All returns and volatilities are shown in percent, so a value of 1.5 is a 1.5% annualized figure.

vol-of-vol factors are considerably lower ($\nu = 0.64$ and $\rho_q = 0.46$). Consequently, the approximation errors are rather small with a maximum error of 4.36% for the equity premium. This result is not surprising as the authors mention in their estimation that the stochastic volatility and the vol-of-vol factors only influence the variance premium and have a negligible influence on the price and return dynamics. And so, 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, Lustig, Van Nieuwerburgh, and Verdelhan (2010) we also find large errors with a maximum error in the equity premium of 35.62% and an overestimation of the premium by about 130 basis points. 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 the large nonlinearities to the model. Koijen, Lustig, Van Nieuwerburgh, and Verdelhan (2010) not only analyze equity markets but also price real and nominal bonds to analyze the martingale component in the stochastic discount factor. In Figure 6 we show the real and nominal yield curve for their model.

Figure 6: Real and Nominal Yield Curve in the Model of Koijen, Lustig, Van Nieuwerburgh, and Verdelhan (2010)



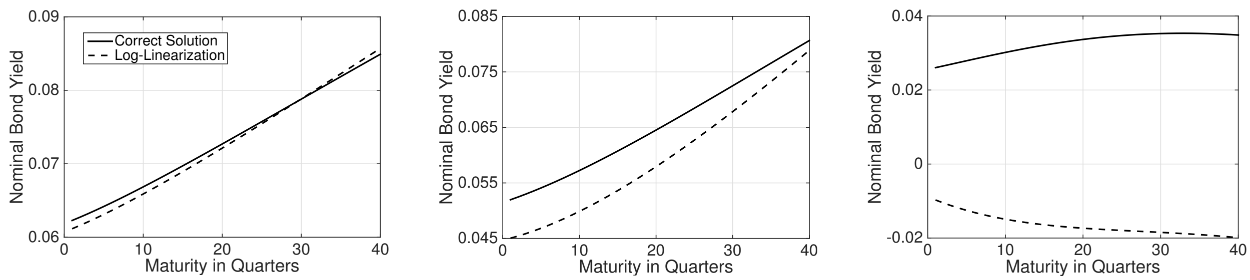

The graph shows the yield curves for real and nominal bonds in the model of Koijen, Lustig, Van Nieuwerburgh, and Verdelhan (2010). Panel (c) shows the nominal yield curve only for 1-20 months bonds (extract of Panel (b)).

We find that the differences between the yield curve obtained by linearizing the model and solving it accurately using the projection approach are small in absolute values. However, the nominal yield curve from the linearized model differs in its shape. While the true nominal yield curve is downwards sloping in the short run and upwards sloping in the long run (see Panel (c)), this pattern does not occur when using log-linearization. So linearizing the model potentially affects the shape of the real curve. The work of Bansal and Shaliastovich (2013) provides further insights to this finding. In Figure 7 we show the nominal yield curve in their model.

Panel (a) shows the yield curve for the parameters in the original study. We observe that

Figure 7: Nominal Yield Curve in the Model of Bansal and Shaliastovich (2013)

(a) $\rho = 0.81$ (b) $\rho = 0.9$ (c) $\rho = 0.975$




The graph shows the yield curve for nominal bonds in the model of Bansal and Shaliastovich (2013).

the difference between the log-linearized solution and the projection solution is negligible with very small errors and also the shape of the yield curve is correct. As Bansal and Shaliastovich (2013) use bond data to estimate the model, they find a 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%.[12] Therefore we increase $\rho$ in panels (b) and (c) to 0.9 and 0.975 correspondingly to increase the premium paid for long-run consumption risk.[13] We find that the errors in the yield curve grow significantly as $\rho$ approaches the value 1. 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 even about the shape of the yield curve.

In sum, we observe 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 (2016), and Koijen, Lustig, Van Nieuwerburgh, and Verdelhan (2010). For these latter models, the poor approximations have a strong effect on the model predictions for key financial statistics. Our observations motivate the next step in our analysis. We want to understand which model characteristics affect the performance of the log-linearization approach; simply put, when can we trust the results of such an approach and

[12] 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 the 2007 working paper of their paper. The process for $\Delta d_{t+1}$ is the same as in Koijen, Lustig, Van Nieuwerburgh, and Verdelhan (2010) (see equation 14). As the 2007 working paper assumes a monthly decision interval and the published version from 2013 has a quarterly interval, we adjust the volatility of dividends $\phi_d$ to match the volatility of dividend growth in the data of approximately 11% annualized.

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

when can we not? And related to this question, we also want to understand which properties of the exact solution lead to a poor performance of a linear method; that is, what exactly goes wrong with the linearized solution?

## 4.3 The Interplay of the State Processes

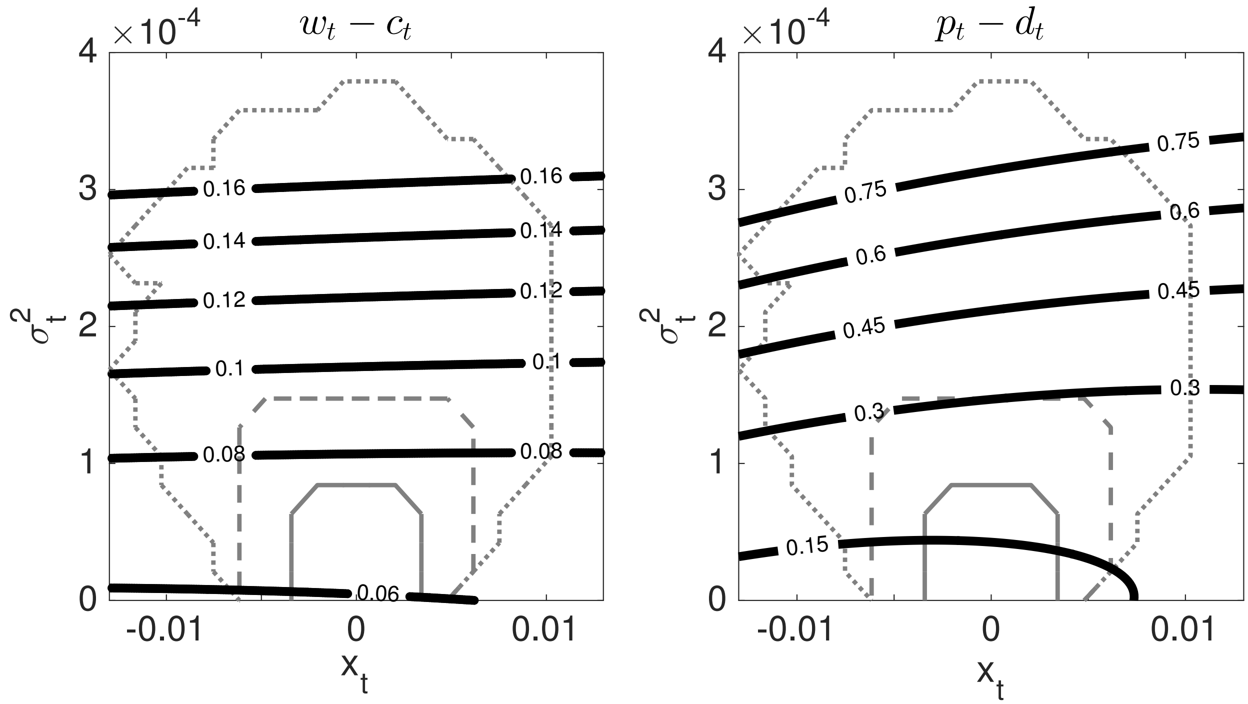
The log-linearization approach assumes that over the state space of the model the first derivatives of the solution are approximately constant and the second derivatives are approximately zero. We now show numerically that this assumption fails to hold for models with more than one highly persistent state process. We demonstrate that for solutions of such models the second derivatives can be very large and so the interplay of the state processes leads to highly nonlinear solutions; therefore, higher-order effects matter for the predictions of such models. The sizable deviations from linearity in the models' solutions is the key reason for the failure of the log-linearization approach.

