Risk Matters: Breaking Certainty Equivalence in Linear Approximations

Juan Carlos Parra-Alvarez[a]
Aarhus University

Hamza Polattimur[b]
Universität Hamburg

Olaf Posch[c]
Universität Hamburg

June 5, 2021

Abstract

In this paper we use the property that certainty equivalence, as implied by a first-order approximation to the solution of stochastic discrete-time models, breaks in its equivalent continuous-time version. We derive a risk-sensitive first-order perturbation solution for a general class of rational expectations models. We show that risk matters economically in a real business cycle (RBC) model with habit formation and capital adjustment costs, and that neglecting risk leads to substantial pricing errors. A first-order perturbation provides a sensible approximation to the effects of risk in continuous-time models. It reduces pricing errors by around 90 percent relative to the certainty equivalent linear approximation.

Keywords: Certainty equivalence, Perturbation methods, Pricing errors.
JEL classification: C02, C61, C63, E13, E32, G12.

[a] *Corresponding author.* Department of Economics and Business Economics, Aarhus University; CREATES; Danish Finance Institute; Fuglesangs Allé 4, 8210 Aarhus V, Denmark; Email address: jparra@econ.au.dk

[b] Faculty of Business, Economics and Social Sciences, Universität Hamburg; Von-Melle-Park 5, 20146 Hamburg, Germany; Email address: Hamza.Polattimur@uni-hamburg.de

[c] Faculty of Business, Economics and Social Sciences, Universität Hamburg; CREATES; Von-Melle-Park 5, 20146 Hamburg, Germany; Email address: Olaf.Posch@uni-hamburg.de

1 Introduction

There is a consensus among economists that uncertainty affects the consumption-saving decision of individuals. Neglecting the effects of risk in macroeconomics and finance often generates substantial pricing errors. Hence, recent research is concerned with the ability of local approximations of nonlinear stochastic macroeconomic models to account for risk, with a particular focus on perturbation methods originally introduced in economics by Judd and Guu (1993). Although perturbation-based methods only provide local precision around a particular point, usually the model's deterministic steady state, many authors suggest that they can generate high levels of accuracy, comparable to that delivered by global approximation techniques, as the order of the approximation is increased (see Judd, 1998; Aruoba et al., 2006; Caldara et al., 2012; Parra-Alvarez, 2018). In many applications, however, we are interested in the first-order perturbation and the resulting linear approximation of the equilibrium conditions. A linear structure not only provides analytical insights and helps to understand key features of the model, but also facilitates its estimation, e.g., by means of the Kalman filter (see Harvey and Stock, 1985; Singer, 1998; Harvey, 2006; Fernández-Villaverde and Rubio-Ramírez, 2007).

A known limitation of the first-order perturbation around the deterministic steady state is that the approximate solution of discrete-time models typically exhibits certainty equivalence (see Simon, 1956; Theil, 1957). In other words, the first-order approximation to the solution of stochastic economic models with forward-looking agents is identical to the solution of the same model under perfect foresight. The direct implication is that the solution becomes invariant to higher-order moments of the underlying shocks. Therefore, this paper addresses the following questions. What are the costs of neglecting the effects of risk in linear approximations? Put differently, what would be the benefits of using a linear approximation that is not certainty equivalent? In particular, by how much could such an approximation reduce the errors that one makes when not accounting for risk? How can these errors be interpreted in an economically meaningful way?

Certainty equivalence prevails in the classical linear-quadratic optimal control problem, popularized in economics by Kydland and Prescott (1982) and Anderson et al. (1996). In the early 1950s the introduction of certainty equivalent stochastic control problems with quadratic utility and linear constraints aimed at providing a practical solution for decision problems under uncertainty. Even today, if risk is negligible for the research question at hand, certainty equivalent solutions are still useful. In this case, one may conclude that "certainty equivalence is a virtue" (see Kimball, 1990a). Conversely, when there is a reason to believe that the effects of risk are important, one notices that "certainty equivalence is a vice". Thus, if risk matters, breaking certainty equivalence is desired in order to account for the effects of risk. As discussed in Fernández-Villaverde et al. (2016), the approximated solution of the model under certainty equivalence (i) makes it difficult to talk about the welfare effects of uncertainty; (ii) cannot generate any risk premia for assets; and (iii) prevents analyzing the consequences of changes in volatility.

To break the property of certainty equivalence in the class of perturbation methods, economists have restored to the computation of higher-order Taylor expansions, the underlying apparatus behind any perturbation method, which translate into nonlinear approximations of the model's solution. Originally proposed in Judd and Guu (1993), higher-order approximations became popular with the work of Schmitt-Grohe and Uribe (2004) for second-order approximations, and that of Andreasen (2012) and Ruge-Murcia (2012) for third-order approximations. More recently, Levintal (2017) extended the perturbation package to include fifth-order approximations. However, the use of high-order approximations for medium-scale macroeconomic models (i) could be computationally expensive, (ii) often results in explosive solutions, and (iii) requires computationally demanding nonlinear estimation methods, such as the particle filter, for the estimation of the model's structural parameters.[1]

In contrast to stochastic discrete-time models, certainty equivalence breaks in the first-order approximation when time is assumed to be continuous (see Judd, 1996; Gaspar and Judd, 1997). This property, which allows us to account for risk in a linear world, is the product of two complementary results. First, while in discrete time the approximation is built inside the system of expectational equations that collects the equilibrium conditions of the economy, in continuous time we may use Itô's lemma to eliminate the expectation operator prior to the construction of the approximation. The resulting *non-expectational* system of equations, although deterministic, will capture the effects of uncertainty by including information on the sensitivity to risk of the yet unknown solution (see Chang, 2009). Second, as shown in Fleming (1971), who provides the mathematical foundations of perturbation methods for continuous-time stochastic optimal control problems, regular perturbation theory produces asymptotically valid approximations of the unknown solution when the variance of the shocks is used as perturbation parameter. As discussed in Jin and Judd (2002), this choice of the perturbation parameter follows basic economic intuition, whereby optimal behavior of agents is affected by the variance in the economy. This is in contrast to discrete-time models where the appropriate perturbation parameter is shown to be the standard deviation (cf. Judd, 1998, Jin and Judd, 2002, Fernández-Villaverde et al., 2016). Combining these two results, the linear approximation to the model's solution, which results from a first-order perturbation around the deterministic steady state, will exhibit a constant correction term that depends on the variance of the shocks that drive the dynamics of the economy.

In this paper, we revisit the ability of a first-order perturbation to capture the ef-

[1]See Kim et al. (2008) for a discussion on point (ii). Lan and Meyer-Gohde (2013) introduce a nonlinear infinite moving-average representation of the policy functions that eliminates the possibility for explosive behavior in higher-order approximations. Meyer-Gohde (2015) shows that it is possible to overcome (iii) by using linear approximations around the stochastic steady state or ergodic mean.

fects of risk. First, we derive a first-order perturbation solution for a general class of dynamic, continuous-time, rational expectations models, thereby formalizing the framework in Gaspar and Judd (1997) and Parra-Alvarez (2018). We show analytically that, as opposed to discrete-time models, the first derivative of the policy function with respect to the perturbation parameter is different from zero at the deterministic steady state implying that the resulting linear approximation is risk-sensitive, i.e., it breaks the certainty-equivalence property. More specifically, the first-order perturbation corrects for risk through an additional constant term that incorporates information on the slope and curvature of the optimal policy functions at the deterministic steady state. Second, we explore how the effects of uncertainty are internalized by this perturbation approach using as a benchmark an RBC model with habit formation and capital adjustment costs *à la* Jermann (1998). By calibrating the parameters of the model to values that are standard in the literature, we compare, along different dimensions, the first-order certainty equivalent (CE) and the risk-sensitive first-order approximations obtained from perturbation to a second-order nonlinear approximation and a global solution obtained by collocation methods. We show that each of the approximations converges to different long-run equilibria in the absence of shocks. While the first-order CE converges to the deterministic steady state, the risk-adjusted solutions converge to their respective risky steady states. This property is reflected in the policy and impulse response functions.

We find that the risk effects captured by the first-order approximation in continuous time are economically significant. We assess the asset pricing implications of the approximations using a partial differential equation approach rather than the standard simulation approach used in discrete time. When relying on the linear CE solution, the pricing errors made on a three month zero-coupon bond are about 1 dollar for each 100 dollar spent at the deterministic steady state. The risk-adjustment of the first-order perturbation approximation leads to errors of about 10 cents for each 100 dollar spent, reducing pricing errors by about 90%. Hence, we conclude that the risk-sensitive first-order perturbation provides a sensible approximation to the effects of risk in continuous-time models.

Our work relates to that of Collard and Juillard (2001), Coeurdacier et al. (2011), de Groot (2013), Meyer-Gohde (2015) and Lopez et al. (2018), who compute first-order approximations around the model's risky steady state instead of the deterministic steady state in order to break certainty equivalence in discrete-time models. Collard and Juillard (2001) consider a bias reduction procedure to compute the approximation around the risky steady state; Coeurdacier et al. (2011), whose approach is generalized by de Groot (2013), approximate the risky steady state based on the second-order solution. Meyer-Gohde (2015) constructs a risk-sensitive linear approximation using policy functions resulting from higher-order perturbations. Lopez et al. (2018) differ from the previous studies as they consider lognormal affine approximations, often used in macro-finance, which are shown to be a special case of a first-order perturbation around the risky steady state. We argue that it is possible to account for risk in an economically meaningful way using *standard* first-order (*linear*) perturbations around the *deterministic* steady state when time evolves continuously.

The rest of the paper is organized as follows. In Section 2, we introduce our model and define the equilibrium conditions used in the perturbation method to approximate the solution. Section 3 summarizes the perturbation approach, presents the main theoretical contribution of the paper, and revisits the certainty equivalence property of linear models. Section 4 derives the pricing implications of the approximated solution and introduces a pricing error measure that can be used to evaluate the accuracy of the approximation. Section 5 discusses the main results by comparing policy functions, impulse-response functions, and pricing errors for the different approximations. Finally, Section 6 concludes.

2 A prototype RBC model

For illustration, we use a continuous-time version of the real business cycle (RBC) model introduced in Jermann (1998). There is a single good in the economy that is produced using a constant-returns-to-scale production technology that is subject to random shocks in productivity. Changes in the economy's aggregate capital stock are subject to adjustment costs, and the household preferences exhibit intertemporal non-separabilities due to internal habit formation in consumption.

Preferences. The economy is inhabited by a large number of identical households that maximize their expected discounted lifetime utility from consumption, C_t,

$$U_0 := \mathbb{E}_0 \left[\int_0^\infty e^{-\rho t} u\left(C_t, X_t\right) \mathrm{d}t \right], \tag{1}$$

where $\mathbb{E}_0\left[\cdot\right]$ is the expectation operator conditional on the information available at time $t = 0$, $\rho \geq 0$ is the household's subjective discount rate, and u is the instantaneous utility function. For simplicity we assume that

$$u\left(C_t, X_t\right) = \frac{\left(C_t - X_t\right)^{1-\gamma}}{1-\gamma}, \tag{2}$$

where γ measures the curvature of the utility function (together with the consumption surplus ratio), such that a higher value of γ yields higher risk aversion. In what follows, we assume that the consumption choice is non-negative, $C_t \geq 0$, and does not fall below a subsistence level of consumption, $C_t \geq X_t$, where X_t denotes habits in consumption. The habit level in consumption is defined endogenously (internal habit) in the model, in contrast to the relative consumption model or 'catching up with the Joneses' (external habit), where habit is given by aggregate consumption and is thus exogenous to the households. In particular, the habit process is given by $X_t = e^{-at}X_0 + b\int_0^t e^{a(s-t)}C_s \mathrm{d}s$, or equivalently,

$$\mathrm{d}X_t = (bC_t - aX_t)\mathrm{d}t, \quad \text{with } X(0) = X_0 \geq 0 \text{ given.} \tag{3}$$

Hence, X_t is a weighted sum of past consumption, with weights declining exponentially into the past. The parameter a measures the degree of persistence in the habit stock, such that the larger is a, the less weight is given to past consumption in determining X_t, and vice versa. The parameter b is a scaling parameter. The special case $b = X_0 = 0$ corresponds to the case of time-separable utility with constant relative risk aversion (see Constantinides, 1990).

Technology. The one good in the economy is produced according to the Cobb-Douglas production function

$$Y_t = \exp(A_t)K_t^{\alpha}L_t^{1-\alpha}, \quad 0 < \alpha < 1, \tag{4}$$

where K_t is the aggregate capital stock, L_t is the perfectly inelastic labor supply (normalized to one for all $t \geq 0$), and A_t is a stochastic process that drives total factor productivity (TFP). The aggregate capital stock in the economy increases if effective investment exceeds depreciation

$$\mathrm{d}K_t = (\Phi(I_t/K_t) - \delta)K_t\mathrm{d}t, \quad \text{with } K(0) = K_0 > 0 \text{ given,} \tag{5}$$

where $\delta \geq 0$ is the depreciation rate, and I_t is aggregate investment. Following Jermann (1998), the capital adjustment cost function is defined by

$$\Phi(I_t/K_t) = \frac{a_1}{1 - 1/\xi}\,(I_t/K_t)^{1-1/\xi} + a_2, \tag{6}$$

where $\xi > 0$ denotes the elasticity of the investment-to-capital ratio with respect to Tobin's q, and $a_1 \geq 0$ and $a_2 \geq 0$ are parameters. In line with Boldrin et al. (2001), we set $a_1 = \delta^{1/\xi}$ and $a_2 = \delta/(1-\xi)$ such that the steady state is invariant to ξ, and hence the long-run investment-to-capital ratio equals the deprecation rate.[2] The process A_t is assumed to follow an Ornstein-Uhlenbeck process with mean reversion $\rho_A > 0$ and variance $\sigma_A > 0$

$$\mathrm{d}A_t = -\rho_A A_t\mathrm{d}t + \sigma_A\mathrm{d}B_{A,t}, \quad \text{with } A(0) = A_0 \in \mathbb{R} \text{ given,} \tag{7}$$

where $B_{A,t}$ is a standard Brownian motion. In equilibrium, the economy satisfies the aggregate resource constraint

$$Y_t = C_t + I_t. \tag{8}$$

[2] Given this parameterization it can be shown that in the steady state: $\Phi(\bar{I}/\bar{K}) = \Phi(\delta) = \delta$, $\Phi'(\bar{I}/\bar{K}) = \Phi'(\delta) = 1$, and $\Phi''(\bar{I}/\bar{K}) = \Phi''(\delta) = -1/(\xi\delta)$, i.e. the slope of Φ' depends negatively on ξ and δ.

Optimality conditions. Consider the problem faced by a social planner who has to choose the path of consumption that maximizes (1) subject to the dynamic constraints (3), (5), and (7), and the static constraints (4), (6), and (8)

$$V(K_0, X_0, A_0) = \max_{\{C_t \geq X_t \in \mathbb{R}^+\}_{t=0}^{\infty}} U_0 \quad \text{s.t.} \quad (3) - (8), \tag{9}$$

in which C_t is the control variable at time $t \in \mathbb{R}^+$, and $V(K_0, X_0, A_0)$ is the value of the optimal plan (value function) from the perspective of time $t = 0$, i.e., when the state of the economy is described by the time $t = 0$ values for the capital stock, K_0, the stock of habits, X_0, and the productivity process, A_0.

As shown in Appendix A, the *Hamilton-Jacobi-Bellman* (HJB) equation associated with the stochastic optimal control problem in (9) is

$$\rho V = \max_{C_t \geq X_t \in \mathbb{R}^+} \left\{ \frac{(C_t - X_t)^{1-\gamma}}{1-\gamma} + (\Phi((\exp(A_t)K_t^{\alpha} - C_t)/K_t) - \delta)K_t V_K \right. \\ \left. + (bC_t - aX_t)V_X - \rho_A A_t V_A + \tfrac{1}{2}\sigma_A^2 V_{AA} \right\}, \tag{10}$$

where $V_K := \partial V(K_t, X_t, A_t)/\partial K_t$, $V_X := \partial V(K_t, X_t, A_t)/\partial X_t$, $V_A := V(K_t, X_t, A_t)/\partial A_t$, and $V_{AA} := \partial^2 V(K_t, X_t, A_t)/\partial A_t^2$ denote the first- and the second-order partial derivatives of the value function $V := V(K_t, X_t, A_t)$ with respect to the states of the economy.[3] The first-order condition for any interior solution reads

$$(C_t - X_t)^{-\gamma} + bV_X = \Phi'\left(\frac{\exp(A_t)K_t^{\alpha} - C_t}{K_t}\right) V_K, \tag{11}$$

making optimal consumption a function of the state variables, $C_t = C(K_t, X_t, A_t)$. The function $C(\cdot)$ maps every possible value of the states of the economy at time t into optimal consumption at time t.

Competitive equilibrium. The general equilibrium in this economy is characterized by the maximized HJB equation

$$\rho V = \frac{(C(K_t, X_t, A_t) - X_t)^{1-\gamma}}{1-\gamma} + (\Phi((\exp(A_t)K_t^{\alpha} - C(K_t, X_t, A_t))/K_t) - \delta)K_t V_K \\ + (bC(K_t, X_t, A_t) - aX_t)V_X - \rho_A A_t V_A + \tfrac{1}{2}\sigma_A^2 V_{AA}, \tag{12}$$

which provides a necessary condition for optimality. Given the dynamics for the state variables in (3), (5), and (7), a solution to the stochastic optimal control problem in (9)

[3] A formal introduction and derivation of the dynamic programming equation in continuous-time can be found in Chang (2009). In the Online Appendix, we provide a step-by-step formulation of the stochastic optimal control problem and the HJB equation in a multidimensional framework.

is given by the unknown value function, $V(K_t, X_t, A_t)$, and the consumption function, $C(K_t, X_t, A_t)$, that solve the functional problem formed by (11) and (12).

A solution to (9) can be alternatively characterized by the tuple $\{V_K, V_X, V_A, C\} := \{V_K(K_t, X_t, A_t), V_X(K_t, X_t, A_t), V_A(K_t, X_t, A_t), C(K_t, X_t, A_t)\}$ that solves the system of second-order quasilinear partial differential equations (PDEs)

$$
\begin{aligned}
0 &= (\rho - \Phi((\exp(A_t)K_t^\alpha - C_t)/K_t) - \Phi'((\exp(A_t)K_t^\alpha - C_t)/K_t)((\alpha-1)\exp(A_t)K_t^{\alpha-1} \\
&\quad + C_t/K_t) + \delta)V_K - (\Phi((\exp(A_t)K_t^\alpha - C_t)/K_t) - \delta)K_t V_{KK} \\
&\quad - (bC_t - aX_t)V_{XK} + \rho_A A_t V_{AK} - \tfrac{1}{2}\sigma_A^2 V_{AAK}, && (13) \\
0 &= (\rho + a)V_X + (C_t - X_t)^{-\gamma} - (\Phi((\exp(A_t)K_t^\alpha - C_t)/K_t) - \delta)K_t V_{KX} \\
&\quad - (bC_t - aX_t)V_{XX} + \rho_A A_t V_{AX} - \tfrac{1}{2}\sigma_A^2 V_{AAX}, && (14) \\
0 &= (\rho_A + \rho)V_A - \Phi'((\exp(A_t)K_t^\alpha - C_t)/K_t)\exp(A_t)K_t^\alpha V_K \\
&\quad - (\Phi((\exp(A_t)K_t^\alpha - C_t)/K_t) - \delta)K_t V_{KA} - (bC_t - aX_t)V_{XA} \\
&\quad + \rho_A A_t V_{AA} - \tfrac{1}{2}\sigma_A^2 V_{AAA}, && (15) \\
0 &= (C_t - X_t)^{-\gamma} + bV_X - \Phi'((\exp(A_t)K_t^\alpha - C_t)/K_t)V_K. && (16)
\end{aligned}
$$

for any admissible values of the state variables (K_t, X_t, A_t). Equations (13)-(15), which are obtained by differentiating (12) with respect to the state variables and by the application of the envelope theorem, characterize the optimal behavior of the costate variables, V_K, V_X, and V_A in the state space. Together with (16), they define a set of necessary conditions for optimality. A detailed derivation of (13)-(16) can be found in Appendix A.

Deterministic steady state. In the absence of uncertainty (i.e., $\sigma_A = 0$), the economy converges over time to a fixed point or deterministic steady state (DSS) in which all variables are idle. Given the assumptions on the capital adjustment cost function (6), the DSS is given by $\overline{A} = 0$, $\overline{K} = [\alpha/(\rho+\delta)]^{\frac{1}{1-\alpha}}$, $\overline{C} = \overline{K}^\alpha - \delta\overline{K}$, $\overline{X} = (b/a)\overline{C}$, $\overline{V_X} = -[1/(\rho+a)]\left(\overline{C} - \overline{X}\right)^{-\gamma}$, $\overline{V_K} = [1 - b/(\rho+a)]\left(\overline{C} - \overline{X}\right)^{-\gamma}$, and $\overline{V_A} = \overline{K}^\alpha \overline{V_K}/(\rho_A + \rho)$, where $\overline{V_K}$, $\overline{V_X}$ and $\overline{V_A}$ are the DSS values of the costate variables for the capital stock, habit formation, and productivity, respectively. For a detailed derivation of the model's DSS see Appendix A.

3 Approximate solution

Most dynamic economic models do not admit an analytical solution, so it usually has to be approximated using numerical methods. Perturbation methods are fast and reliable, and provide an approximate solution to the stochastic optimal control problem in (9) based on the implicit function theorem and Taylor's theorem. The perturbed solution consists of a polynomial that approximates the true solution of the problem locally in a neighborhood of an *a priori* known solution. In what follows, we build the perturbation solution to the equilibrium system of PDEs in (13)-(16) around the DSS.

Let $\eta > 0$ denote a *perturbation parameter* that rescales the amount of risk in the economy. For continuous-time stochastic optimal control problems, Fleming (1971) showed that by choosing η to control the *variance* of the exogenous disturbances, it is possible to use regular perturbation theory to obtain asymptotically valid approximations to the unknown policy functions (see Judd, 1996; Gaspar and Judd, 1997).[4] For the model in Section 2 this amounts to write the exogenous stochastic processes that defines TFP (7) as

$$\mathrm{d}A_t = -\rho_A A_t \mathrm{d}t + \sqrt{\eta\sigma_A^2}\mathrm{d}B_{A,t} = -\rho_A A_t \mathrm{d}t + \sqrt{\eta}\sigma_A \mathrm{d}B_{A,t},$$

where $\eta = 0$ makes the model deterministic and $\eta = 1$ recovers the true stochastic process in (7).

Following Judd (1998), the perturbation method can be summarized as follows:

1. Express the problem of interest as a continuum of problems parameterized by the added perturbation parameter η, with the $\eta = 0$ case known.

2. Differentiate the continuum of problems with respect to the state variables and the perturbation parameter η.

3. Solve the resulting equation for the implicitly defined derivatives at the known solution of the state variables and $\eta = 0$.

4. Compute the desired order of approximation by means of Taylor's theorem. Set $\eta = 1$ to recover the approximation to the original model.

In what follows, we present a framework that generalizes the perturbation method in Gaspar and Judd (1997) and Parra-Alvarez (2018) for a class of dynamic, continuous-time, rational expectations models.[5] Following Gomme and Klein (2011), Binning (2013a,b) and Lan and Meyer-Gohde (2013, 2014), we avoid the use of tensor notation and instead derive our perturbation solution using standard methods from linear algebra based on the rules for multidimensional calculus in Vetter (1973).[6] We then provide an illustrative example of the method by using a simplified version of our prototype model. Subsequently, we explain why the property of certainty equivalence, that usually results from

[4]This is in contrast to discrete-time stochastic problems, where the perturbation parameter rescales the *standard deviation* of the shocks (see Schmitt-Grohe and Uribe, 2004; Fernández-Villaverde et al., 2016). Choosing instead the variance of the shocks as the perturbation parameter produces approximations with undesirable stochastic properties in discrete time as shown in Jin and Judd (2002).

[5]A brief description of the perturbation method in discrete time is provided in the Online Appendix.

[6]The Online Appendix provides a summary with definitions of all the matrix structures, as well as the rules for differentiation used throughout the paper.

any first-order perturbation approximation to discrete-time models, breaks in continuous-time models. Finally, we introduce the notion of the risky steady state (RSS), which will become relevant for understanding the transition paths of the model's variables.

3.1 Solving the model: A general framework

3.1.1 Problem statement

We consider continuous-time dynamic stochastic models where the rational expectations equilibrium can be summarized by the n-dimensional vector-valued functional

$$\widetilde{\mathcal{H}}\left(\mathbf{x}, \mathbf{u}, \mathbf{y}, \mathbf{y}_{\mathbf{x}}, \mathbf{y}_{\mathbf{xx}}\right) = \mathbf{0}, \tag{17}$$

where $\mathbf{x} \in \mathbb{X}$ is an $n_x \times 1$ vector of state variables from the state space $\mathbb{X} \subseteq \mathbb{R}^{n_x}$, $\mathbf{u} \in \mathbb{U}$ is an $n_u \times 1$ vector of control variables from the control region $\mathbb{U} \subseteq \mathbb{R}^{n_u}$, $\mathbf{y} \in \mathbb{Y} \subseteq \mathbb{R}^{n_y}$ is an $n_y \times 1$ a vector of non-control and non-state variables (e.g., the optimal value function and/or costate variables). In the following, let $n = n_y + n_u$. Moreover, $\mathbf{y}_{\mathbf{x}} := \mathcal{D}_{\mathbf{x}^\top}\{\mathbf{y}\}$ is an $n_y \times n_x$ matrix of first-order partial derivatives, and $\mathbf{y}_{\mathbf{xx}} := \mathcal{D}_{(\mathbf{x}^\top)^2}\{\mathbf{y}\}$ is the corresponding $n_y \times n_x^2$ matrix of second-order partial derivatives. In general, we let $\mathbf{y}_{\mathbf{x}^k} := \mathcal{D}_{(\mathbf{x}^\top)^k}\{\mathbf{y}\}$ denote the $n_y \times n_x^k$ matrix containing the k-th order derivative of $\mathbf{y}$ with respect to the transpose of the state vector, $\mathbf{x}^\top$.

The state vector can be partitioned as $\mathbf{x} = \left[\mathbf{x}_1^\top; \mathbf{x}_2^\top\right]^\top$, where $\mathbf{x}_1$ is an $(n_x - n_w) \times 1$ vector of endogenous state variables, and $\mathbf{x}_2$ is an $n_w \times 1$ vector of exogenous state variables. Let $\tilde{n}_x = n_x - n_w$. The state variables $\mathbf{x}_1$ and $\mathbf{x}_2$ are assumed to evolve over time according to the following system of controlled stochastic differential equations (SDEs)

$$\begin{aligned}
\mathrm{d}\mathbf{x}_1 &= \mathbf{b}_1\left(\mathbf{x}, \mathbf{u}; \eta\right) \mathrm{d}t + \sqrt{\eta}\boldsymbol{\sigma}_1\left(\mathbf{x}, \mathbf{u}\right) \mathrm{d}\mathbf{w}_1, \quad \mathbf{x}_1(0) = \mathbf{x}_{10} \text{ given}, \\
\mathrm{d}\mathbf{x}_2 &= \mathbf{b}_2\left(\mathbf{x}_2; \eta\right) \mathrm{d}t + \sqrt{\eta}\boldsymbol{\sigma}_2\left(\mathbf{x}_2\right) \mathrm{d}\mathbf{w}_2, \quad \mathbf{x}_2(0) = \mathbf{x}_{20} \text{ given},
\end{aligned}$$

where $\mathbf{b}_1\left(\mathbf{x}, \mathbf{u}; \eta\right)$ and $\mathbf{b}_2\left(\mathbf{x}_2; \eta\right)$ are, respectively, the $\tilde{n}_x \times 1$ and $n_{w_2} \times 1$ real-valued drift vector functions; $\boldsymbol{\sigma}_1\left(\mathbf{x}, \mathbf{u}\right)$ is the $\tilde{n}_x \times n_{w_1}$ real-valued diffusion matrix for the endogenous states, which can potentially depend on the state and/or the control variables (see, e.g., Merton, 1971, Steger, 2005, Wälde, 2011, Brunnermeier and Sannikov, 2014, and Christensen et al., 2016); $\boldsymbol{\sigma}_2\left(\mathbf{x}_2\right)$ is the $n_{w_2} \times n_{w_2}$ real-valued diffusion matrix for the exogenous states; and $\mathbf{w} = \left[\mathbf{w}_1^\top; \mathbf{w}_2^\top\right]^\top$ is an $n_w \times 1$ vector of mutually independent standard Brownian motions that represent zero-mean exogenous innovations with $\mathbf{w}_1$ and $\mathbf{w}_2$ denoting $n_{w_1} \times 1$ and $n_{w_2} \times 1$ vectors so that $n_w = n_{w_1} + n_{w_2}$. The joint dynamics of the state vector $\mathbf{x}$ can be compactly written as

$$\mathrm{d}\mathbf{x} = \mathbf{b}\left(\mathbf{x}, \mathbf{u}; \eta\right) \mathrm{d}t + \sqrt{\eta}\boldsymbol{\sigma}\left(\mathbf{x}, \mathbf{u}\right) \mathrm{d}\mathbf{w}, \quad \mathbf{x}(0) = \mathbf{x}_0 \text{ given}, \tag{18}$$

where $\mathbf{b} = [\mathbf{b}_1^\top; \mathbf{b}_2^\top]^\top$ is the $n_x \times 1$ real-valued drift vector, and $\boldsymbol{\sigma} = \left[\boldsymbol{\sigma}_1^\top; \boldsymbol{\sigma}_2^\top\right]^\top$ is the $n_x \times n_w$ real-valued diffusion matrix determining the $n_x \times n_x$ variance-covariance matrix $\boldsymbol{\Sigma}(\mathbf{x}, \mathbf{u}) = \boldsymbol{\sigma}(\mathbf{x}, \mathbf{u}) \boldsymbol{\sigma}(\mathbf{x}, \mathbf{u})^\top$ of the state variables.

Remark 1. We allow the drift of the state vector in (18) to depend on the perturbation parameter, $\eta \geq 0$. This accommodates situations in which functionals of the state vector induce risk corrected drifts. For example, if the dynamics of a state variable follows an Ornstein-Uhlenbeck process in logs, $\mathrm{d}\log x = -\rho \log x \mathrm{d}t + \sqrt{\eta}\sigma \mathrm{d}w$, then Itô's lemma implies that the level of x is given by $\mathrm{d}x = -\left(\rho \log x - \frac{1}{2}\eta\sigma^2\right) x_t \mathrm{d}t + \sqrt{\eta}\sigma x \mathrm{d}w$, where the drift includes a risk correction term and thus depends on the perturbation parameter. ■

The vector of optimal control variables is, in general, defined implicitly by the algebraic first-order conditions to the stochastic optimal control problem, and it can be expressed in terms of the state and non-state variables, i.e., $\mathbf{u} = \mathcal{U}(\mathbf{x}, \mathbf{y}, \mathbf{y}_\mathbf{x})$, with $\mathcal{U} : \mathbb{X} \times \mathbb{Y} \times \mathbb{Y} \to \mathbb{U}$. Similar to Anderson et al. (1996) and Gaspar and Judd (1997), we simplify our approach by making use of these optimality conditions to substitute the control variables from the problem, so that (17) is reduced to an n_y-dimensional vector-valued system of equilibrium conditions in $\mathbf{x}$ and $\mathbf{y}$ only, i.e.,

$$\widetilde{\mathcal{H}}(\mathbf{x}, \mathcal{U}(\mathbf{x}, \mathbf{y}, \mathbf{y}_\mathbf{x}), \mathbf{y}, \mathbf{y}_\mathbf{x}, \mathbf{y}_{\mathbf{xx}}) = \mathcal{H}(\mathbf{x}, \mathbf{y}, \mathbf{y}_\mathbf{x}, \mathbf{y}_{\mathbf{xx}}) = \mathbf{0}. \tag{19}$$

By the same token, the drift vector and diffusion matrices that characterize the dynamics of the state vector can be written as $\mathbf{b}(\mathbf{x}, \mathcal{U}(\mathbf{x}, \mathbf{y}, \mathbf{y}_\mathbf{x}); \eta) := \mathbf{b}(\mathbf{x}, \mathbf{y}, \mathbf{y}_\mathbf{x}; \eta)$ and $\boldsymbol{\sigma}(\mathbf{x}, \mathcal{U}(\mathbf{x}, \mathbf{y}, \mathbf{y}_\mathbf{x})) := \boldsymbol{\sigma}(\mathbf{x}, \mathbf{y}, \mathbf{y}_\mathbf{x})$, respectively.