For the purpose of making these points, we concentrate on the two fundamental factors of long-run risk and stochastic volatility (see equation (7)). We use the calibration of Bansal, Kiku, and Yaron (2012). Figure 8 shows isolines 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 $x$ and $\sigma^2$ (black solid lines). For example along a line marked with '0.1', the absolute error of the log-linearization is 0.1. The figure also shows the regions into which 50%, 90%, and 100% of the observations fall. These regions show the subsets of the state space that the model actually visits and in which regions it "spends most of its time" during long simulations. Corresponding errors for the first derivatives with respect to the state variables are shown in Figure 9 and for the second derivative in Figure 10.

We find that the errors in the log wealth-consumption are rather small, with maximum values of about 0.16 within the observation range. For the log price-dividend ratio, the errors are also small in the area close to the long-run mean of the processes, but they increase significantly with $\sigma^2$ and reach values of up to 0.3 in the 90% observation range, see Figure 8. In other words, the price-dividend ratio obtained by the log-linearization is off by a factor of $e^{0.3} \approx 1.35$ for almost 10% of the time and can be off by a factor larger than 2 for extreme values reached in the simulations.

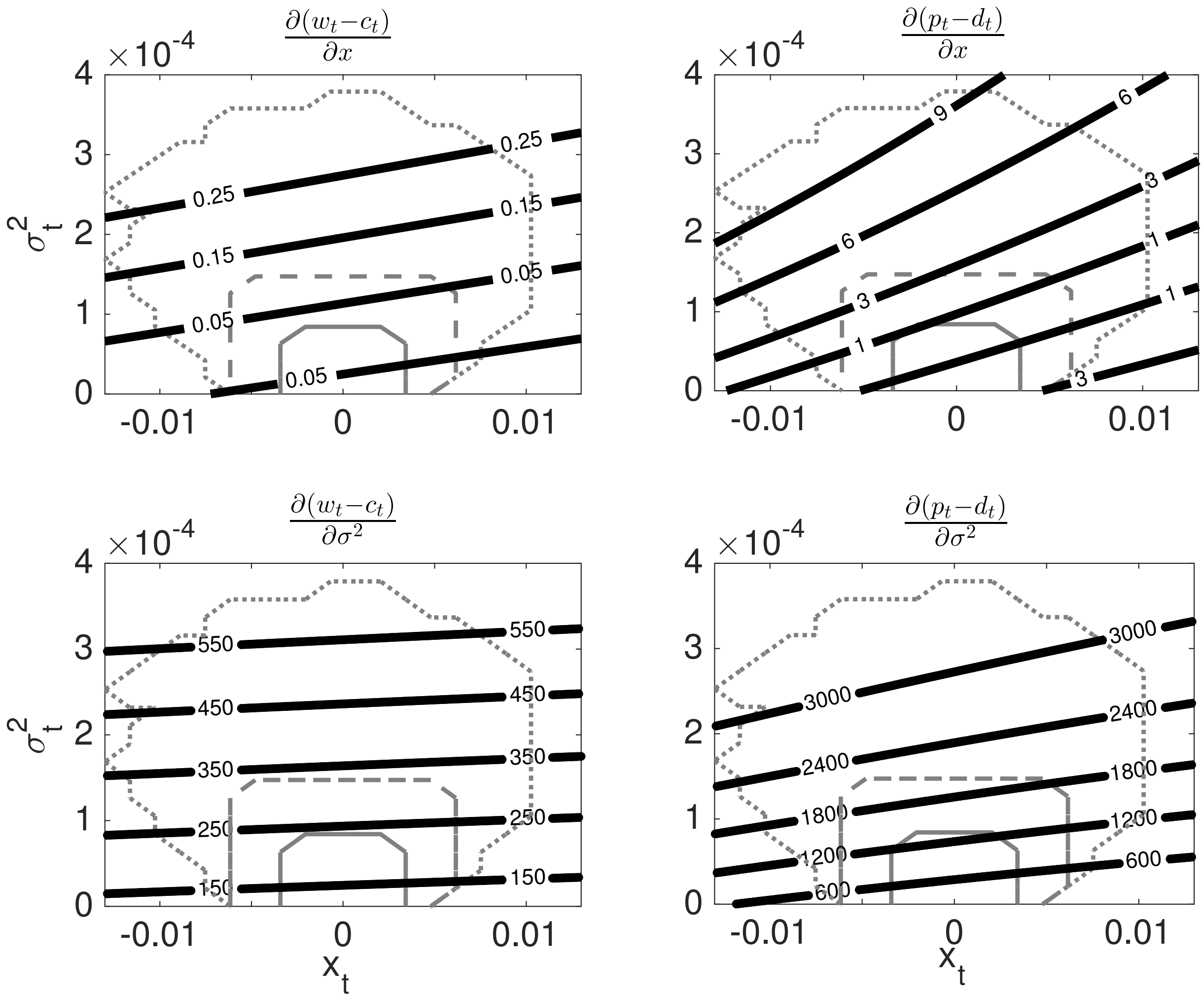
The errors in the first derivatives show similar patterns. Again the errors in the derivatives of the price-dividend ratio are significantly larger than the errors in the derivatives of the wealth-consumption ratio and the errors increase monotonically with $\sigma^2$ for the BKY (2012) calibration. We observe in Figure 9 that the errors in the derivatives with respect to $\sigma^2$ are especially large, with errors as large as 3000 for the second derivative of the price-dividend ratio. As mentioned above, the main purpose of the BKY (2012) calibration is to amplify the

Figure 8: Approximation Errors in the log Wealth-Consumption and log Price-Dividend Ratio of the Log-Linearization




The graph shows isolines 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 $x$ and $\sigma^2$ (black solid lines). The (grey) dotted, dashed and solid lines mark the respective areas into which 100%, 90% and 50% of the observations from $10^6$ simulated data points fall. The parameter values are from the calibration of Bansal, Kiku, and Yaron (2012), see Table 1.

Figure 9: Approximation Errors in the First Derivatives of the log Wealth-Consumption and log Price-Dividend Ratio of the Log-Linearization




The graph shows 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 $x$ and $\sigma^2$ of the log-linearization (black solid lines). The (grey) dotted, dashed and solid lines mark the respective areas into which 100%, 90% and 50% of the observations from $10^6$ simulated data points fall. The parameter values are from the calibration of Bansal, Kiku, and Yaron (2012), see Table 1.

role of the stochastic volatility channel by increasing its persistence. But as demonstrated in the figures, this effect introduces large nonlinearities to the model that cannot be captured by the log-linearization and hence causes large approximation errors. Figure 10 shows that the second derivatives in the model are substantially different from zero (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 $\sigma^2$, which is another reason for the large approximation errors reported in Table 2.