3.1.2 Model class

For a wide class of dynamic macroeconomic models, the rational expectation equilibrium in (19) can be represented by an $n_y \times 1$ system of second-order *quasilinear* PDEs of the form

$$\mathcal{H}(\mathbf{x}, \mathbf{y}, \mathbf{y}_\mathbf{x}, \mathbf{y}_{\mathbf{xx}}; \eta) := \mathbf{a}(\mathbf{x}, \mathbf{y}, \eta\mathbf{y}_\mathbf{x}) + \mathbf{y}_\mathbf{x}\mathbf{b}(\mathbf{x}, \mathbf{y}, \eta\mathbf{y}_\mathbf{x}; \eta) + \eta\mathbf{y}_{\mathbf{xx}}\mathbf{c}(\mathbf{x}, \mathbf{y}, \eta\mathbf{y}_\mathbf{x}) = \mathbf{0}, \tag{20}$$

where $\mathbf{a} \in \mathbb{R}^{n_y}$, $\mathbf{b} \in \mathbb{R}^{n_x}$, and $\mathbf{c} \in \mathbb{R}^{n_x^2}$ are continuous functions of the state and the non-state variables, with $\mathbf{x}$ evolving over time according to (18). Their dependence on $\eta\mathbf{y}_\mathbf{x}$ accommodates any potential effect of the variance-covariance matrix of the shocks, $\boldsymbol{\Sigma}$.

Hereafter, let the vector-valued function $\mathbf{y}$ contain only costate variables, so that $n_y = n_x$. Then, the system in (20) stacks all the PDEs that characterize the optimal behavior of the costate variables associated with a discounted stochastic optimal control problem, with discount rate $\rho \geq 0$. These equations correspond to the first-order derivatives of the maximized Hamilton-Jacobi-Bellman equation with respect to the state variables, and provide a set of necessary conditions for optimality. Moreover, as shown in the Online Appendix, $\mathbf{a}(\mathbf{x},\mathbf{y},\eta\mathbf{y}_{\mathbf{x}})$ and $\mathbf{c}(\mathbf{x},\mathbf{y},\eta\mathbf{y}_{\mathbf{x}})$ can be decomposed as

$$\mathbf{a}(\mathbf{x},\mathbf{y},\eta\mathbf{y}_{\mathbf{x}}) = \widetilde{\mathbf{a}}(\mathbf{x},\eta\mathbf{y}_{\mathbf{x}}) + \mathbf{b}_{\mathbf{x}}^{\top}(\mathbf{x},\mathbf{y},\eta\mathbf{y}_{\mathbf{x}};\eta)\,\mathbf{y} - \rho\mathbf{y}, \tag{21}$$

$$\mathbf{c}(\mathbf{x},\mathbf{y},\eta\mathbf{y}_{\mathbf{x}}) = \tfrac{1}{2}\mathrm{vec}\,\boldsymbol{\Sigma}(\mathbf{x},\mathbf{y},\eta\mathbf{y}_{\mathbf{x}}), \tag{22}$$

where $\widetilde{\mathbf{a}}$ is an $n_x \times 1$ vector, and $\mathbf{b}_{\mathbf{x}} := \mathcal{D}_{\mathbf{x}^{\top}}\{\mathbf{b}\}$ is an $n_x \times n_x$ matrix.

Remark 2. For $\boldsymbol{\Sigma}$ independent of the state and control variables, and hence constant, the functions $\mathbf{a}$, $\mathbf{b}$ and $\mathbf{c}$ in (20) do not depend on $\eta\mathbf{y}_{\mathbf{x}}$. All the effects from risk are captured through the resulting constant $\mathbf{c}$ which is determined by the constant variance-covariance matrix $\boldsymbol{\Sigma}$. This is the case for the prototype RBC economy in Section 2, where (20) simplifies to

$$\mathcal{H}(\mathbf{x},\mathbf{y},\mathbf{y}_{\mathbf{x}},\mathbf{y}_{\mathbf{xx}};\eta) := \mathbf{a}(\mathbf{x},\mathbf{y}) + \mathbf{y}_{\mathbf{x}}\mathbf{b}(\mathbf{x},\mathbf{y}) + \eta\mathbf{y}_{\mathbf{xx}}\mathbf{c} = \mathbf{0},$$

with $\mathbf{y} = [V_K, V_X, V_A]^{\top}$ and $\mathbf{x} = [K_t, X_t, A_t]^{\top}$. The explicit definitions of the matrices $\mathbf{a}(\mathbf{x},\mathbf{y})$, $\mathbf{b}(\mathbf{x},\mathbf{y})$, and $\mathbf{c}$ are shown in Appendix A. ■

A solution to the class of models defined in (20) is given by the set of policy functions

$$\mathbf{y} = \mathbf{g}(\mathbf{x};\eta), \tag{23}$$

with $\mathbf{g} : \mathbb{R}^{n_x+1} \to \mathbb{R}^{n_x}$ such that

$$\begin{aligned}\mathbf{F}(\mathbf{x};\eta) := \mathbf{a}(\mathbf{x},\mathbf{g}(\mathbf{x};\eta),\eta\mathbf{g}_{\mathbf{x}}(\mathbf{x};\eta)) + \mathbf{g}_{\mathbf{x}}(\mathbf{x};\eta)\,\mathbf{b}(\mathbf{x},\mathbf{g}(\mathbf{x};\eta),\eta\mathbf{g}_{\mathbf{x}}(\mathbf{x};\eta);\eta)\\ + \eta\mathbf{g}_{\mathbf{xx}}(\mathbf{x};\eta)\,\mathbf{c}(\mathbf{x},\mathbf{g}(\mathbf{x};\eta),\eta\mathbf{g}_{\mathbf{x}}(\mathbf{x};\eta)) = \mathbf{0},\end{aligned} \tag{24}$$

where $\mathbf{F} : \mathbb{R}^{n_x+1} \to \mathbb{R}^{n_x}$ is a new functional operator in the state space that results from the substitution of (23) into (20), i.e., $\mathbf{F}(\mathbf{x};\eta) := \mathcal{H}(\mathbf{x},\mathbf{g}(\mathbf{x};\eta),\mathbf{g}_{\mathbf{x}}(\mathbf{x};\eta),\mathbf{g}_{\mathbf{xx}}(\mathbf{x};\eta);\eta)$, with $\mathbf{g}_{\mathbf{x}} = \mathcal{D}_{\mathbf{x}^{\top}}\{\mathbf{g}\}$ and $\mathbf{g}_{\mathbf{xx}} = \mathcal{D}_{(\mathbf{x}^{\top})^2}\{\mathbf{g}\}$.

Once (23) is obtained, it is straightforward to compute the optimal controls as $\mathbf{u} = \mathcal{U}(\mathbf{x},\mathbf{g}(\mathbf{x};\eta),\eta\mathbf{g}_{\mathbf{x}}(\mathbf{x};\eta))$, and the optimally controlled states as the solution to the system of SDEs in (18), $\mathbf{x}_1 = \mathbf{x}_{1,0} + \int_0^t \mathbf{b}_1(\mathbf{x},\mathbf{u})\,\mathrm{d}s + \sqrt{\eta}\int_0^t \boldsymbol{\sigma}_1(\mathbf{x},\mathbf{u})\,\mathrm{d}\mathbf{w}_1$.

3.1.3 Perturbation solution

In general, the solution (23) is not available in closed form. Therefore, we approximate $\mathbf{g}(\mathbf{x};\eta)$ by means of a Taylor series expansion around the DSS of the model, which is defined by the fixed point $(\mathbf{x},\mathbf{y};\eta) = (\overline{\mathbf{x}},\overline{\mathbf{y}};0)$ that solves (19) in the absence of uncertainty.

More specifically, $(\overline{\mathbf{x}}, \overline{\mathbf{y}})$ is obtained as the solution to the extended system of equations

$$\begin{bmatrix} \mathcal{H}\left(\overline{\mathbf{x}}, \overline{\mathbf{y}}, \mathbf{0}, \mathbf{0}; 0\right) \\ \mathrm{d}\overline{\mathbf{x}}/\mathrm{d}t \end{bmatrix} := \begin{bmatrix} \mathbf{a}\left(\overline{\mathbf{x}}, \overline{\mathbf{y}}, \mathbf{0}\right) \\ \mathbf{b}\left(\overline{\mathbf{x}}, \overline{\mathbf{y}}, \mathbf{0}; 0\right) \end{bmatrix} = \begin{bmatrix} \mathbf{0} \\ \mathbf{0} \end{bmatrix}, \tag{25}$$

where the second equation imposes the idle, or no growth, condition on the vector of the state variables in the deterministic case.

Hence, a k-th order approximation to (23) around the DSS is given by (see Lan and Meyer-Gohde, 2014)

$$\mathbf{g}\left(\mathbf{x}; \eta\right) = \sum_{j=0}^{k} \frac{1}{j!} \left[\sum_{i=0}^{k-j} \frac{1}{i!} \mathcal{D}_{(\mathbf{x}^\top)^j \eta^i} \{\mathbf{g}\left(\overline{\mathbf{x}}; 0\right)\} \eta^i \right] \left(\mathbf{x} - \overline{\mathbf{x}}\right)^{\otimes [j]}, \tag{26}$$

where $\mathcal{D}_{(\mathbf{x}^\top)^j \eta^i}$ is the derivative of the policy function taken j times with respect to the state vector and i times with respect to the perturbation parameter, evaluated at the DSS.

First-Order Perturbation. For $k = 1$, we define the *First-Order* approximation of $\mathbf{g}\left(\mathbf{x}; \eta\right)$ around $\left(\mathbf{x}; \eta\right) = \left(\overline{\mathbf{x}}; 0\right)$ as

$$\mathbf{g}\left(\mathbf{x}; \eta\right) = \mathbf{g}\left(\overline{\mathbf{x}}; 0\right) + \mathbf{g}_{\mathbf{x}}\left(\overline{\mathbf{x}}; 0\right)\left(\mathbf{x} - \overline{\mathbf{x}}\right) + \mathbf{g}_{\eta}\left(\overline{\mathbf{x}}; 0\right)\eta. \tag{27}$$

By the definition of the DSS, it follows that $\overline{\mathbf{y}} = \mathbf{g}\left(\overline{\mathbf{x}}; 0\right)$. To compute the remaining $n_x \times n_x$ coefficients in $\overline{\mathbf{g}}_{\mathbf{x}} := \mathbf{g}_{\mathbf{x}}\left(\overline{\mathbf{x}}; 0\right)$ and the $n_x \times 1$ coefficients in $\overline{\mathbf{g}}_{\eta} := \mathbf{g}_{\eta}\left(\overline{\mathbf{x}}; 0\right)$ that define (27), we successively differentiate (24) with respect to $\mathbf{x}$ and η.

Proposition 1. *The matrix of coefficients $\overline{\mathbf{g}}_{\mathbf{x}}$ in (27) satisfies the $n_x \times n_x$ continuous-time algebraic Riccati equation (CARE)*

$$\mathbf{A}^\top \overline{\mathbf{g}}_{\mathbf{x}} + \overline{\mathbf{g}}_{\mathbf{x}} \mathbf{A} + \overline{\mathbf{g}}_{\mathbf{x}} \mathbf{C} \overline{\mathbf{g}}_{\mathbf{x}} + \mathbf{B} = \mathbf{0}, \tag{28}$$

where $\mathbf{A}$, $\mathbf{B}$, $\mathbf{C} \in \mathbb{R}^{n_x \times n_x}$. For discounted stochastic optimal control problems, with discount rate $\rho > 0$, the coefficient matrices are given by $\mathbf{A} = \left(\overline{\mathbf{b}}_{\mathbf{x}} - \frac{\rho}{2}\mathbf{I}_{n_x}\right)$, $\mathbf{C} = \overline{\mathbf{b}}_{\mathbf{y}}$, and $\mathbf{B} = \overline{\tilde{\mathbf{a}}}_{\mathbf{x}} + \overline{\mathbf{b}}_{\mathbf{xx}}\left(\mathbf{I}_{n_x} \otimes \overline{\mathbf{y}}\right)$, with $\mathbf{I}_{n_x}$ an identity matrix of size n_x. In addition, $\mathbf{C} = \mathbf{C}^\top$ and $\mathbf{B} = \mathbf{B}^\top$. A bar on top of a variable indicates its corresponding value at the DSS.

Proof. See Appendix B. ∎

To solve the matrix quadratic equation (28), we introduce the $2n_x \times 2n_x$ Hamiltonian matrix

$$\mathbf{H}_\infty = \begin{bmatrix} \mathbf{A} & \mathbf{C} \\ -\mathbf{B} & -\mathbf{A}^\top \end{bmatrix}. \tag{29}$$

Following Anderson et al. (1996), Zhou et al. (1996), and Hansen and Sargent (2014), the solution $\overline{\mathbf{g}}_{\mathbf{x}}$ to (28) is obtained by locating the stable invariant subspace of (29) by means of an ordered Jordan decomposition, i.e.,

$$\mathbf{H}_{\infty} = \mathbf{P} \begin{bmatrix} \mathbf{\Lambda}_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{\Lambda}_2 \end{bmatrix} \mathbf{P}^{-1}, \tag{30}$$

where $\mathbf{\Lambda}_1$ is a diagonal matrix containing all the stable eigenvalues of $\mathbf{H}_{\infty}$, i.e., all the eigenvalues with negative real parts, $\mathbf{\Lambda}_2$ is a diagonal matrix containing all the unstable eigenvalues, i.e., all the eigenvalues with positive real parts, and

$$\mathbf{P} = \begin{bmatrix} \mathbf{P}_{11} & \mathbf{P}_{12} \\ \mathbf{P}_{21} & \mathbf{P}_{22} \end{bmatrix},$$

is the matrix of corresponding eigenvectors partitioned into blocks of equal size. If the number of stable eigenvalues equals the number of state variables, $\#\mathbf{\Lambda}_1 = n_x$, the solution is said to be stable (see Buiter, 1984 and Judd, 1996) and is given by

$$\mathbf{g}_{\mathbf{x}}(\overline{\mathbf{x}}; 0) = -\mathbf{P}_{22}^{-1}\mathbf{P}_{21}, \tag{31}$$

which together with the initial condition $\mathbf{x}(0) = \mathbf{x}_0$ ensures that the vectors of state and costate variables converge asymptotically to a stationary point, and at the same time the costate variables satisfy appropriate transversality conditions.

To recover $\overline{\mathbf{g}}_{\eta}$, we differentiate (24) with respect to η to obtain the system of n_x linear equations that leads to the following result.

Theorem 1. *The risk correction term $\overline{\mathbf{g}}_{\eta}$ in (27) is determined by a linear system of n_x inhomogeneous equations with solution*

$$\mathbf{g}_{\eta}(\overline{\mathbf{x}}; 0) = -\left(\overline{\mathbf{a}}_{\mathbf{y}} + \overline{\mathbf{g}}_{\mathbf{x}}\overline{\mathbf{b}}_{\mathbf{y}}\right)^{-1}\left(\overline{\mathbf{\Omega}}_{\eta}^{\mathbf{a}} + \overline{\mathbf{g}}_{\mathbf{x}}\left(\overline{\mathbf{\Omega}}_{\eta}^{\mathbf{b}} + \overline{\mathbf{b}}_{\eta}\right) + \overline{\mathbf{g}}_{\mathbf{xx}}\overline{\mathbf{c}}\right), \tag{32}$$

where $\overline{\mathbf{\Omega}}_{\eta}^{\mathbf{a}} := \left((\operatorname{vec}\mathbf{I}_{n_x})^{\top}(\mathbf{I}_{n_x} \otimes \overline{\mathbf{g}}_{\mathbf{x}}) \otimes \mathbf{I}_{n_x}\right)\overline{\mathbf{a}}_{(\operatorname{vec}\eta\mathbf{y}_{\mathbf{x}})^{\top}}$ is the DSS value of the $n_x \times 1$ matrix $\mathbf{\Omega}_{\eta}^{\mathbf{a}} = \mathcal{D}_{\eta}\{\mathbf{a}(\cdot,\cdot,\eta\mathbf{y}_{\mathbf{x}})\}$, and $\overline{\mathbf{\Omega}}_{\eta}^{\mathbf{b}} := \left((\operatorname{vec}\mathbf{I}_{n_x})^{\top}(\mathbf{I}_{n_x} \otimes \overline{\mathbf{g}}_{\mathbf{x}}) \otimes \mathbf{I}_{n_x}\right)\overline{\mathbf{b}}_{(\operatorname{vec}\eta\mathbf{y}_{\mathbf{x}})^{\top}}$ is the DSS value of the $n_x \times 1$ matrix $\mathbf{\Omega}_{\eta}^{\mathbf{b}} = \mathcal{D}_{\eta}\{\mathbf{b}(\cdot,\cdot,\eta\mathbf{y}_{\mathbf{x}})\}$.

Proof. See Appendix B. ■

Theorem 1 shows that the First-Order approximation to the policy functions for continuous-time models belonging to the class given in (20) includes a non-zero constant correction term $\mathbf{g}_{\eta}(\overline{\mathbf{x}}; 0) \neq \mathbf{0}$, as long as the model is not deterministic, i.e., $\mathbf{c}(\overline{\mathbf{x}}, \overline{\mathbf{y}}) \neq 0$. Thus, the expected values of the endogenous variables are *not* equal to their DSS values. Notice that this contrasts with the results in Schmitt-Grohe and Uribe (2004, Theorem 1), where $\mathbf{g}_{\eta}(\overline{\mathbf{x}}; 0) = \mathbf{0}$ for first-order approximations around the DSS in discrete time. The constant $\mathbf{g}_{\eta}$ captures the effects of risk and hence $\overline{\mathbf{g}}_{\eta} \neq \mathbf{0}$ makes the approximation risk-sensitive and breaks CE, while for $\overline{\mathbf{g}}_{\eta} = \mathbf{0}$ CE prevails (see Judd, 1996). Below, we will refer to the approximation in (27) with $\mathbf{g}_{\eta}(\overline{\mathbf{x}};0)$ set equal to zero as a *First-Order Certainty Equivalent* (CE) approximation.

Notice that the computation of the risk-correction term demands information on the slope, $\mathbf{g}_{\mathbf{x}}(\overline{\mathbf{x}};0)$, and the curvature, $\mathbf{g}_{\mathbf{xx}}(\overline{\mathbf{x}};0)$, of the optimal policy functions at the DSS (cf. Judd and Guu, 1993 and Judd, 1998). Despite the need for this additional information, the approximated policy function (27) remains, however, linear in the state variables. Whilst the information on the slope is already available from (31), the $n_x \times n_x^2$ coefficients in $\overline{\mathbf{g}}_{\mathbf{xx}} := \mathbf{g}_{\mathbf{xx}}(\overline{\mathbf{x}};0)$ are obtained by differentiating (24) with respect to $\mathbf{x}$ twice which yields the following results.

Proposition 2. *The matrix of coefficients $\overline{\mathbf{g}}_{\mathbf{xx}}$ in (32) satisfies the $n_x \times n_x^2$ Sylvester equation*

$$\mathbf{R}\,\overline{\mathbf{g}}_{\mathbf{xx}} + \overline{\mathbf{g}}_{\mathbf{xx}}\,\mathbf{Q} + \mathbf{S} = \mathbf{0}, \tag{33}$$

where $\mathbf{R}$, $\mathbf{Q}$, and $\mathbf{S}$ are, respectively, $n_x \times n_x$, $n_x^2 \times n_x^2$, and $n_x \times n_x^2$ matrices given by

$$\begin{aligned}
\mathbf{R} &= \overline{\mathbf{a}}_{\mathbf{y}} + \overline{\mathbf{g}}_{\mathbf{x}}\overline{\mathbf{b}}_{\mathbf{y}}, \\
\mathbf{Q} &= \mathbf{I}_{n_x} \otimes \left(\overline{\mathbf{b}}_{\mathbf{x}} + \overline{\mathbf{b}}_{\mathbf{y}}\overline{\mathbf{g}}_{\mathbf{x}}\right) + \left(\mathbf{I}_{n_x} \otimes \left(\overline{\mathbf{b}}_{\mathbf{x}} + \overline{\mathbf{b}}_{\mathbf{y}}\overline{\mathbf{g}}_{\mathbf{x}}\right)\right)\mathbf{K}_{n_x,n_x}, \\
\mathbf{S} &= \overline{\mathbf{a}}_{\mathbf{xx}} + \overline{\mathbf{a}}_{\mathbf{xy}}\left(\overline{\mathbf{g}}_{\mathbf{x}} \otimes \mathbf{I}_{n_x}\right) + \left(\overline{\mathbf{a}}_{\mathbf{yx}} + \overline{\mathbf{a}}_{\mathbf{yy}}\left(\overline{\mathbf{g}}_{\mathbf{x}} \otimes \mathbf{I}_{n_y}\right)\right)\left(\mathbf{I}_{n_x} \otimes \overline{\mathbf{g}}_{\mathbf{x}}\right) \\
&\quad + \overline{\mathbf{g}}_{\mathbf{x}}\left(\overline{\mathbf{b}}_{\mathbf{xx}} + \overline{\mathbf{b}}_{\mathbf{xy}}\left(\overline{\mathbf{g}}_{\mathbf{x}} \otimes \mathbf{I}_{n_x}\right) + \left(\overline{\mathbf{b}}_{\mathbf{yx}} + \overline{\mathbf{b}}_{\mathbf{yy}}\left(\overline{\mathbf{g}}_{\mathbf{x}} \otimes \mathbf{I}_{n_y}\right)\right)\left(\mathbf{I}_{n_x} \otimes \overline{\mathbf{g}}_{\mathbf{x}}\right)\right),
\end{aligned}$$

with $\mathbf{K}_{n_x,n_x}$ denoting an $n_x^2 \times n_x^2$ commutation matrix (see Magnus and Neudecker, 2019).

Proof. See Appendix B. ∎

A unique solution $\overline{\mathbf{g}}_{\mathbf{xx}}$ to (33) is given by

$$\operatorname{vec}\left(\mathbf{g}_{\mathbf{xx}}\left(\overline{\mathbf{x}};0\right)\right) = -\left[\mathbf{I}_{n_x^2} \otimes \mathbf{R} + \mathbf{Q}^{\top} \otimes \mathbf{I}_{n_x}\right]^{-1} \operatorname{vec}\left(\mathbf{S}\right), \tag{34}$$

as long as the inverse exists, and the matrices $\mathbf{R}$ and $-\mathbf{Q}$ do not share any eigenvalues (see e.g., Anderson et al., 1996).

Having built the approximation in (27), it is straightforward to compute an approximation to the optimal controls using the first-order conditions from the underlying stochastic optimal control problem. In some applications, these optimality conditions provide an explicit solution for the controls in which case $\mathbf{u}(\mathbf{x};\eta) = \mathcal{U}\left(\mathbf{x}, \mathbf{g}(\mathbf{x};\eta), \eta\mathbf{g}_{\mathbf{x}}(\mathbf{x};\eta)\right)$. If, on the contrary, the optimal controls are only defined implicitly we use the implicit function theorem and define a first-order approximation to the optimal control $\mathbf{u}(\mathbf{x};\eta)$ around the DSS as

$$\mathbf{u}(\mathbf{x};\eta) = \overline{\mathbf{u}} + \mathbf{u}_{\mathbf{x}}\left(\overline{\mathbf{x}};0\right)\left(\mathbf{x} - \overline{\mathbf{x}}\right) + \mathbf{u}_{\eta}\left(\overline{\mathbf{x}};0\right)\eta, \tag{35}$$

where $\overline{\mathbf{u}} := \mathcal{U}\left(\overline{\mathbf{x}}, \mathbf{g}\left(\overline{\mathbf{x}};0\right), \mathbf{0}\right)$, $\mathbf{u}_{\mathbf{x}}\left(\overline{\mathbf{x}};0\right) := \mathcal{U}_{\mathbf{x}}\left(\overline{\mathbf{x}}, \mathbf{g}\left(\overline{\mathbf{x}};0\right), \mathbf{0}\right) + \mathcal{U}_{\mathbf{y}}\left(\overline{\mathbf{x}}, \mathbf{g}\left(\overline{\mathbf{x}};0\right), \mathbf{0}\right)\mathbf{g}_{\mathbf{x}}\left(\overline{\mathbf{x}};0\right)$, and $\mathbf{u}_\eta(\overline{\mathbf{x}};0) := \mathcal{U}_\mathbf{y}(\overline{\mathbf{x}},\mathbf{g}(\overline{\mathbf{x}};0),\mathbf{0})\,\mathbf{g}_\eta(\overline{\mathbf{x}};0) + \mathcal{U}_{\eta\mathbf{y}_\mathbf{x}}(\overline{\mathbf{x}},\mathbf{g}(\overline{\mathbf{x}};0),\mathbf{0})\,\mathbf{g}_\mathbf{x}(\overline{\mathbf{x}};0)$.

Second- and higher-order Perturbations. By further differentiating (24) with respect to the state variables and the perturbation parameter, it is possible to construct higher-order approximations to the unknown policy functions using (26). For example, we define a *Second-Order* approximation to the unknown policy function around the DSS as

$$\begin{aligned}\mathbf{g}(\mathbf{x};\eta) = \mathbf{g}(\overline{\mathbf{x}};0) + \mathbf{g}_\mathbf{x}(\overline{\mathbf{x}};0)(\mathbf{x}-\overline{\mathbf{x}}) + \mathbf{g}_\eta(\overline{\mathbf{x}};0)\eta \\ + \tfrac{1}{2}\mathbf{g}_{\mathbf{xx}}(\overline{\mathbf{x}};0)(\mathbf{x}-\overline{\mathbf{x}})\otimes(\mathbf{x}-\overline{\mathbf{x}}) + \mathbf{g}_{\mathbf{x}\eta}(\overline{\mathbf{x}};0)(\mathbf{x}-\overline{\mathbf{x}})\eta + \tfrac{1}{2}\mathbf{g}_{\eta\eta}(\overline{\mathbf{x}};0)\eta^2. \quad (36)\end{aligned}$$

Similar to the First-Order approximation, the computation of the risk-correction terms $\mathbf{g}_{\mathbf{x}\eta}(\overline{\mathbf{x}};0)$ and $\mathbf{g}_{\eta\eta}(\overline{\mathbf{x}};0)$ in (36) requires information on the derivatives of $\mathbf{g}(\mathbf{x},\eta)$ with respect to $\mathbf{x}$ at the DSS beyond that already provided by $\mathbf{g}_\mathbf{x}(\overline{\mathbf{x}};0)$ and $\mathbf{g}_{\mathbf{xx}}(\overline{\mathbf{x}};0)$. As shown in Gaspar and Judd (1997), as a general rule, knowledge of the first $(k+2)$ derivatives of $\mathbf{g}(\mathbf{x};\eta)$ with respect to $\mathbf{x}$ only provides information on the first k derivatives of $\mathbf{g}_\eta(\mathbf{x};\eta)$ with respect to $\mathbf{x}$. From a computational perspective, this implies that the cost of building a k-th order perturbation for continuous-time stochastic models, in terms of the number of matrix operations required, is of order $\mathcal{O}(n_x^{k+2})$.[7] According to this rule, to obtain $\mathbf{g}_{\mathbf{x}\eta}(\overline{\mathbf{x}};0)$ and $\mathbf{g}_{\eta\eta}(\overline{\mathbf{x}};0)$ we need to compute $\mathbf{g}_{\mathbf{xxx}}(\overline{\mathbf{x}};0)$, $\mathbf{g}_{\mathbf{xxxx}}(\overline{\mathbf{x}};0)$ and $\mathbf{g}_{\mathbf{xx}\eta}(\overline{\mathbf{x}};0)$. An attractive feature of the perturbation approach is that all these unknown coefficients are obtained as the solution to sequential linear system of equations with constants that depend on lower-order derivatives of the policy function at the DSS.

Remark 3. The choice of not including the value function $V(\mathbf{x};\eta)$ in the vector $\mathbf{y}$ is because of our interest in computing a risk-corrected first-order, and thus linear, approximation to the control variables (35). Since the latter are usually determined by the first-order derivatives of the value function (i.e. the costates), any first-order approximation to the value function would imply optimal control variables that are constant. As a by-product of the First-Order approximation to the costate variables (27) it is straightforward to define first- and a second-order approximations to the value function V as

$$\begin{aligned}V(\mathbf{x};\eta) = \overline{V} + \mathbf{g}(\overline{\mathbf{x}};0)(\mathbf{x}-\overline{\mathbf{x}}) + \overline{V}_\eta\eta \\ + \tfrac{1}{2}\mathbf{g}_\mathbf{x}(\overline{\mathbf{x}};0)(\mathbf{x}-\overline{\mathbf{x}})\otimes(\mathbf{x}-\overline{\mathbf{x}}) + \mathbf{g}_\eta(\overline{\mathbf{x}};0)(\mathbf{x}-\overline{\mathbf{x}})\eta + \tfrac{1}{2}\overline{V}_{\eta\eta}\eta^2, \quad (37)\end{aligned}$$

where $\overline{V}$, $\overline{V}_\eta$, and $\overline{V}_{\eta\eta}$ are computed from the DSS value of the maximized HJB equation and the application of the envelope theorem. ■

[7] This contrasts with the computational cost of a k-th order approximation to discrete-time models which is of order $\mathcal{O}(n_x^k)$.

3.2 An illustration: The stochastic growth model

To illustrate how the procedure works, consider the stochastic growth model which results from setting $X_0 = b = 0$ and letting $\xi \to \infty$ in the prototype model of Section 2. A detailed derivation of all the results can be found in the accompanying Online Appendix.

Competitive equilibrium. For any given admissible values of the state variables (K_t, A_t), optimal consumption in this economy is entirely determined by the costate variable for capital, $C_t = V_K^{-1/\gamma}$, which in equilibrium satisfies the second-order quasilinear PDE

$$(\exp(A_t)K_t^\alpha - V_K^{-1/\gamma} - \delta K_t)V_{KK} + (\alpha \exp(A_t)K_t^{\alpha-1} - \delta)V_K \\ - \rho_A A_t V_{AK} + \tfrac{1}{2}\eta\sigma_A^2 V_{AAK} - \rho V_K = 0, \quad (38)$$

with solution

$$V_K = V_K(K_t, A_t; \eta). \quad (39)$$

Let $\mathbf{y} = V_K$ and $\mathbf{x} = [K_t, A_t]^\top$. Then, the equilibrium condition (38) belongs to the class of models in (20) with

$$\begin{aligned}
\mathbf{a}(\mathbf{x}, \mathbf{y}) &= \left(\alpha \exp(A_t)K_t^{\alpha-1} - \delta - \rho\right) V_K, \\
\mathbf{b}(\mathbf{x}, \mathbf{y}) &= \begin{bmatrix} \exp(A_t)K_t^\alpha - V_K^{-1/\gamma} - \delta K_t \\ -\rho_A A_t \end{bmatrix}, \\
\mathbf{c}(\mathbf{x}, \mathbf{y}) &= \left[0,\ 0,\ 0,\ \tfrac{1}{2}\sigma_A^2\right]^\top.
\end{aligned} \quad (40)$$

Substitution of (39) into (40) yields the functional equation $\mathbf{F}(\mathbf{x}; \eta) = \mathbf{0}$.