In general, the figures show that the stochastic volatility channel highly influences the nonlinear aspects of the model. But is it only the stochastic volatility that matters? Caldara, Fernandez-Villaverde, Rubio-Ramirez, and Yao (2012) analyze the accuracy of several solution methods in a neoclassical growth model with Epstein-Zin preferences and stochastic volatility. They report that higher-order approximations are needed to capture the nonlinearities of the model. Bansal, Kiku, and Yaron (2016) report approximation errors for the long-run risks model in their estimation study by comparing the results of the log-linearization to the results obtained by the discretization method of Tauchen and Hussey (1991) (see Table A.1 of their paper). In their original paper, they use a simplified version of their model that only features long-run risks (and no stochastic volatility). They find rather small approximation errors. But in the long-run risk model, there are two sources of nonlinearities: the stochastic volatility channel and the long-run risk channel. Hence, when solving the model, it is essential to understand whether and how the interplay of the two components drives the nonlinearities.

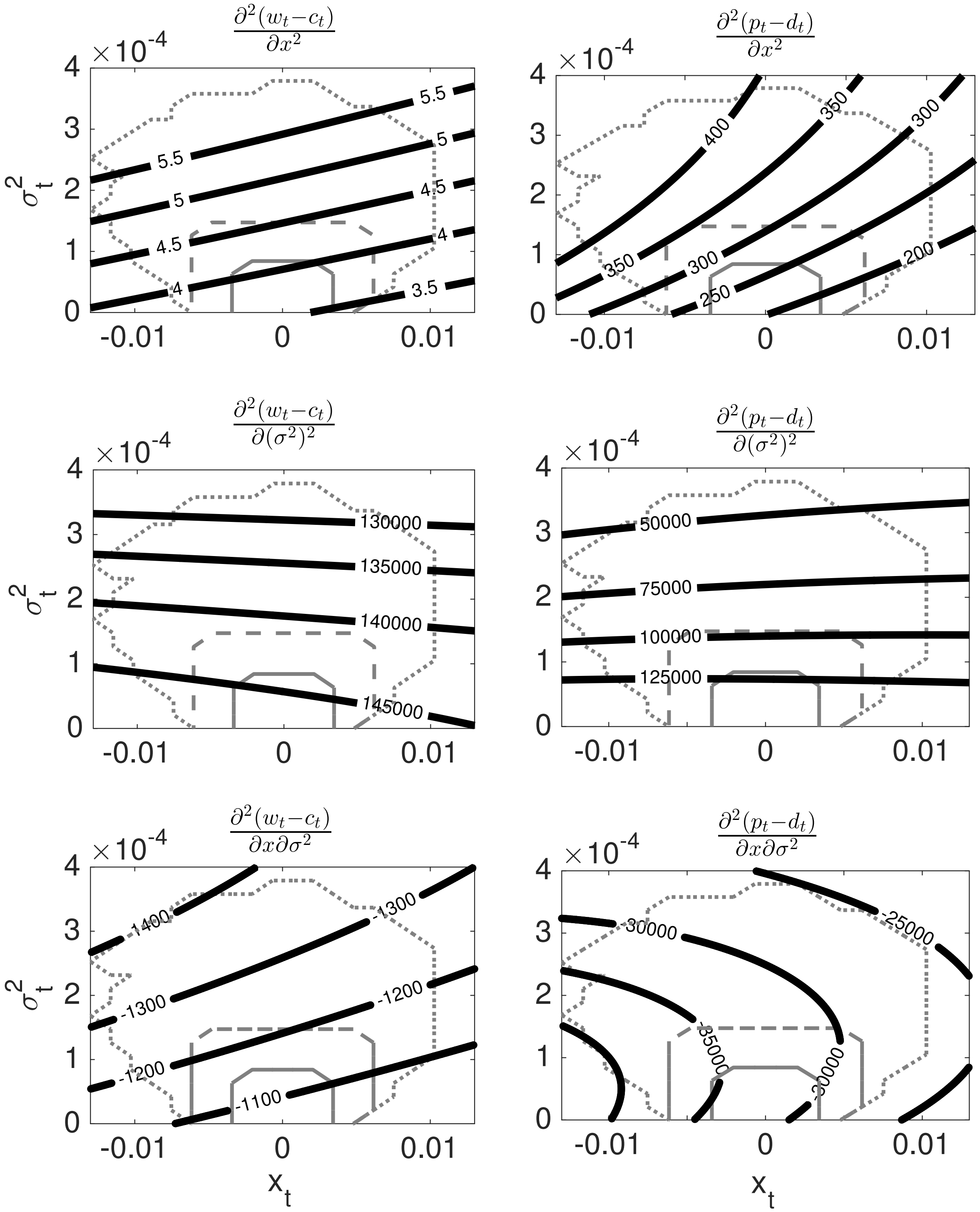
To obtain such an understanding, we analyze the approximation errors implied by the log-linearization for each of the two state variables of the model separately. In particular we first fix the stochastic volatility to its long-run mean, $\sigma_t = \bar{\sigma}^2 \ \forall \ t$, and secondly, we solve the model without long-run risk, $x_t = 0 \ \forall \ t$. Table 3 shows the corresponding errors[14] in

Table 3: Approximation Errors for Each State of the Long-Run Risks Model Separately

| | $E(w_t - c_t)$ | $\sigma(w_t - c_t)$ | $E(p_t - d_t)$ | $\sigma(p_t - d_t)$ |
|---|---|---|---|---|
| State: $x_t$ | 0.003% | 0.024% | 0.084% | 0.21% |
| State: $\sigma_t^2$ | 0.14% | 4.49% | 2.62% | 7.05% |
| Both States: | 1.05% | 12.25% | 3.15% | 26.90% |

*The table shows approximation errors in the unconditional mean and standard deviation of the (monthly) log wealth-consumption and log price-dividend ratio induced by the log-linearization in the long-run risk model for each of the two state variables $x_t$ and $\sigma_t$ separately. For the case with only $x_t$, the state $\sigma_t$ is simply set constant at its long-run mean $\bar{\sigma}^2$ (or equivalently $\nu = \sigma_w = 0$). For the case with only $\sigma_t$, $x_t$ is set to 0 (or equivalently $\rho = \phi_x = \Phi = 0$). The parameter values are from the calibration of Bansal, Kiku, and Yaron (2012), see Table 1.*

Figure 10: Approximation Errors in the Second Derivatives of the log Wealth-Consumption and Price-Dividend Ratio




The graph shows 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 $x$ and $\sigma^2$ of the log-linearization (black solid lines). The (grey) dotted, dashed and solid lines mark the respective areas into which 100%, 90% and 50% of the observations from $10^6$ simulated data points fall. The parameter values are from the calibration of Bansal, Kiku, and Yaron (2012), see Table 1.

the unconditional mean and standard deviation of the log wealth-consumption and log price-dividend ratio for the two cases. We find that, in line with the test results from Bansal, Kiku, and Yaron (2016), for the one-dimensional model with only long-run risks, the approximation errors are very small with a maximum error of 0.21%. For the second case, without long-run risks and only stochastic volatility, the errors are slightly larger but still remain below 7.1%. However, for the full model with long-run risk and stochastic volatility approximation errors increase dramatically with a maximum error of 26.9% for the volatility of the log price-dividend ratio. This finding suggests that neither the stochastic volatility alone nor the long-run risks component alone introduces the nonlinearities in the model; instead it is the simultaneous presence and interplay of the two features which makes the model so difficult to solve.