Deterministic steady state. For $\eta = 0$, the model's DSS is given by the fixed point $(\overline{\mathbf{x}}, \overline{\mathbf{y}}) = \left(\overline{K}, \overline{A}, \overline{V_K}\right)$ that solves the system of equations

$$\begin{bmatrix} \mathbf{a}(\overline{\mathbf{x}}, \overline{\mathbf{y}}) \\ \mathbf{b}(\overline{\mathbf{x}}, \overline{\mathbf{y}}) \end{bmatrix} = \begin{bmatrix} \left(\alpha \exp(\overline{A})\overline{K}^{\alpha-1} - \delta - \rho\right) \overline{V_K} \\ \exp(\overline{A})\overline{K}^\alpha - \overline{V_K}^{-1/\gamma} - \delta \overline{K} \\ -\rho_A \overline{A} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}. \quad (41)$$

In particular, the DSS is given by $\overline{A} = 0$, $\overline{K} = [\alpha/(\rho+\delta)]^{\frac{1}{1-\alpha}}$, and $\overline{V_K} = \left(\overline{K}^\alpha - \delta\overline{K}\right)^{-\gamma}$.

Approximate solution. The First-Order approximation to the costate variable for capital around the DSS is defined as

$$V_K(K, A; \eta) = \overline{V_K} + \overline{V_{KK}}\left(K - \overline{K}\right) + \overline{V_{KA}}\left(A - \overline{A}\right) + \overline{V_{K\eta}}\eta, \quad (42)$$

where $\overline{V_K}$ is obtained from (41), while the remaining constants, $\overline{V_{KK}} := \partial V_K\left(\overline{K}, \overline{A}; 0\right)/\partial K$ and $\overline{V_{KA}} := \partial V_K\left(\overline{K}, \overline{A}; 0\right)/\partial A$ are the solution to the quadratic system of equations formed by

$$\mathbf{F}_{\mathbf{x}}\left(\overline{\mathbf{x}},0\right)=\left[\begin{array}{c} F_K(\overline{K},\overline{A};0) \\ F_A(\overline{K},\overline{A};0)\end{array}\right]=\left[\begin{array}{c}0\\0\end{array}\right].$$

In particular, these constants are given by

$$\overline{V_{KK}} = \frac{-\rho \pm \sqrt{\rho^2 - 4\frac{1}{\gamma}\alpha(\alpha-1)\overline{K}^{\alpha-2}\overline{V_K}^{-1/\gamma}}}{2\frac{1}{\gamma}\overline{V_K}^{-1/\gamma-1}}, \tag{43}$$

$$\overline{V_{KA}} = -\left(\frac{1}{\gamma}\overline{V_K}^{-1/\gamma-1}\overline{V_{KK}}-\rho_A\right)^{-1}\left(\overline{K}^{\alpha}\overline{V_{KK}}+(\delta+\rho)\overline{V_K}\right). \tag{44}$$

To ensure that the problem's value function is strictly concave along the capital stock lattice, we choose the root in (43) that ensures $\overline{V_{KK}}<0$ (see Parra-Alvarez, 2018).

The remaining constant, $\overline{V_{K\eta}} := \partial V_K\left(\overline{K},\overline{A};0\right)/\partial\eta$, corresponds to the solution to the inhomogeneous linear equation

$$\mathbf{F}_{\eta}\left(\overline{\mathbf{x}},0\right)=F_{\eta}(\overline{K},\overline{A};0)=0,$$

which is given by

$$\overline{V_{K\eta}} = -\tfrac{1}{2}\sigma_A^2\frac{\overline{V_{AAK}}}{\frac{1}{\gamma}\overline{V_K}^{-1/\gamma-1}\overline{V_{KK}}}. \tag{45}$$

Having obtained (42), it is now possible to compute a linear approximation to the optimal consumption function by linearizing the first-order condition, $C=V_K^{-1/\gamma}$, around the DSS

$$C(K,A,\eta)=\overline{C}+\overline{C_K}\left(K-\overline{K}\right)+\overline{C_A}\left(A-\overline{A}\right)+\overline{C_\eta}\eta, \tag{46}$$

where $\overline{C}=\overline{V_K}^{-1/\gamma}$, with $\overline{C_K} := \partial C\left(\overline{K},\overline{A};0\right)/\partial K$, $\overline{C_A} := \partial C\left(\overline{K},\overline{A};0\right)/\partial A$, and $\overline{C_\eta} := \partial C\left(\overline{K},\overline{A};0\right)/\partial\eta$ given by

$$\overline{C_K} = -\frac{1}{\gamma}\overline{V_K}^{-1/\gamma-1}\overline{V_{KK}} = \frac{\rho}{2}\mp\sqrt{\left(\frac{\rho}{2}\right)^2-\frac{1}{\gamma}\alpha(\alpha-1)\overline{K}^{\alpha-2}\overline{C}}, \tag{47}$$

$$\overline{C_A} = -\frac{1}{\gamma}\overline{V_K}^{-1/\gamma-1}\overline{V_{KA}} = \frac{1}{(\overline{C_K}+\rho_A)}\left(\overline{K}^{\alpha}\overline{C_K}-\frac{(\delta+\rho)\overline{C}}{\gamma}\right), \tag{48}$$

$$\overline{C_\eta} = -\frac{1}{\gamma}\overline{V_K}^{-1/\gamma-1}\overline{V_{K\eta}} = -(\overline{C_K})^{-1}\tfrac{1}{2}\left[(1+\gamma)\overline{C}\left(\frac{\overline{C_A}}{\overline{C}}\right)^2-\overline{C_{AA}}\right]\sigma_A^2. \tag{49}$$

Expression (49) shows that $\overline{C_\eta}\neq 0$, suggesting that a first-order perturbation approximation to the optimal consumption function includes an adjustment for risk. Given the concavity of the consumption function, the constant $\overline{C_\eta}$ is negative. This implies that risk averse agents will consume less in the presence of risk due to precautionary savings. The risk-correction term requires information on both the slope and the curvature of the optimal consumption function at the DSS. While $\overline{C_K}$ and $\overline{C_A}$ are given by (47) and (48), $\overline{C_{AA}}$ is still unknown but can be obtained by first computing $\overline{V_{AAK}}$ in (45) as the solution to the linear system of equations formed by

$$\mathbf{F_{xx}}\left(\overline{\mathbf{x}},0\right)=\left[F_{KK}(\overline{K},\overline{A};0),\ F_{KA}(\overline{K},\overline{A};0),\ F_{AA}(\overline{K},\overline{A};0)\right]^{\top}=\mathbf{0}.$$

Notice that certainty equivalence will still hold, i.e., $\overline{C_\eta}=0$, under the following assumptions on the stochastic growth model: (i) zero risk, $\sigma_A=0$; (ii) quadratic utility, $(1+\gamma)=0$ and $C_{AA}=0$ (see Judd, 1996).

3.3 An intuition: Why does certainty equivalence break?

The solution to stochastic economic models is said to be certainty equivalent if the resulting policy functions are invariant to higher-order moments of the model's underlying exogenous shocks. In other words, the solution of an economic model under uncertainty is identical to the solution of the same model under certainty.

For discrete-time stochastic models, certainty equivalence holds for any first-order (linear) approximation around the DSS. In general, the optimality conditions that characterize equilibria in these models can be summarized by a system of stochastic difference equations, where expectations regarding the future value of the control variables need to be formed. Given that the policy functions are *a priori* unknown, the computation of such expectations can only be done *ex-post* once the optimal controls have been approximated. Hence, for a first-order perturbation this is equivalent to calculating the expected value of a set of linear functions which, according to the linearity property of the expectation operator, implies that only first-order moments will enter the approximated solution.

However, as exemplified by equation (49), this is not the case for continuous-time stochastic models. To illustrate this point let us consider the equilibrium dynamics of consumption in the stochastic growth model. The Euler equation for consumption in this case is given by (see Online Appendix)

$$\mathrm{d}C_t=\left[\frac{\alpha\exp(A_t)K_t^{\alpha-1}-\delta-\rho}{\gamma}+\tfrac{1}{2}(1+\gamma)\left(\frac{C_A}{C_t}\right)^2\eta\sigma_A^2\right]C_t\mathrm{d}t+C_A\sqrt{\eta\sigma_A^2}\mathrm{d}B_{A,t}. \quad (50)$$

Notice that (50) includes some features that account for the model's underlying risk. In particular, consider the quadratic term $\frac{1}{2}(1+\gamma)C\left(C_A/C\right)^2\eta\sigma_A^2$, which also appears in (49). The first thing to note is that it contains the marginal response of optimal consumption to changes in the exogenous driving force of the model, C_A, which is closely related to risk aversion. To see this recall that in equilibrium the optimal consumption function, C, is related to the marginal utility of consumption, $u'\left(C\right)$, and thus C_A is related to the first-order derivative of the marginal utility, $u''\left(C\right)$. This contrasts to the Euler equation

Risk effects on:	$\partial^n u/(\partial c)^n$	related to:	Cont. Time		Discrete Time		
			1st	2nd	1st	2nd	3rd
—	$n = 2$	risk aversion: $-u''/u'$	✓	✓	✓	✓	✓
level of C	$n = 3$	prudence: $-u'''/u''$	✓	✓		✓	✓
slope of C	$n = 4$	temperance: $u^{(4)} < 0$		✓			✓

Table 1. Effects of risk in perturbation solutions. The table indicates the order of the derivative of the utility function $\partial^n u/(\partial c)^n$ necessary to account for a particular effect of risk on optimal consumption, as well as the order of approximation needed to capture it both in continuous-time and discrete-time stochastic models.

for consumption in discrete time which only includes terms related to $u'(C)$ due to the presence of expected values that cannot be computed *a priori*. Also note that (50) contains the perturbation and the variance parameters which jointly capture the amount of risk in the model, $\eta\sigma_A^2$. Finally, note that the term $(1 + \gamma)$ is the coefficient of relative prudence for the case of CRRA utility functions.

How this relates to certainty equivalence becomes clear when taking a closer look at the precautionary motive, or prudence, that describes the optimal reaction of consumption to risk. Prudence is related to the third derivative of the utility function, $u'''(C)$, and its absence leads to certainty equivalence.[8] Hence, a policy function that only contains $u''(C)$ will account for risk aversion, i.e., how much an agent dislikes risk, but not for prudence and, thus, will be certainty equivalent. If, in addition, the policy function involves the fourth derivative of the utility function, $u^{(4)}(C) < 0$, then it will also account for temperance, i.e., how the marginal propensity to consume responds to risk. Thus, while the effects of risk on the level of consumption are captured by $u'''(C)$, the effects on the slope are captured by $u^{(4)}(C)$ (see Kimball, 1990a and Zeldes, 1989).

In terms of the approximation method, note that a first-order (linear) perturbation to the unknown consumption function requires computing the first derivative of the Euler equation. Since its discrete-time version only contains $u'(C)$, a first-order approximation will just include terms up to the second derivative of the utility function and hence, it will account for risk aversion but not for prudence. Therefore, a first-order approximation in discrete time will be certainty equivalent. In contrast, the continuous-time Euler equation (50) already includes terms related to $u'(C)$ and $u''(C)$, so its first-order approximation will account for both risk aversion *and* prudence. The resulting policy functions in continuous time will not only depend on the mean of the exogenous shock but also on its variance – *breaking certainty equivalence*. To break certainty equivalence in discrete time, a second-order approximation is needed, which in continuous time already leads to correction terms in the slopes. Nevertheless, notice that the ability of a first-order approximation to break the certainty-equivalence property in continuous-time models comes with

[8] Absolute prudence is defined as $-u'''(C)/u''(C)$, while relative prudence is defined as $-u'''(C)C/u''(C)$ (see Kimball, 1990b).

larger computational costs than those implied by a first-order approximation to discrete-time models (see footnote 7). Table 1 summarizes the discussion above by indicating which order of approximation is required in order to account for a given risk effect.

3.4 Risky steady state

Similar to the concept of the DSS, we may define the risky or stochastic steady state (RSS) as the fixed point to which the economic system converges in the presence of uncertainty, but in the absence of shocks. Knowledge of the RSS facilitates the understanding of the transitional dynamics implied by the perturbation-based approximations. Furthermore, the RSS is relevant to the extent that it incorporates information regarding the future risk prospects of risk-averse economic agents (see Coeurdacier et al., 2011).

Unfortunately, the computation of the RSS is not straightforward. Following its definition, we require information about how risk, as measured by the variance of economic shocks, affects the policy functions, $\mathbf{g}(\mathbf{x};\eta)$, which, *ex-ante*, are also unknown. However, it is still possible to approximate its value by using the perturbation-based approximation of the policy functions around the DSS.

In particular, we define the RSS of the state variables, $\widehat{\mathbf{x}}$, by the solution to the system of equations formed by

$$\mathbf{b}\left(\widehat{\mathbf{x}}, \mathbf{g}\left(\widehat{\mathbf{x}};1\right), \mathbf{g}_{\mathbf{x}}\left(\widehat{\mathbf{x}};1\right);1\right) = \mathbf{0}, \tag{51}$$

where $\mathbf{b}(\cdot)$ is the drift in (18) that results from (i) replacing $\mathbf{y} = \mathbf{g}(\mathbf{x};\eta)$ and $\mathbf{y}_{\mathbf{x}} = \mathbf{g}_{\mathbf{x}}(\mathbf{x};\eta)$, with the corresponding k-th order perturbation-based approximation (26) evaluated at $\widehat{\mathbf{x}}$; (ii) setting any future realization of economic shocks to zero, i.e. $\mathrm{d}\mathbf{w} = \mathbf{0}$; and (iii) imposing the stationarity condition $\mathrm{d}\widehat{\mathbf{x}}/\mathrm{d}t = \mathbf{0}$. Thus, we refer to $\widehat{\mathbf{x}}$ as the k-th order approximation to the RSS value of the state vector. Once $\widehat{\mathbf{x}}$ is computed, we obtain the corresponding k-th order approximation to the RSS of the costate and control variables as $\widehat{\mathbf{y}} = \mathbf{g}(\widehat{\mathbf{x}};\eta = 1)$, and $\widehat{\mathbf{u}} = \mathcal{U}(\widehat{\mathbf{x}}, \widehat{\mathbf{y}}, \widehat{\mathbf{y}}_{\mathbf{x}})$, respectively.

An attractive feature of the perturbation method of Section 3.1 is that since it is possible to account for risk using a first-order approximation, we can use (51) to build an approximation to the RSS already within a linear framework, i.e., for $k = 1$.[9] In this case, the first-order approximation to the RSS is given by the tuple $(\widehat{\mathbf{x}}, \widehat{\mathbf{y}}, \widehat{\mathbf{u}}; \eta = 1)$ that solves

$$\begin{aligned}
\mathbf{0} &= \mathbf{b}\left(\widehat{\mathbf{x}}, \mathbf{g}(\overline{\mathbf{x}};0) + \mathbf{g}_{\mathbf{x}}(\bar{\mathbf{x}};0)(\widehat{\mathbf{x}} - \overline{\mathbf{x}}) + \mathbf{g}_{\eta}(\overline{\mathbf{x}};0), \mathbf{g}_{\mathbf{x}}(\overline{\mathbf{x}};0);1\right), &(52)\\
\widehat{\mathbf{y}} &= \mathbf{g}(\overline{\mathbf{x}};0) + \mathbf{g}_{\mathbf{x}}(\overline{\mathbf{x}};0)(\widehat{\mathbf{x}} - \overline{\mathbf{x}}) + \mathbf{g}_{\eta}(\overline{\mathbf{x}};0), &(53)\\
\widehat{\mathbf{u}} &= \overline{\mathbf{u}} + \mathbf{u}_{\mathbf{x}}\left(\overline{\mathbf{x}};0\right)\left(\widehat{\mathbf{x}} - \overline{\mathbf{x}}\right) + \mathbf{u}_{\eta}\left(\overline{\mathbf{x}};0\right). &(54)
\end{aligned}$$

A similar procedure can be used to build a k-th order approximation of the RSS, for $k > 1$.

[9] de Groot (2013) uses a similar approach for discrete-time models for the case of perturbations of order $k = 2$ and higher (see Online Appendix).

As an example, let us consider the stochastic growth model in Section 3.2. Using the approximated policy function (42), we can compute the RSS value for the state variables, $(\widehat{K}, \widehat{A})$, as the solution to

$$\begin{bmatrix} \exp(\hat{A})\widehat{K}^{\alpha} - \left(\overline{V_K} + \overline{V_{KK}}\left(\widehat{K} - \overline{K}\right) + \overline{V_{KA}}\left(\widehat{A} - \overline{A}\right) + \overline{V_{K\eta}}\right)^{-1/\gamma} - \delta\widehat{K} \\ -\rho_A\widehat{A} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

It follows that $\widehat{A} = \overline{A} = 0$, while the first-order approximation to the RSS of the capital stock, $\widehat{K}$, is given as the solution to the nonlinear equation

$$\widehat{K}^{\alpha} - \left(\overline{V_K} + \overline{V_{KK}}\left(\widehat{K} - \overline{K}\right) + \overline{V_{K\eta}}\right)^{-1/\gamma} - \delta\widehat{K} = 0.$$

Next, substitute $(\widehat{K}, \widehat{A})$ in (42) and (46) to obtain an approximation to the RSS values for $\widehat{V_K}$ and $\widehat{C}$, respectively.[10]

4 Economic implications

This section investigates the economic implications of the approximated solutions in (26) by measuring the pricing errors when using the First-Order CE approximation, the First-Order approximation, and the Second-Order approximation defined in Section 3.1. The pricing errors are compared relative to the solution obtained from a global nonlinear collocation method based on a Chebyshev polynomial approximation of the unknown value function. Though this approach delivers highly accurate solutions, it is costly in terms of computational efficiency (see Parra-Alvarez, 2018; Posch, 2020). By comparing the economic pricing errors, we can study how the risk adjustment matters quantitatively.

4.1 Stochastic discount factor

We define the stochastic discount factor (SDF) as the process m_s/m_t , such that, for any security with price P_t, and a single payoff χ_s at some future date $s \geq t$, we obtain

$$m_t P_t = \mathbb{E}_t\left[m_s \chi_s\right] \quad \Rightarrow \quad 1 = \mathbb{E}_t\left[(m_s/m_t)R_s\right], \tag{55}$$

where $R_s := \chi_s/P_t$ is the security's gross return, and m_t is the present (discounted) value of a unit of consumption in period t. Hence, condition (55) can be used to discount expected payoffs on *any* asset with a single payoff to find its equilibrium price. From the expected discounted life-time utility (1), the instantaneous utility function (2), and the first-order condition (11), we obtain the SDF for $s > t$ following Detemple and Zapatero

[10] As shown in the Online Appendix, an approximate closed-form solution to the RSS can be obtained by first linearizing $\mathbf{b}(\mathbf{x}, \mathbf{y}; \eta)$ around the DSS.

(1991) as (see Appendix A)

$$m_s/m_t = e^{-\rho(s-t)} \frac{(C_s - X_s)^{-\gamma} + bV_{X,s}}{(C_t - X_t)^{-\gamma} + bV_{X,t}} \quad \text{for} \quad s > t, \tag{56}$$

where $m_t = e^{-\rho t}(C_t - X_t)^{-\gamma} + bV_{X,t}$. Hence, the accuracy of the approximation to the models' solution matters for the pricing kernel in (56), where C_t and $V_{X,t}$ are replaced by the approximate solutions, respectively.

4.2 Pricing errors and their decomposition

In what follows we define pricing errors as (see Lettau and Ludvigson, 2009)

$$\varepsilon := \mathbb{E}_t[(m_s/m_t)R_s] - 1 \tag{57}$$

based on the gross return on any tradable asset with instantaneous return, R_s at $s > t$, and where $\varepsilon = 0$ for the model's true (but unknown) policy functions. This (unit free) quantity compares to the measure of numerical accuracy based on first-order principles, often referred to as 'Euler equation errrors' (cf. Aruoba et al., 2006). In fact, it compares to the discrete-time intertemporal condition and is interpreted as the relative error incurred by the use of the approximated policy functions, measured in terms of hypothetical pricing errors of a given asset.[11]

In the following, we focus on pricing *zero-coupon bonds* with sure payoff $\chi_{t+N} = 1$ at period $t + N$. From (55), we obtain the price of this zero-coupon bond as $P_t = \mathbb{E}_t\left[(m_{t+N}/m_t)\right]$ such that the gross return of this asset is $R_t = 1/P_t$, conditional on the information set at time t. Unfortunately, we do not readily observe the instantaneous risk-free rate for $N \to 0$ (with corresponding $r_t := \lim_{N\to 0} - \ln P_t/N$). Any equilibrium return from a zero-coupon bond carries a term premium for a given time-to-maturity N.

To compute the price of a zero-coupon bond for a given time-to-maturity N, we use the PDE approach (Cochrane, 2005, chap. 19.4). In the absence of arbitrage opportunities, the fundamental price of a zero-coupon bond P_t with fixed maturity N satisfies

$$\underbrace{\mathbb{E}_t\left(\frac{\mathrm{d}P_t}{P_t}\right) - \frac{1}{P_t}\frac{\partial P_t}{\partial N}\mathrm{d}t}_{\text{holding period return}} = \underbrace{-\mathbb{E}_t\left(\frac{\mathrm{d}m_t}{m_t}\right)}_{\text{risk-free return}} + \underbrace{\left(-\frac{1}{\mathrm{d}t}\mathbb{E}_t\left(\left(\frac{\mathrm{d}P_t}{P_t}\right)\left(\frac{\mathrm{d}m_t}{m_t}\right)\right)\right)}_{\text{term premium}}, \tag{58}$$

with $-\mathbb{E}_t\left(\mathrm{d}m_t/m_t\right) = r_t\mathrm{d}t$, and where the SDF evolves according to

$$\frac{\mathrm{d}m_t}{m_t} = \mu_m\mathrm{d}t + \sigma_m\mathrm{d}B_{A,t}, \tag{59}$$

[11]Following Judd and Guu (1993), we provide in the Online Appendix a measure of the numerical accuracy of the approximated policy functions in terms of residuals to the HJB equation (12).

defining the drift, μ_m, and diffusion, σ_m, terms in (59) as functions of the policy functions (see Appendix A). Hence, they both depend on the accuracy of the approximate solution to the economic model.[12]

Given the decomposition in (58), we are now prepared to shed light on the different sources of pricing errors. The arbitrage-free price of the zero-coupon bond P_t is defined in terms of the (true) policy functions from the data generating process (DGP), for which $\varepsilon = 0$. For an approximation to the model's policy functions we decompose any potential pricing error into three categories. First, the holding period return is inaccurate. Second, the risk-free rate is poorly approximated. And third, the term premium is poorly captured.

Since in equilibrium all the time dependence of the zero-coupon bond price with given time-to-maturity N comes through the law of motions of the state variables that drive the economy, i.e., $P_t = P(K_t, X_t, A_t)$, an application of Itô's lemma shows that the dynamics of the bond's price is given by (cf. Posch, 2020)

$$\mathrm{d}P_t = \frac{\partial P_t}{\partial K_t}\mathrm{d}K_t + \frac{\partial P_t}{\partial X_t}\mathrm{d}X_t + \frac{\partial P_t}{\partial A_t}\mathrm{d}A_t + \tfrac{1}{2}\frac{\partial^2 P_t}{\partial A_t^2}\left(\mathrm{d}A_t\right)^2. \tag{60}$$

Inserting (60), together with (3), (5), and (7) into (58) yields the fundamental pricing equation as

$$\begin{aligned}\frac{\partial P_t}{\partial N} \;=\;& \mu_m P_t + \frac{\partial P_t}{\partial K_t}\Big(\Phi((\exp(A_t)K_t^\alpha - C(K_t, X_t, A_t))/K_t) - \delta\Big)K_t \\ &+\frac{\partial P_t}{\partial X_t}\Big(bC(K_t, X_t, A_t) - aX_t\Big) - \rho_A A_t\frac{\partial P_t}{\partial A_t} + \tfrac{1}{2}\frac{\partial^2 P_t}{\partial A_t^2}\sigma_A^2 + \frac{\partial P_t}{\partial A_t}\sigma_A\sigma_m.\end{aligned} \tag{61}$$

Again, we can decompose the pricing error (57) into a *direct* effect of the approximated solution $C(K_t, X_t, A_t)$ on the holding period return, and two *indirect* effects through the risk-free rate channel, μ_m, and through the covariance of the price with the SDF, σ_m.

The functional form of the solution to the PDE (61) is unknown. In order to study the effects of approximation errors on the pricing of the zero-coupon bond, we use collocation methods to approximate the price as a function of the time-to-maturity and the state variables with the polynomial $P_t \approx \boldsymbol{\phi}(N, K_t, X_t, A_t)\boldsymbol{\nu}$, in which $\boldsymbol{\nu}$ is a vector of unknown coefficients, and $\boldsymbol{\phi}(\cdot)$ denotes the Chebyshev basis matrix with associated Chebyshev nodes. We extend the PDE in (61) with the boundary condition $\boldsymbol{\phi}(0, K_t, X_t, A_t)\boldsymbol{\nu} = \mathbf{1}_p$, where p denotes the degree of the approximation. This approach provides accurate results, is simple to implement as it requires only the solution of a linear system, and thus allows us to avoid tedious simulations for the different approximated solutions.

[12] Alternatively, one could include the bond prices at different maturities in $\mathcal{H}$ and construct a perturbation approximation around the DSS. Andreasen and Zabczyk (2015) propose an efficient way to implement this approach using a perturbation-on-perturbation framework for discrete-time models.

From (57), we obtain for the zero-coupon bond with time-to-maturity N,

$$\varepsilon^{(N)} := \mathbb{E}_t(m_{t+N}/m_t)/P_t - 1,$$

which is zero for the correct pricing kernel $\mathbb{E}_t(m_{t+N}/m_t) = P_t$ implying the arbitrage-free price P_t, but would be different from zero if replaced by an investor's subjective pricing kernel built from an approximated solution to the DGP. Let P_t^a be the resulting price, i.e.,

$$P_t^a := \mathbb{E}_t\Big[(m_{t+N}/m_t)\Big|C(K_t, X_t, A_t) = \mathbf{u}(\mathbf{x};\eta), \mu_m = \tilde{\mu}_m, \sigma_m = \tilde{\sigma}_m\Big]$$

based on the subjective solution for the consumption function $C(K_t, X_t, A_t) = \mathbf{u}(\mathbf{x};\eta)$, for the drift $\tilde{\mu}_m := \mu_m(\mathbf{x}, \mathbf{u}(\mathbf{x};\eta), \mathbf{g}(\mathbf{x};\eta))$ and for the diffusion $\tilde{\sigma}_m := \sigma_m(\mathbf{x}, \mathbf{u}(\mathbf{x};\eta), \mathbf{g}(\mathbf{x};\eta))$. Below, we study the consequences of misspricing generated by the First-Order (CE), the First-Order, or the Second-Order approximation on both the bond-price dynamics in (60) and/or the SDF dynamics in (59).

Let $\varepsilon_a^{(N)}$ represent a measure of *ex-ante* pricing errors on a zero-coupon bond for given time-to-maturity N, defined as the (absolute) percentage deviation of the bond price using the subjective pricing kernel relative to the arbitrage-free price

$$\varepsilon_a^{(N)} := P_t^a/P_t - 1. \tag{62}$$

To analyze the different sources of pricing mismatch, we define the *ex-post* pricing errors as the (absolute) percentage deviation of the price under the subjective pricing kernel relative to the no-arbitrage price, when the investor can observe the true SDF dynamics either partially (drift only) or completely (drift and diffusion), respectively,

$$\varepsilon_b^{(N)} := \mathbb{E}_t\Big[(m_{t+N}/m_t)\Big|C(K_t, X_t, A_t) = \mathbf{u}(\mathbf{x};\eta), \sigma_m = \tilde{\sigma}_m\Big]/P_t - 1, \tag{63}$$

$$\varepsilon_c^{(N)} := \mathbb{E}_t\Big[(m_{t+N}/m_t)\Big|C(K_t, X_t, A_t) = \mathbf{u}(\mathbf{x};\eta)\Big]/P_t - 1. \tag{64}$$

Hence, the measures in (63) and (64) are focusing on the pricing error obtained by providing further information on the SDF dynamics. While $\varepsilon_b^{(N)}$ measures the error when shutting down the indirect risk-free rate channel in the error decomposition, the second $\varepsilon_c^{(N)}$ measures the error when shutting down both indirect channels, mainly focusing on the direct pricing errors. For example in (64), the investor infers the correct SDF from the data, and solves the corresponding PDE

$$\frac{\partial P_t^c}{\partial N} = \mu_m P_t^c + \frac{1}{\mathrm{d}t}\mathbb{E}_t\Big[\mathrm{d}P_t^c\Big|C(K_t, X_t, A_t) = \mathbf{u}(\mathbf{x};\eta)\Big] + \left(\frac{\partial P_t^c}{\partial A_t}\right) P_t^c \sigma_A \sigma_m,$$

with the approximated price P_t^c (in the same way we define P_t^b). This enables us to study the hypothetical error an investor would face ex-post when trading the asset at the

Parameter	Value	Source / Target
Discounting, ρ	0.0410	Jermann (1998)
Risk aversion, γ	2.0000	Aruoba et al. (2006)
Depreciation rate, δ	0.0963	Jermann (1998)
Capital share in output, α	0.3600	Jermann (1998)
Persistence TFP, ρ_A	0.2052	Aruoba et al. (2006)
Volatility TFP, σ_A	0.0307	U.S. real GDP growth volatility
Adjustment cost, ξ	0.3261	Short-term return on government bonds reported in Jermann (1998)
Habit current cons., b	0.8200	Jermann (1998)
Habit past cons., a	1.0000	Jermann (1998)

Table 2. Parameter values. The parameters of the model are calibrated to an annual frequency and their values should be interpreted accordingly.

subjective (approximated) price instead of true P_t, yet knowing the SDF dynamics.

5 Results

5.1 Calibration

To quantitatively evaluate the extent to which the First-Order approximation can account for the effects of risk, we proceed to calibrate the prototype model of Section 2 to an annual frequency. Therefore, all the parameter values should be interpreted accordingly. Many of the parameter values are chosen to replicate the parameterization to the U.S. economy used in the discrete-time models of Jermann (1998) and Aruoba et al. (2006). A complete summary of the model's calibration is provided in Table 2.

In particular, we set the risk aversion parameter and the share of capital income to $\gamma = 2$ and $\alpha = 0.36$, respectively. The values for the subjective discount rate, the depreciation rate and the habit process are set to $\rho = 0.041$, $\delta = 0.0963$, and $a = 1$ and $b = 0.82$, respectively. These parameter values are consistent with steady-state values for the capital-output ratio, and the consumption and investment shares in aggregate output of around 2.5, and 76% and 24%, respectively. We fix the adjustment cost parameter to $\xi = 0.3261$ such that the model produces an average real return on short term government bonds close to that reported in Jermann (1998). Finally, the persistence of the underlying productivity process is set to $\rho_A = 0.2052$ which corresponds to the continuously compounded value of that in Aruoba et al. (2006), while its volatility is set to $\sigma_A = 0.0307$ to target the relative growth volatilities (relative standard deviations) of consumption and investment to output, and which is consistent with the observed volatility of real GDP growth in the U.S. for the period 1954-1989.

Table 3 reports some of the moments implied by different parameterizations of our RBC model when solved by the First-Order perturbation and a global approximation

Model version	σ_C/σ_Y		σ_I/σ_Y		$R_t^{(0.25)}$	
	Pert.	Global	Pert.	Global	Pert.	Global
Benchmark	0.45	0.44	2.65	2.69	0.37 (4.99)	0.68 (5.17)
No habits, no adj. costs	0.34	0.34	3.00	3.01	4.09 (0.19)	4.11 (0.19)
Habit, no adj. costs	0.13	0.14	3.72	3.72	3.86 (0.26)	4.10 (0.19)
Adj. costs, no habits	1.12	1.11	0.68	0.66	3.77 (0.60)	3.85 (0.61)
U.S. Data (1954-1989)	0.51		2.65		0.80 (5.67)	

Table 3. Moments from simulated data. The different moments are computed using 100,000 draws starting at the deterministic steady state (DSS). The policy functions are computed using both a First-Order perturbation (Pert.) and a global approximation (Global). For comparison, U.S. data moments correspond to those in Jermann (1998). We report the standard deviation (sd) of quarterly growth rates for consumption, σ_C, and investment, σ_I, relative to output, σ_Y, after 10 years; and the mean of the distribution of the three-month yield $R_t^{(0.25)}$ around the DSS after one quarter with sd in brackets (annualized, percentage terms).

method based on collocations. Along with the prototype model of Section 2 (Benchmark), we report the moments for the model without habit formation and no capital adjustment cost of Section 3.2 (No habits, no adj. costs, i.e. $b = X_0 = 0$ and $\xi \to \infty$), no capital adjustment cost (Habit, no adj. costs, i.e, $\xi \to \infty$), and without habits (Adj. costs, no habits, i.e., $b = X_0 = 0$). The last row in the table shows the moments reported by Jermann (1998) for the U.S. between 1954 and 1989. The relative standard deviations for quarterly consumption and investment growth correspond to averages over 100,000 samples generated through an Euler-Maruyama discretization scheme with precision $\Delta = 0.0125$, each of them consisting of 10 years of simulated data, initialized at the DSS. Finally, the table also includes the three-month simulated yield-to-maturity for a zero-coupon bond and the standard deviation of its simulated distribution.