# 5 Conclusion

In this paper we have presented an analysis of higher-order effects in asset pricing models with long-run risk. We have shown that solutions of models that build on the framework of Bansal and Yaron (2004) are potentially very nonlinear and that for very persistent exogenous processes the approximation errors introduced by the Campbell-Shiller log-linearization method can be large and economically significant. For example, in the most recent calibration of the Bansal-Yaron long-run risk model, see Bansal, Kiku, and Yaron (2012), the approximation errors in the annual volatility of the log price-dividend ratio exceed 22%; similarly, the errors in the equity premium exceed 75% in the estimation study of Schorfheide, Song, and Yaron (2016). Models with lower persistence, such as the original Bansal and Yaron (2004) model or the volatility-of-volatility model of Bollerslev, Xu, and Zhou (2015), have much smaller approximation errors. The results for nominal bonds in Bansal and Shaliastovich (2013) and Koijen, Lustig, Van Nieuwerburgh, and Verdelhan (2010) are particularly interesting; for the high level of persistence necessary to explain the equity premium, the log-linear approximation can actually produce a downward sloping yield curve, when, on the contrary, the true yield curve is upward sloping.

In addition to numerical errors, the Campbell-Shiller approach can lead to misleading economic conclusions. For example, conventional wisdom is that long-run risk cannot generate time-varying risk premia, and that an additional mechanism such as stochastic volatility is required. We show that once nonlinear effects are accounted for, long-run risk can generate sizable time-varying risk premia. Interestingly, the time variation generated by long-run risk is procyclical—the risk premium is higher when growth is high, and lower when growth is low.

In light of the tremendous impact of long-run risk models on the literature of asset pricing,

[14] Note that the reported errors in Table 3 are for monthly statistics while the numerical errors in Table 2 are provided for annualized statistics.

our results strongly suggest that more sophisticated solution methods, such as projection methods, should be used for the analysis of asset pricing models with highly persistent state processes. The solution methods applied to the analysis of these models should be able to account for the potentially very nonlinear relationships between the endogenous model quantities and the exogenous state variables.

# Appendix

## A Computational Methods for Asset Pricing Models with Recursive Preferences

One of the common approaches to solve 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 the risk aversion has no influence (as there is no risk). But if the 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, for example, Juillard (2011), de Groot (2013) or Meyer-Gohde (2014)).[15] Another drawback of the standard log-linearization is that the policies are independent of the volatility of the model (see Caldara, Fernandez-Villaverde, Rubio-Ramirez, and Yao (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 features, 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 which, therefore, has been used extensively for solving asset pricing models with recursive preferences (Segal, Shaliastovich, and Yaron (2015), Bansal, Kiku, and Yaron (2010), Bansal, Kiku, and Yaron (2012), Bollerslev, Tauchen, and Zhou (2009), Kaltenbrunner and Lochstoer (2010), Koijen, Lustig, Van Nieuwerburgh, and Verdelhan (2010), Drechsler and Yaron (2011), Bansal and Shaliastovich (2013), Constantinides and

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

Ghosh (2011), Bansal, Kiku, Shaliastovich, and Yaron (2014) or Beeler and Campbell (2012), among others).[16] One reason for its popularity 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 in Section A.1.

This study analyzes the log-linearized model solution with regard to the influence of higher order dynamics on equilibrium outcomes that can, by construction, not 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, there are to the best of our knowledge no such solutions. Therefore, we need a highly accurate solution method which is capable to correctly capture higher-order features of the asset returns. A convenient choice is projection methods that allow to choose the approximation degree as well as the size of the approximation interval in order to be able to capture higher-order dynamics that are driven by the tails of the distribution.[17] Projection methods are a general-purpose tool for solving functional equations. They were first introduced by physicists and engineers 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 how to apply projection methods to asset pricing models.) A detailed description of projection methods and how they can applied to solve the equilibrium conditions (2) and (6) is given in Appendix A.2.

## A.1 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

$$z_{i,t} = A_{0,i} + A_i y_t \tag{16}$$

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

[16] Another approach, proposed by Kogan and Uppal (2001) and used for example in Hansen, Heaton, Lee, and Roussanov (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 follow-up work in the long-run risk literature has focused on the log-linearization used in Bansal and Yaron (2004), so we concentrate on this particular approximation.

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

the unknown linearization coefficients. The log return of the asset $i$, $r_{i,t+1}$ is then defined as

$$r_{i,t+1} = \log\left(e^{z_{i,t+1}} + 1\right) - z_{i,t} + \Delta d_{i,t+1} \tag{17}$$

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

$$r_{i,t+1} \approx \kappa_{i,0} + \kappa_{i,1} z_{i,t+1} - z_{i,t} + \Delta d_{i,t+1} \tag{18}$$

with the linearizing constants

$$\kappa_{i,1} = \frac{e^{\bar{z}_i}}{1 + e^{\bar{z}_i}} \tag{19}$$

$$\kappa_{i,0} = -\log\left((1 - \kappa_{i,1})^{1-\kappa_{i,1}} \kappa_{i,1}^{\kappa_{i,1}}\right) \tag{20}$$

that only depend 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 (18) into the equilibrium condition (6) yields

$$E_t\left[e^{\theta \log \delta + (\theta - \frac{\theta}{\psi})\Delta c_{t+1} + \theta(\kappa_{w,0} + \kappa_{w,1} z_{w,t+1} - z_{w,t})}\right] = 1. \tag{21}$$

The equilibrium condition now only depends 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 the expectation can be evaluated analytically, as for example for processes with normal innovations as in Bansal and Yaron (2004) or Bollerslev, Tauchen, and Zhou (2009). This allows for approximate closed-form solutions for the linearization coefficients that only depend 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.

## A.2 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

$$(\mathcal{G}z)(x) = 0, \tag{22}$$

where the variable $x$ 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 given operator $\mathcal{G}$ is a continuous mapping between two function spaces. Note that solving equation (22) 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 $X$ by 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.[18] For a set $\{\Lambda_k\}_{k\in\{0,1,\ldots,n\}}$ of chosen basis functions the approximation $\hat{z}$ of $z$ is

$$\hat{z}(x;\boldsymbol{\alpha}) = \sum_{k=0}^{n} \alpha_k \Lambda_k(x), \tag{23}$$

where $\boldsymbol{\alpha} = [\alpha_0, \alpha_1, \ldots, \alpha_n]$ are unknown coefficients. Replacing the function $z$ in equation (22) by its approximation $\hat{z}$, we can define the residual function $\hat{F}(x;\boldsymbol{\alpha})$ as the error in the original equation,

$$\hat{F}(x;\boldsymbol{\alpha}) = (\mathcal{G}\hat{z})(x;\boldsymbol{\alpha}). \tag{24}$$

Instead of solving equation (22) 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.