We confirm that only the model with both, habit formation and capital adjustment costs, generates the historical consumption and investment volatility relative to output, and three-month bond yields with sufficient variability. Hence, in this model risk matters quantitatively and we can use the parameterization in Table 2 to investigate the asset pricing errors for the different solution methods (at the deterministic steady state).

5.2 Approximated policy functions

Figure 1 shows the First- and Second-order perturbation approximations to the policy function for consumption around the DSS for our prototype model using the calibration in Table 2. The left panel shows optimal consumption along the capital stock lattice for values 15% below and above its DSS, while keeping the remaining state variables fixed at their deterministic steady state values. The right panel plots optimal consumption along the habit formation lattice covering values that are 15% above and below its DSS. The figures also indicate the DSS and RSS values for consumption, capital stock and habit.

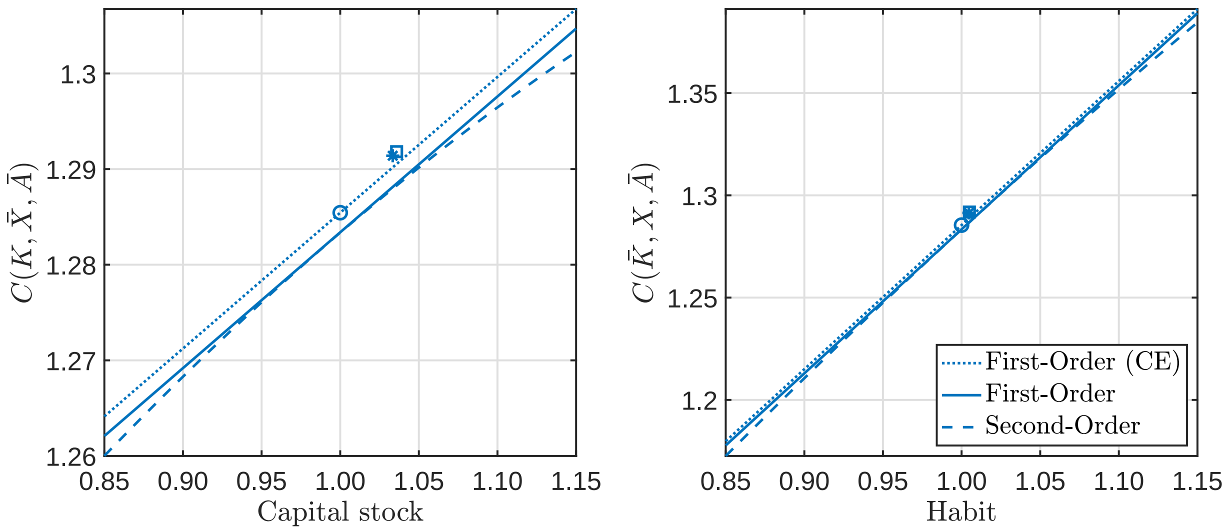

Figure 1. Approximated policy function for consumption: First- and Second-Order approximations to the policy function for consumption around the deterministic steady state (DSS) along the capital stock lattice (left panel) and the habit lattice (right panel), while keeping the remaining state variables at their corresponding DSS values. Values on the horizontal axis represent deviations from DSS. A circle denotes the DSS, a star denotes the first-order approximation to the risky steady state (RSS), and a square denotes the second-order approximation to the RSS.

Their values are reported in Table 4, where we have also included a measure of the RSS computed from a global approximation based on projection methods.

The plot depicts two types of a first-order (linear) approximation to the optimal consumption function. First, it shows the First-Order (CE) by the dotted line, which by construction, is invariant to the amount of volatility in the model, and hence is certainty equivalent. Second, the solid line depicts the First-Order approximation that corresponds to the first-order perturbation solution that breaks certainty equivalence as it includes the (constant) risk correction term $\overline{C}_\eta := C_\eta\left(\overline{K}, \overline{X}, \overline{A}; 0\right) \neq 0$. Hence, while still being a linear approximate solution, it is risk sensitive as its intercept depends on the amount of uncertainty in the model. For comparison, we plot the Second-Order approximation (dashed line) to illustrate the additional risk correction attainable when using higher orders of approximation.

Two things are worth mentioning at this point. First, note that the First-Order (CE) policy function for consumption, which by construction passes through the deterministic steady state, lays above the other two alternative approximations. The reason is that the latter account for the effects of risk, and hence imply lower consumption levels along the entire state space. In particular, the First-Order approximation is parallel to the First-Order (CE), and for values of the state space in a neighborhood of the DSS, it will imply levels of consumption that are relatively close to those suggested by the Second-Order, and hence, nonlinear approximation. Second, the RSS computed from the First- and Second-Order approximations command higher values for the capital stock, habits,

Variable	DSS	RSS			Unconditional Mean		
		First	Second	Global	First	Second	Global
A	0	0	0	0	0	0	0
X	1.0541	1.0589	1.0593	1.0592	1.0535	1.0537	1.0533
K	4.5077	4.6582	4.6693	4.6655	4.5658	4.5722	4.5662
C	1.2854	1.2914	1.2918	1.2917	1.2849	1.2851	1.2846

Table 4. Steady states values and unconditional means. The table reports steady-state values of all state variables and consumption in the model. It includes the deterministic steady-state values (DSS), the first- and second-order approximated risky steady-state (RSS) values, as well as a global approximation to the RSS. The table also reports the unconditional mean of the ergodic distribution. The latter is obtained from simulating the model using a first-order, a second-order and a global approximation to the solution.

and consumption over the long-run, relative to those implied by the deterministic case. This result can be confirmed by looking at Table 4, where we also report the unconditional means of the model variables. Notice that due to a precautionary savings motive, the RSS value and the unconditional mean for the capital stock, K, are larger than the corresponding DSS value. However, the relation between the RSS and the unconditional mean of consumption with its DSS value is ambiguous. Increased savings on the one hand reduce consumption for a given capital stock, while on the other hand they imply a larger capital stock in the future, which in turn increases consumption, *ceteris paribus*. In fact, while the RSS is higher, the unconditional mean of consumption is below its DSS.

A detailed summary of the approximated policy function for consumption is presented in Table 5, where we report the loadings from the first- and second-order perturbations associated to each of the state variables. Columns 2 and 4 show the coefficients for the continuous-time model, while columns 3 and 5 do the same for a discrete-time version of the model described in the Online Appendix.[13] Comparing columns 2 and 3 confirms that a first-order approximation to continuous-time models breaks certainty equivalence. Following our previous discussion, the constant risk correction of -0.0020 implied by our calibration, which is otherwise absent in the solution to the discrete-time model, suggests that a First-Order (CE) approximation overestimates optimal consumption in the presence of uncertainty along the entire state space. Note how a similar risk correction of -0.0025 is obtained in a discrete-time framework when using a second-order, and hence nonlinear, approximation. Comparing columns 4 and 5 reveals that a second-order perturbation in continuous time includes not only an additional adjustment in the constant term of the approximation, $\overline{C}_{\eta\eta} \neq 0$, but also in the slopes of the policy function, implying a time-varying risk correction.[14] As suggested in Andreasen (2012), these two

[13]The first- and second-order approximations to the policy functions that solve the corresponding discrete-time model are computed using `Dynare`.

[14]Based on our results, we conjecture that all the coefficients in (36) are different from zero. Hence, a second-order approximation provides additional risk correction not only through the additional constant

	First-Order		Second-Order	
	Continuous time	Discrete time	Continuous time	Discrete time
$\overline{C}$	1.2854	1.2854	1.2854	1.2854
$\overline{C}_{\eta}$	$\mathbf{-0.0020}$	**0**	-0.0020	0
$\overline{C}_{K}$	0.0315	0.0290	0.0315	0.0290
$\overline{C}_{X}$	0.6680	0.7042	0.6680	0.7042
$\overline{C}_{A}$	0.5370	0.4899	0.5370	0.4899
$\overline{C}_{\eta\eta}$	-	-	$\mathbf{-0.0000}$	$\mathbf{-0.0025}$
$\overline{C}_{K\eta}$	-	-	$\mathbf{-0.0003}$	**0**
$\overline{C}_{X\eta}$	-	-	**0.0020**	**0**
$\overline{C}_{A\eta}$	-	-	$\mathbf{-0.0063}$	**0**
$\overline{C}_{KK}$	-	-	-0.0049	-0.0046
$\overline{C}_{XX}$	-	-	-0.1930	-0.2089
$\overline{C}_{AA}$	-	-	-0.3119	-0.3663
$\overline{C}_{KX}$	-	-	0.0402	0.0389
$\overline{C}_{KA}$	-	-	-0.0282	-0.0286
$\overline{C}_{AX}$	-	-	0.6508	0.6942

Table 5. Loadings of policy function for consumption. The table reports the coefficients from first- and second-order approximations to the policy function for consumption, $C = C(K, X, A; \eta)$, around the deterministic steady state $\left(\overline{C}, \overline{K}, \overline{A}\right)$ for the model in Section 2 and its equivalent discrete-time version.

additional effects can only be achieved in discrete-time models by computing third-order approximations (see Table 1).

5.3 Impulse response functions

Having approximated the unknown policy function, we now compute the impulse-response functions (IRF) in order to compare how the different degrees of approximation capture the amplification and propagation mechanisms of the prototype economy to a temporary shock on the level of TFP. The results are presented in Figure 2, where we plot the transitional dynamics of consumption, capital stock, habits, and output over the course of 60 years after a one-time unexpected increase in TFP equal to σ_A. Prior to the shock, all the variables are assumed equal to their respective stationary values. Thus, while the First-Order (CE) solution is initially resting at the DSS, the First- and Second-Order solutions are resting at their respective (approximate) RSS. Again, the IRF for the CE solution resembles that from a first-order perturbation to an equivalent discrete-time model. For comparison purposes, we report the first- and second-order IRFs for the discrete-time model in the Online Appendix.

Note that the IRFs for the First-Order (CE) lay below the risk-sensitive approximations in Figure 2. Intuitively, since the constant correction term for the First-Order

term, $\mathbf{g}_{\eta\eta}(\overline{\mathbf{x}}; 0)$, but also through the time-varying component, $\mathbf{g}_{\mathbf{x}\eta}(\overline{\mathbf{x}}; 0)$.

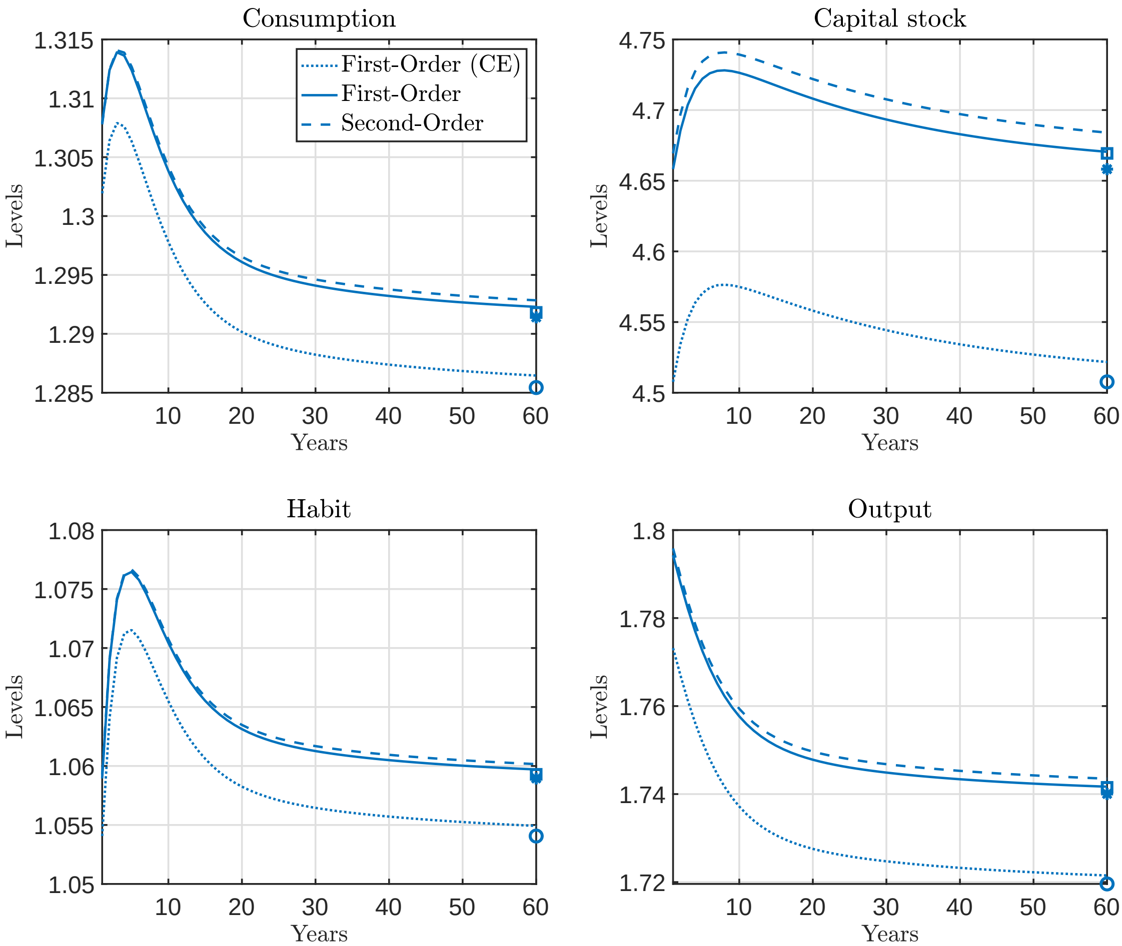

Figure 2. Impulse-Response function to a TFP shock: Responses for the annual levels of consumption, capital stock, habit formation and output to a shock in TFP equivalent to one standard deviation, σ_A. All the variables are assumed to be in their corresponding steady states before the shock. A circle denotes the deterministic steady state, a star denotes the first-order approximation to the risky steady state, and a square denotes the second-order approximation to the risky steady state.

approximation is negative, $\overline{C}_\eta < 0$, one may expect that the consumption response approximated by a First-Order will be below the one approximated by First-Order (CE). However, as shown in Figure 1 and Table 4, the risk-correction in consumption induced by the former leads to a higher risky steady-state capital stock and, thereby, a higher risky steady-state level of consumption. Thus, the fact that the First-Order (CE) is below the First-Order and Second-Order responses is explained by the differences in their fixed points, or long-run levels, and hence cannot be readily interpreted as an indication that certainty equivalent approximations underestimate the response of macroeconomic variables to aggregate shocks. Furthermore, note that the additional risk-corrections provided by the Second-Order approximation have only minor effects on the optimal reaction of consumption to a TFP shock. In other words, the risk-correction in the First-Order approximation provides a sensible approximation to the effects of risk in continuous-time models.