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 $\alpha$ are now the projection conditions,

$$\int_X \hat{F}(x;\boldsymbol{\alpha}) g_k(x) w(x) dx = 0, \;\; k = 0, 1, \ldots, n. \tag{25}$$

[18]For the ease of notation, we demonstrate the projection approach using polynomials for the basis functions. Alternative specifications are for example rational basis functions, or piecewise polynomial approximations. For example to solve the studies with high-dimensional state spaces like the model of Schorfheide, Song, and Yaron (2016), we use cubic splines as there is an efficient and fast implementation in Matlab.

Observe that this system of equations imposes $n+1$ conditions on the $(n+1)$-dimensional vector $\alpha$. 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 and the Galerkin method.

A collocation method chooses $n+1$ distinct nodes in the domain, $\{x_k\}_{k=0}^{n}$, and defines the test functions $g_k$ by

$$g_k(x) = \begin{cases} 0 \text{ if } x \neq x_k \\ 1 \text{ if } x = x_k. \end{cases}$$

With a weight term $w(x) \equiv 1$, the projection conditions (25) simplify to

$$\hat{F}(x_k; \boldsymbol{\alpha}) = 0, \;\; k = 0, 1, \ldots, n. \tag{26}$$

Simply put, the collocation method determines the coefficients in the approximation (23) by solving the square system (26) 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 (25) become

$$\int_X \hat{F}(x; \boldsymbol{\alpha}) \Lambda_k(x) \frac{1}{\sqrt{1-x^2}} dx = 0, \;\; k = 0, 1, \ldots, n. \tag{27}$$

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

### A.2.1 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 (22). For this purpose, we need to choose an appropriate state space and perform the usual transformation from an equilibrium described by infinite sequences (with a 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 subsequent state in the next period by $x'$. (For example in the original model by 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 $\sigma_t^2$), 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 a density $f_x$.

First note that 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, in a second step, we can solve for any asset return by applying the projection approach to equation (2). For the first step, write equation (6) in state-space representation

$$E\left[\exp\left(\theta\log\delta - \frac{\theta}{\psi}\Delta c(x'|x) + \theta r_w(x'|x)\right)\bigg| x\right] = 1, \quad \forall x, \tag{28}$$

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 $z_w(x) = \log\left(\frac{W(x)}{C(x)}\right)$. Next, write the state-dependent log return of the aggregate consumption claim as

$$\begin{aligned} r_w(x'|x) &= \log\left(\frac{W(x')}{W(x)-C(x)}\right) = \log\left(\frac{\frac{W(x')}{C(x')}}{\frac{W(x)}{C(x)}-1} \times \frac{C(x')}{C(x)}\right) \\ &= z_w(x') - \log\left(e^{z_w(x)} - 1\right) + \Delta c(x'|x). \end{aligned} \tag{29}$$

Inserting the last term in equation (28) yields

$$E\left[\exp\left(\theta\left(\log\delta + (1-\frac{1}{\psi})\Delta c(x'|x) + z_w(x') - \log\left(e^{z_w(x)}-1\right)\right)\right) - 1\bigg| x\right] = 0, \quad \forall x. \tag{30}$$

Equivalently,

$$0 = \int_X \left[\exp\left(\theta\left(\log\delta + (1-\frac{1}{\psi})\Delta c(x'|x) + z_w(x') - \log\left(e^{z_w(x)}-1\right)\right)\right) - 1\right] \mathrm{d}f_x \tag{31}$$

which is a functional equation of the form (22) and allows us to apply the projection approach.

The unknown solution function to this equilibrium condition, $z_w$, is an element of a function space which is an infinite-dimensional vector space. A key feature of every projection method is to approximate the solution function $z_w$ by an element from a finite-dimensional space. Specifically, we use the approximation $\hat{z}_w(x;\boldsymbol{\alpha_w}) = \sum_{k=0}^{n} \alpha_{w,k}\Lambda_k(x)$, where $\{\Lambda_k\}_{k\in\{0,1,\dots,n\}}$ is a set of chosen (known) basis functions and $\boldsymbol{\alpha_w} = [\alpha_{w,0}, \alpha_{w,1}, \dots, \alpha_{w,n}]$ 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 (31), which is defined by

$$\hat{F}_w(x;\boldsymbol{\alpha_w}) = \int_X \left[\exp\left(\theta\left(\log\delta + (1-\frac{1}{\psi})\Delta c(x'|x) + \hat{z}_w(x') - \log\left(e^{\hat{z}_w(x)}-1\right)\right)\right) - 1\right] \mathrm{d}f_x. \tag{32}$$

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 different such projection conditions, the collocation and the Galerkin method, see Appendix A.2. 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, $\hat{r}_w(x'|x;\boldsymbol{\alpha_w}) = \hat{z}_w(x';\boldsymbol{\alpha_w}) - \log\left(e^{\hat{z}_w(x;\boldsymbol{\alpha_w})} - 1\right) + \Delta c(x'|x)$.

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 $z_i(x) = \log\left(\frac{P(x)}{D(x)}\right)$ and rewrite the state-dependent log return of asset $i$ as

$$\begin{aligned} r_i(x'|x) &= \log\left(\frac{P_i(x') + D_i(x')}{P_i(x)}\right) = \log\left(\frac{\frac{P_i(x')}{D_i(x')} + 1}{\frac{P_i(x)}{D_i(x)}} \times \frac{D_i(x')}{D_i(x)}\right) \\ &= \log\left(e^{z_i(x')} + 1\right) - z_i(x) + \Delta d_i(x'|x). \end{aligned} \tag{33}$$

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

$$E\left[\exp\left(\theta\log\delta - \frac{\theta}{\psi}\Delta c(x'|x) + (\theta - 1)r_w(x'|x) + r_i(x'|x)\right)\Bigg|\, x\right] = 1. \tag{34}$$

Substituting the return expressions (29) and (33) into this equations 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

$$\begin{aligned} \hat{F}_i(x;\boldsymbol{\alpha_i}) = \int_X \Bigg[\exp\Bigg(&\theta\log\delta - \frac{\theta}{\psi}\Delta c(x'|x) + (\theta - 1)\hat{r}_w(x'|x;\boldsymbol{\alpha_w}) \\ &+ \log\left(e^{\hat{z}_i(x';\boldsymbol{\alpha_i})} + 1\right) - \hat{z}_i(x;\boldsymbol{\alpha_i}) + \Delta d_i(x'|x)\Bigg) - 1\Bigg] \mathrm{d}f_x \end{aligned} \tag{35}$$

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 (32). Once $\boldsymbol{\alpha_w}$ is known, the projections can be applied to equation (35) 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* (29) *to derive the residual function for the return on wealth*

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

*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* (33) *to derive the residual function for asset* $i$*,*

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

*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 actually implement the algorithm, we need to specify additional algorithmic details such as the choices for basis functions and the integration technique.

### A.2.2 Algorithmic Ingredients

In the **Initialization** step, 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 $z_i$ by 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

$$\Lambda_k(x) = T_k\left(2\left(\frac{x - x_{min}}{x_{max} - x_{min}}\right) - 1\right) \tag{36}$$

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

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,

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

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

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

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 (32) and (35) 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, see Appendix A.2, 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 errors (RMSE) and maximum absolute errors (MAE) in the residual function (35); these errors are

$$\text{RMSE}_i = \sqrt{\frac{1}{m^e}\sum_{j=1}^{m^e} \hat{F}_i(x_j^e|\boldsymbol{\alpha_i})^2}, \tag{37}$$

$$\text{MAE}_i = \max_{j=1,2,\ldots,m^e} |\hat{F}_i(x_j^e|\boldsymbol{\alpha_i})|, \tag{38}$$

respectively, with

$$x_j^e = x_{min} + \frac{x_{max} - x_{min}}{m^e - 1}(j-1), \qquad j = 1, \ldots, m^e. \tag{39}$$

# B Computational Details and Accuracy of the Projection Method

All the results in the paper are computed using Matlab. We use the solver 'fmincon' with the SQP algorithm. As 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 (2016) and Koijen, Lustig, Van Nieuwerburgh, and Verdelhan (2010), we use cubic splines with not-a-knot end conditions instead of Chebyshev polynomials due to the faster implementation in Matlab.

Table 4 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) with different numbers of discretization nodes.[19] Another method, popular in macroeconomics, is perturbation methods (see, for example, Caldara, Fernandez-Villaverde, Rubio-Ramirez, and Yao (2012)). de Groot (2015) compared these methods to the analytical solutions, and finds they perform worse than even the log-linearization for long-run risk models, so we do

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

not consider them further. We find that for the calibration with $\rho = 0.95$ already a first order approximation with an approximation interval of $n_\sigma = 1$ standard deviation around the unconditional mean of $x_t$ provides a very accurate solution with an approximation error of 1.51e-5 for $\mathrm{E}(\frac{W}{C})$ and 2.37e-6 for $\mathrm{std}(\frac{W}{C})$ for the case with recursive utility and $\gamma = 10$. For the high persistence case with $\rho = 0.99$ a larger degree is required and the degree four polynomial is sufficient to compute a highly accurate solution. Overall we observe, that the projection method provides highly accurate solutions for all specifications considered in this example. Although not reported, we have performed similar exercises for the six asset pricing models used in this study to ensure highly accurate solutions.

# C The Volatility of Volatility Factor

This section analyzes 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) or Bollerslev, Xu, and Zhou (2015). For this purpose we use the parsimonious model formulation as in Bollerslev, Tauchen, and Zhou (2009) who take the basic model setup (9) without the long-run risk factor, so $\phi_x = 0$, and add vol-of-vol modeled by a square root process $q_t$:

$$\begin{aligned}
\sigma_{t+1}^2 &= \bar{\sigma}^2(1-\nu) + \nu\sigma_t^2 + \sqrt{q_t}\eta_{\sigma,t+1} \\
q_{t+1} &= \mu_q(1-\rho_q) + \rho_q q_t + \phi_q\sqrt{q_t}\eta_{q,t+1} \\
&\quad \eta_{\sigma,t+1}, \eta_{q,t+1} \sim i.i.d.\ N(0,1).
\end{aligned} \tag{40}$$

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 has also been used for the stochastic volatility process in the original Bansal and Yaron (2004) study and many subsequent papers). However, to compute model solutions, the assumption of a non-truncated 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 11 shows model outcomes for CRRA preferences with $\psi = 1.5$ (Panel (a)) and the corresponding EZ case with $\gamma = 10$ (Panel (b)) for various persistence and volatility parameters of the vol-of-vol process $\rho_q$ and $\phi_q$. The black numbers show the true mean wealth-consumption ratio under the assumption of a reflecting boundary for $q_t$ at zero. Blue values are the results

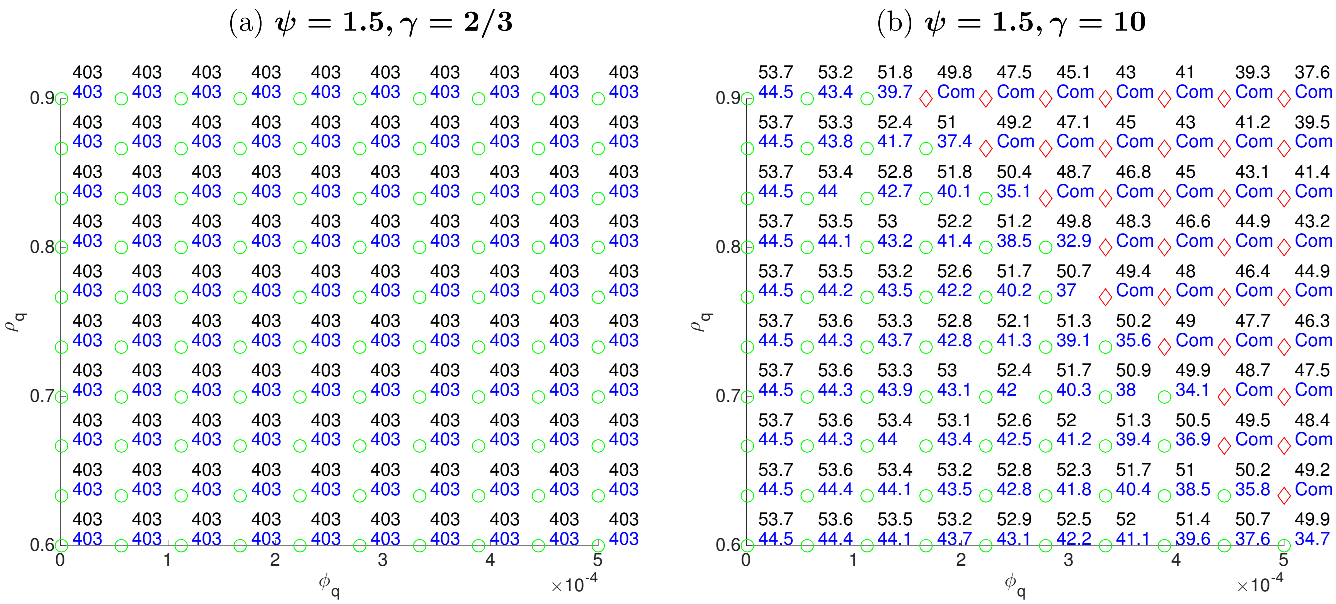
Table 4: Accuracy of the Projection Method

| | **Closed-F.** | **Log-Lin** | **Projection** | | | **Discretization** | | |
|---|---|---|---|---|---|---|---|---|
| | | | $n = 1$ $n_\sigma = 1$ | $n = 4$ $n_\sigma = 4$ | $n = 16$ $n_\sigma = 32$ | $n_D = 5$ | $n_D = 10$ | $n_D = 50$ |
| | | | | $\boldsymbol{\psi = 1.5, \gamma = 1/\psi}$ | | | | |
| | | | | $\rho = 0.95$ | | | | |
| $\mathrm{E}(\frac{W}{C})$ | 1681.20 | 1681.16 | 1681.18 | 1681.20 | 1681.20 | 1669.99 | 1670.75 | 1671.00 |
| Error | 0 | 2.11e-5 | 1.19e-5 | 2.61e-8 | 2.60e-8 | 0.0067 | 0.0062 | 0.0060 |
| $\mathrm{std}(\frac{W}{C})$ | 12.1815 | 12.1812 | 12.1813 | 12.1815 | 12.1815 | 11.2206 | 11.2384 | 11.5015 |
| Error | 0 | 2.20e-5 | 1.27e-5 | 2.65e-8 | 2.64e-8 | 0.0789 | 0.0774 | 0.0558 |
| | | | | $\rho = 0.99$ | | | | |
| $\mathrm{E}(\frac{W}{C})$ | 1868.36 | 1862.93 | 1865.54 | 1868.36 | 1868.36 | 3404.73 | 2121.64 | 1852.27 |
| Error | 0 | 0.0029 | 0.0015 | 1.21e-7 | 7.65e-11 | 0.8223 | 0.1356 | 0.0086 |
| $\mathrm{std}(\frac{W}{C})$ | 144.14 | 143.65 | 143.86 | 144.14 | 144.14 | 1966.49 | 381.82 | 141.39 |
| Error | 0 | 0.0034 | 0.0020 | 1.17e-7 | 8.17e-11 | 12.6427 | 1.6489 | 0.0191 |
| | | | | $\boldsymbol{\psi = 1.5, \gamma = 10}$ | | | | |
| | | | | $\rho = 0.95$ | | | | |
| $\mathrm{E}(\frac{W}{C})$ | - | 1314.39 | 1314.59 | 1314.61 | 1314.61 | 1532.25 | 1514.25 | 1508.08 |
| Error | - | 1.66e-4 | 1.51e-5 | 4.37e-11 | 2.12e-12 | 0.1655 | 0.1518 | 0.1472 |
| $\mathrm{std}(\frac{W}{C})$ | - | 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 | 10.0744 | 0.0689 | 0.0918 |
| | | | | $\rho = 0.99$ | | | | |
| $\mathrm{E}(\frac{W}{C})$ | - | 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 |
| $\mathrm{std}(\frac{W}{C})$ | - | 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 |

The table shows the mean wealth-consumption ratio for the long-run risk model of Bansal and Yaron (2004) with constant volatility (Equations (7) with $\sigma_{c,t} = \sigma_{x,t} = \bar{\sigma}$ and $\eta_{c,t+1}$, $\eta_{x,t+1}$ i.i.d. normal.). Results are shown for the log-linearization, the projection as well as the discretization by Tauchen and Hussey (1991) with the extension of Floden (2007) that performs better for highly persistent processes. For the projection method solutions with three different degrees $n$ where the approximation interval is set up $n_\sigma$ standard deviations around the unconditional mean of the long-run risk process $x_t$ are provided. For the discretization results are shown for three different numbers of approximation nodes $n_D$. The table also shows the relative error of the solutions, where in the case of $\gamma = 1/\psi$ the closed-form solution is taken from de Groot (2015) and 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 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 $\phi_x = 0.044$.

from log-linearization under the assumption of a standard non-truncated normal distribution. Green circles denote convergence of both, the projection and the log-linearization approach. Red 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 risk aversion, using the standard log-linearization technique can lead to complex solutions. This is for example the case for the calibration in Bollerslev, Tauchen, and Zhou (2009) with $\rho_q = 0.8$ and $\phi_q = $ 1e-3.[20]

Figure 11: Sensitivity Analysis and Existence Results in the Vol-of-Vol Model



The graph 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 $\rho_q$ and volatility parameters $\phi_q$. Panel (a) depicts the case of CRRA utility with $\psi = 1.5$, while panel (b) depicts the corresponding cases 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 blue numbers show the values obtained by the standard log-linearization with normal shocks. Green circles denote convergence of both, the projection and the log-linearization approach. Red 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.

So what are the determinants of the complexity of the linearized solution? The square-root specification of $q_t$ implies that the coefficient for $q_t$ is determined by a quadratic equation and hence may have more than one solution. The log-linear approximation of the log wealth-

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

consumption ratio $z_{w,t}$ has the following form

$$z_{w,t} = A_0 + A_\sigma \sigma_t^2 + A_q q_t \tag{41}$$

with the linearization coefficients (see Appendix A.1 for the derivation) given by

$$\begin{aligned} A_\sigma &= \frac{(1-\gamma)^2}{2\theta(1-k_1\nu)} \\ A_0 &= \frac{\log\delta + (1-\frac{1}{\psi})\mu_c + k_0 + k_1\left[A_\sigma\bar{\sigma}^2(1-\nu) + A_q\mu_q(1-\rho_q)\right]}{(1-k_1)} \\ A_q &= \frac{1 - k_1\rho_q \pm \sqrt{(1-k_1\rho_q)^2 - \theta^2 k_1^4 \phi_q^2 A_\sigma^2}}{\theta k_1^2 \phi_q^2} \end{aligned} \tag{42}$$

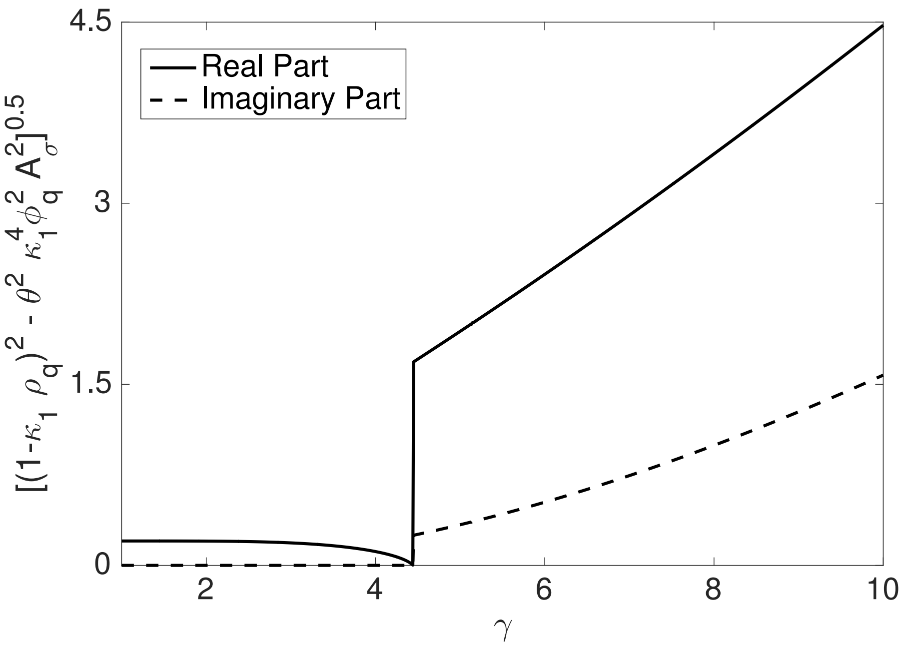
We find that the coefficient for the vol-of-vol factor $A_q$ has indeed two solutions. As Bollerslev, Tauchen, and Zhou (2009) show in their paper by the no arbitrage argument, the minus term is the economically meaningful root and the positive solution can be neglected. Complexity of the solution is determined by the term inside the square root in equation (42) 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 12 shows the values of the square root term as a function of the risk aversion $\gamma$. In line with the results above, we find that for small $\gamma$ the solution is well behaved with only a real and no imaginary part. However if we increase $\gamma$, $\theta$ becomes significantly larger (it goes from -3 for $\gamma = 2$ to -27 for $\gamma = 10$) and hence the real part of the term decreases. For a certain threshold (about 4.4 in this example) the term hits zero and the solution thereafter consists of a significant imaginary part. Also Panel (b) in Figure 11 shows, that the larger the persistence or the larger the volatility of the vol-of-vol process solutions become 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.

# References


Atkinson, K. E. (1992): "A Survey of Numerical Methods for Solving Nonlinear Integral Equations," *Journal of Integral Equations and Applications*, 4, 15–46.

Bansal, R., D. Kiku, I. Shaliastovich, and A. Yaron (2014): "Volatility, the Macroeconomy, and Asset Prices," *Journal of Finance*, 69, 2471–2511.

Figure 12: Analysis of Square-Root Term in the Vol-of-Vol Model




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


Bansal, R., D. Kiku, and A. Yaron (2010): "Long-Run Risks, the Macroeconomy, and Asset Prices," *American Economic Review*, 100, 542–546.

——— (2012): "An Empirical Evaluation of the Long-Run Risks Model for Asset Prices," *Critical Finance Review*, 1, 183–221.

——— (2016): "Risks For the Long Run: Estimation with Time Aggregation," *Journal of Monetary Economics*, 82, 52–69.

Bansal, R. and I. Shaliastovich (2013): "A Long-Run Risks Explanation of Predictability Puzzles in Bond and Currency Markets," *Review of Financial Studies*, 26, 1–33.

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

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

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

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

Caldara, D., J. Fernandez-Villaverde, J. Rubio-Ramirez, and W. Yao (2012): "Computing DSGE Models with Recursive Preferences and Stochastic Volatility," *Review of Economic Dynamics*, 15, 188–206.

Campbell, J. Y. and R. J. Shiller (1988): "The Dividend-Price Ratio and Expectations of Future Dividends and Discount Factors," *Review of Financial Studies*, 1, 195–228.

Chen, Y., T. F. Cosimano, and A. A. Himonas (2014): "On Formulating and Solving Portfolio Decision and Asset Pricing Problems," in *Handbook of the Computational Economics, Volume 3*, ed. by K. Schmedders and K. L. Judd, North-Holland, Amsterdam, 161–223.

Collard, F. and M. Juillard (2001): "Accuracy of Stochastic Perturbation Methods: The Case of Asset Pricing Models," *Journal of Economic Dynamics and Control*, 25, 979–999.

Constantinides, G. M. and A. Ghosh (2011): "Asset Pricing Tests with Long-Run Risks in Consumption Growth," *Review of Asset Pricing Studies*, 1, 96–136.

de Groot, O. (2013): "Computing the Risky Steady State of DSGE Models," *Economics Letters*, 120, 566–569.

——— (2015): "Solving Asset Pricing Models with Stochastic Volatility," *Journal of Economic Dynamics and Control*, 52, 308–321.

Dew-Becker, I., S. W. Giglio, A. Le, and M. Rodriguez (2015): "The Price of Variance Risk," Available at SSRN: `http://ssrn.com/abstract=2471079`.

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

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

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

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

Fama, E. F. and K. R. French (1989): "Business Conditions and Expected Returns on Stocks and Bonds," *Journal of Financial Economics*, 25, 23–49.

FLODEN, M. (2007): "A Note on the Accuracy of Markov-Chain Approximations to Highly Persistent AR(1) Processes," *Economics Letters*, 99, 516–520.

GROSSMAN, S. J. AND R. J. SHILLER (1981): "The Determinants of the Variability of Stock Market Prices," *The American Economic Review*, 71, 222–227.

HANSEN, L. P., J. HEATON, J. LEE, AND N. ROUSSANOV (2007): "Intertemporal Substitution and Risk Aversion," in *Handbook of Econometrics*, ed. by J. J. Heckman and E. E. Leamer, Elsevier, vol. 6, Part A, Chapter 61, 3967–4056.

HANSEN, L. P., J. HEATON, AND N. LI (2008): "Consumption Strikes Back?: Measuring Long Run Risk," *Journal of Political Economy*, 116, 260–302.

HANSEN, L. P. AND K. J. SINGLETON (1982): "Generalized Instrumental Variables Estimation of Nonlinear Rational Expectations Models," *Econometrica*, 50, 1269–1286.

JU, N. AND J. MIAO (2012): "Ambiguity, Learning, and Asset Returns," *Econometrica*, 80, 559–591.

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

——— (1996): "Approximation, Perturbation, and Projection Methods in Economic Analysis," in *Handbook of the Computational Economics, Volume 1*, ed. by H. Amman, D. Kendrick, and J. Rust, North-Holland, Amsterdam, 511–585.

——— (1998): *Numerical Methods in Economics*, The MIT Press: Cambridge, Massachusetts.

JUILLARD, M. (2011): "Local Approximation of DSGE Models Around the Risky Steady State," Available at IDEAS: `http://ideas.repec.org/p/ter/wpaper/0087.html`.

KALTENBRUNNER, G. AND L. A. LOCHSTOER (2010): "Long-Run Risk Through Consumption Smoothing," *Review of Financial Studies*, 23, 3190–3224.

KOGAN, L. AND R. UPPAL (2001): "Risk Aversion and Optimal Portfolio Policies in Partial and General Equilibrium Economies," Available at NBER: `http://www.nber.org/papers/w8609`.

KOIJEN, R. S. J., H. LUSTIG, S. VAN NIEUWERBURGH, AND A. VERDELHAN (2010): "Long-Run Risk, the Wealth-Consumption Ratio, and the Temporal Pricing of Risk," *American Economic Review*, 100, 552–556.

MEHRA, R. AND E. PRESCOTT (1985): "The Equity Premium: A Puzzle," *Journal of Monetary Economics*, 15, 145–161.

Meyer-Gohde, A. (2014): "Risky Linear Approximations," Available at IDEAS: `http://ideas.repec.org/p/hum/wpaper/sfb649dp2014-034.html`.

Pohl, W., K. Schmedders, and O. Wilms (2015): "Higher-Order Effects in Asset-Pricing Models with Long-Run Risks," Available at IDEAS: `https://ideas.repec.org/p/red/sed016/306.html`.

Schorfheide, F., D. Song, and A. Yaron (2016): "Identifying Long-Run Risks: A Bayesian Mixed-Frequency Approach," Available at: `http://www.economics.illinois.edu/seminars/documents/Frank.pdf`.

Segal, G., I. Shaliastovich, and A. Yaron (2015): "Good and Bad Uncertainty: Macroeconomic and Financial Market Implications," *Journal of Financial Economics*, 117, 369–397.

Tauchen, G. (2011): "Stochastic Volatility in General Equilibrium," *Quarterly Journal of Finance*, 1, 707–731.

Tauchen, G. and R. Hussey (1991): "Quadrature-Based Methods for Obtaining Approximate Solutions to Nonlinear Asset Pricing Models," *Econometrica*, 59, 317–396.

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

Zhou, G. and Y. Zhu (2015): "Macroeconomic Volatilities and Long-Run Risks of Asset Prices," *Management Science*, 61, 413–430.