	First-Order (CE)	First-Order	Second-Order
$\varepsilon_a^{(0.25)}$	0.0095	0.0014 ($-$**85.7**)	0.0003 ($-$**96.4**)
$\varepsilon_b^{(0.25)}$	0.0001	0.0001 ($-$45.6)	0.0000 ($-$98.2)
$\varepsilon_c^{(0.25)}$	0.0001	0.0000 ($-$99.2)	0.0000 ($-$98.7)
$\varepsilon_a^{(1)}$	0.0494	0.0065 ($-$**86.8**)	0.0009 ($-$**98.2**)
$\varepsilon_b^{(1)}$	0.0017	0.0009 ($-$47.7)	0.0001 ($-$92.0)
$\varepsilon_c^{(1)}$	0.0018	0.0000 ($-$97.3)	0.0000 ($-$98.8)
$\varepsilon_a^{(5)}$	0.4268	0.1742 ($-$**59.2**)	0.0050 ($-$**98.8**)
$\varepsilon_b^{(5)}$	0.0243	0.0096 ($-$60.3)	0.0069 ($-$71.5)
$\varepsilon_c^{(5)}$	0.0302	0.0017 ($-$94.3)	0.0000 ($-$99.9)

Table 6. Pricing errors for zero-coupon bonds. The table reports the pricing errors $\varepsilon^{(N)}$ and their reduction relative to First-Order (CE) solution at the deterministic steady state for pricing zero-coupon bonds with time-to-maturity of three months ($N = 0.25$), one year ($N = 1$) and five years ($N = 5$), when SDF dynamics are (a) are not observed, (b) partially observed (drift only) and (c) completely observed (drift and diffusion).

5.4 Asset pricing implications

In this section we investigate the asset implications of the different approximations on pricing errors defined in Section 4.2. In particular, we quantify the pricing errors an investor would make if the optimal policy functions are approximated by different perturbation solutions. We assess to what extent the First-Order solution reduces pricing errors relative to the First-Order (CE) solution.

Table 6 reports the absolute pricing errors of zero-coupon bonds with 3-months, 1 year and 5 years time-to-maturity at the deterministic steady state, $\varepsilon^{(N)} := \varepsilon^{(N)}\left(\overline{K}, \overline{X}, \overline{A}\right)$. Columns $2-4$ report errors resulting from First-Order (CE), First-Order, and Second-Order solutions with the reduction in pricing errors relative to First-Order (CE) in parenthesis. Moreover, for each time-to-maturity, we report errors for the cases in which the SDF dynamics are (a) not observed as in (62), (b) partially observed (drift only) as in (63), or (c) fully observed (drift and diffusion) as in (64).

Our results suggest that there are substantial gains from using the (risk-sensitive) First-Order relative to the First-Order (CE) version. Consider the case of a three-month zero-coupon bond with effective price of $P = 0.9983$. An investor using the certainty equivalent linear solution (First-Order (CE)) to price zero-coupon bonds will incur in pricing errors of about 1 dollar for each 100 dollar spent ($\varepsilon_a^{(0.25)} = 0.95\%$ in Column 2). If instead the investor uses the First-Order approximation, the pricing error will be of the order of 10 cents for each 100 dollars spent ($\varepsilon_a^{(0.25)} = 0.14\%$ in Column 3). Hence, breaking certainty equivalence reduces the potential price mismatch by 85.7%, while still remaining in the linear world. The Second-Order approximation further reduces pricing errors which fall to about 3 cents per 100 dollars ($\varepsilon_a^{(0.25)} = 0.03\%$ in Column 4). Sizable gains are also observed for bonds with longer time-to-maturities and/or for the cases when inferring the dynamics of the SDF (drift and diffusion) from the data.

Our decomposition shows that the pricing mismatch primarily results from a poorly approximated risk-free rate. This is important as that misspecified pricing kernel would be used to price any asset in the economy. We find that although pricing errors increase as N increases, they can be substantially reduced if the investor uses the true risk-free rate (or drift of the SDF), which is case (b). In fact, the pricing error for the First-Order approximation is below 10 cents ($\varepsilon_b^{(1)} = 0.09\%$ in Column 3) for the one year time-to-maturity zero-coupon bond. Relative to the First-Order (CE) solution, the risk-adjusted First-Order approximation reduces pricing errors by about 50 percent when using the true risk-free rate. If the investor further knows the true diffusion of the SDF, case (c), then the First-Order approximation also performs well for maturities of 5 years. Here, the error is about 20 cents ($\varepsilon_c^{(5)} = 0.17\%$ in Column 3), which is about 90% lower than the error from First-Order (CE) ($\varepsilon_c^{(5)} = 3.02\%$ in Column 2).

Figure 3 confirms our results for the case of *ex-ante* absolute pricing errors. Investors using a First-Order (CE) solution, when the DGP is the true solution (as approximated by a collocation method), would accept large and persistent pricing errors (see Lettau and Ludvigson, 2009): they range between one to five percent for bonds with time-to-maturities from a quarter to a year ($\varepsilon_a^{(0.25)} = 0.95\%$, $\varepsilon_a^{(1)} = 4.94\%$). On the contrary, those using the First-Order approximation can reduce these pricing errors by more than 85 percent ($\varepsilon_a^{(0.25)} = 0.14\%$, $\varepsilon_a^{(1)} = 0.65\%$). Again, the pricing errors should just illustrate the economic consequences of the approximated policy functions. While the absolute pricing errors could be decreased by including the asset return in the perturbation solution, the relative pricing error reduction by accounting for risk would prevail (compare assets with different maturities).

For illustration, the right panel of Figure 3 decomposes the pricing mismatch made when using the First-Order (CE) solution into: (i) the error stemming from linearization in the presence of uncertainty; and (ii) the error from imposing certainty equivalence in the linear world. As (i) is given by the error resulting from the First-Order approximation, (ii) results as the (absolute) difference of the errors from the First-Order and the First-Order (CE) approximation, and hence provides a measure of pricing error reduction (relative to the true solution).[15] Therefore, the red area measures the pricing error that can be attributed to imposing certainty equivalence in the First-Order (CE) solution, while the blue area measures the error that can be attributed to linearization. As

[15]The accompanying Online Appendix presents an alternative decomposition according to which we decompose the pricing error into: (i) the error stemming from certainty equivalence in the nonlinear world, which would result from a nonlinear certainty equivalent solution; and (ii) the error stemming from linearization under certainty equivalence. We conclude that the errors induced by certainty equivalence and those by linearization are similar under both decompositions, which suggests that the entire error from certainty equivalence is removed by the First-Order approximation.

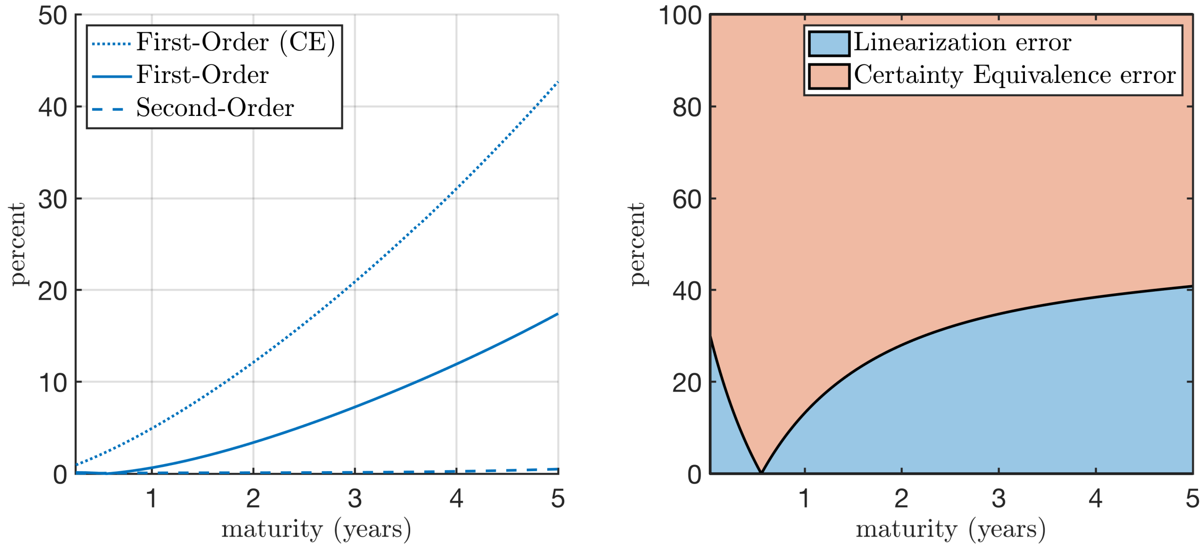

Figure 3. Pricing errors: The left panel plots the absolute pricing errors for different approximations. The right panel decomposes the pricing error incurred by using the First-Order CE solution into a certainty-equivalence component and a linearization component, where the line represents $|(\text{First-Order})/(\text{First-Order (CE)}) - 1|$.

the red area suggests, between 60% and 100% of the error stemming from the certainty-equivalence solution for maturities below five years can be reduced by the risk adjustment of the First-Order solution. For a maturity of 1 year, for instance, 86% of the error in the First-Order (CE) solution can be attributed to the presence of certainty equivalence itself and hence reduced by the First-Order approximation. The remaining error resulting from the First-Order solution, as indicated by the blue area, is inevitable in linear models.

Our results shed light on the key source of the weakness of the linear approximation from perturbation methods that exhibit the certainty-equivalence property. Thus, we can conclude that once the vice of certainty equivalence is discarded, similar to the First-Order approximation in continuous-time models, one may stay with linear models and at the same time account for risk in a reasonable manner.

6 Conclusions

In this paper we use the fact that certainty equivalence (CE) breaks in continuous-time stochastic nonlinear models when their rational expectation solution is linearly approximated around the deterministic steady state. To this end, we generalize the perturbation framework in Judd and Guu (1993) and Parra-Alvarez (2018) by deriving a first-order perturbation approximation to the policy functions of a general class of continuous-time dynamic, stochastic, rational expectation models. We show analytically that the first derivative of the policy function with respect to the perturbation parameter is different from zero at the deterministic steady state. Thus, the resulting linear approximation is risk-sensitive, i.e., breaks the certainty-equivalence property.

Using an otherwise standard RBC model with internal habit formation and capital adjustment costs, which is known to generate substantial risk effects (see Jermann, 1998), we study the economic implications of breaking certainty equivalence in the linear approximation. First, we illustrate the differences between the risk-sensitive and the certainty equivalent first-order approximation by means of policy functions and impulse response functions. We show that the risk-sensitive linear approximation differs substantially from its CE version, and is very close to the second-order approximation. Then, to quantify the risk effects economically, we consider the asset pricing implications and compute pricing errors. This reveals that the risk-sensitive first-order approximation reduces errors of pricing a zero-coupon bond by about 90 percent relative to the CE solution which, by construction, neglects the effects of risk.

We provide intuition for why the first-order perturbation solution in continuous time accounts for prudence and, hence, is not certainty equivalent. This is the result of two complementary points. First, in continuous time it is possible to use Itô's lemma to compute expectations before building the perturbation solution. Second, the perturbation approximation is built around the variance of the shocks that drive the economy, and not around the standard deviation, as it is done in discrete time. The implication is that the linear approximation exhibits a constant correction term that depends on the variance of the shocks.

Our results encourage the use of continuous-time perturbation to account for risk in the class of (approximate) linear models, which are especially useful for the computation and estimation of large-scale macroeconomic models. Given the advantages of perturbation in continuous time, future work should make these advantages more accessible by developing a toolbox that automates perturbation in continuous-time models.

Acknowledgements

We thank Serguei Mailiar, Johannes Pfeiffer, Alexander Meyer-Gohde, Martin M. Andreasen, Oskar Arnt Juul, participants of the 23rd International Conference in Computing in Economics and Finance (New York, 2017), the 13th Dynare Conference (Tokyo, 2017), the 11th International Conference on Computational and Financial Econometrics (London, 2017), the 26th Annual SNDE Symposium (Tokyo, 2018), the 23rd Annual LACEA Meeting (Guayaquil, 2018), the 6th International Conference on Economics of the Turkish Economic Association (Antalya, 2018), the 2020 EEA Annual Congress; and seminar participants for their valuable comments and discussions. We are also grateful to the anonymous reviewers for helpful comments.

Funding

We appreciate financial support from the German Research Foundation (DFG, Project No. 268475812).

Declarations of interest

None.

Appendix

A Deriving the prototype RBC economy

A.1 The HJB equation and the first-order conditions

The benevolent planner chooses a path for consumption in order to maximize the expected discounted life-time utility of a representative household. Define the value of the optimal program as

$$V(K_0, X_0, A_0) = \max_{\{C_t \geq X_t \in \mathbb{R}^+\}_{t=0}^{\infty}} U_0 \quad \text{s.t.} \quad (3) - (8)$$

in which $C_t \geq X_t \in \mathbb{R}^+$ denotes the control variable at instant $t \in \mathbb{R}^+$.

As a first step, we define the *Hamilton-Jacobi-Bellman equation* (HJB) for any $t \in [0, \infty)$

$$\rho V(K_t, X_t, A_t) = \max_{C_t \geq X_t \in \mathbb{R}^+} \left\{ \frac{(C_t - X_t)^{1-\gamma}}{1-\gamma} + \frac{1}{\mathrm{d}t} \mathbb{E}_t \mathrm{d}V(K_t, X_t, A_t) \right\}.$$

Itô's lemma implies (see e.g., Chang, 2009, ch.3, or Wälde, 2012, ch.10)

$$\begin{aligned} \mathrm{d}V(K_t, X_t, A_t) &= V_K(K_t, X_t, A_t)\mathrm{d}K_t + V_X(K_t, X_t, A_t)\mathrm{d}X_t \\ &\quad + V_A(K_t, X_t, A_t)\mathrm{d}A_t + \tfrac{1}{2} V_{AA}(K_t, X_t, A_t)\sigma_A^2 \mathrm{d}t, \end{aligned}$$

where $V_i(K_t, X_t, A_t) := \frac{\partial V_i(K_t, X_t, A_t)}{\partial i}$, and $V_{ij}(K_t, X_t, A_t) := \frac{\partial^2 V(K_t, X_t, A_t)}{\partial i \partial j}$ for $i, j = K_t, X_t, A_t$. If we apply the expectation operator to the integral form, substitute $\mathrm{d}K_t$, $\mathrm{d}X_t$, $\mathrm{d}A_t$, and use the property of stochastic integrals we obtain

$$\begin{aligned} \mathbb{E}_t \mathrm{d}V(K_t, X_t, A_t) = \Bigg[& \left(\Phi\left(\frac{\exp(A_t)K_t^{\alpha} - C_t}{K_t} \right) - \delta \right) K_t V_K(K_t, X_t, A_t) \\ & + (bC_t - aX_t)V_X(K_t, X_t, A_t) - \rho_A A_t V_A(K_t, X_t, A_t) + \tfrac{1}{2}\sigma_A^2 V_{AA}(K_t, X_t, A_t) \Bigg] \mathrm{d}t. \end{aligned}$$

Inserting into the HJB equation yields

$$\begin{aligned} 0 = \max_{C_t \geq X_t \in \mathbb{R}^+} \Bigg\{ & \frac{(C_t - X_t)^{1-\gamma}}{1-\gamma} + \left(\Phi\left(\frac{\exp(A_t)K_t^{\alpha} - C_t}{K_t} \right) - \delta \right) K_t V_K(K_t, X_t, A_t) \\ & + (bC_t - aX_t)V_X(K_t, X_t, A_t) - \rho_A A_t V_A(K_t, X_t, A_t) \\ & + \tfrac{1}{2}\sigma_A^2 V_{AA}(K_t, X_t, A_t) - \rho V(K_t, X_t, A_t) \Bigg\}. \end{aligned}$$

The first-order condition for any interior solution reads

$$(C_t - X_t)^{-\gamma} + bV_X(K_t, X_t, A_t) = \Phi'\left(\frac{\exp(A_t)K_t^\alpha - C_t}{K_t}\right)V_K(K_t, X_t, A_t), \tag{A.1}$$

making optimal consumption an implicit function of the state variables, $C_t = C(K_t, X_t, A_t)$.

A.2 Competitive equilibrium

The maximized (concentrated) HJB equation reads

$$\begin{aligned}0 = \frac{(C(K_t, X_t, A_t) - X_t)^{1-\gamma}}{1-\gamma} + \left(\Phi\left(\frac{\exp(A_t)K_t^\alpha - C(K_t, X_t, A_t)}{K_t}\right) - \delta\right)K_t V_K(K_t, X_t, A_t) \\ + (bC(K_t, X_t, A_t) - aX_t)V_X(K_t, X_t, A_t) - \rho_A A_t V_A(K_t, X_t, A_t) \\ + \tfrac{1}{2}\sigma_A^2 V_{AA}(K_t, X_t, A_t) - \rho V(K_t, X_t, A_t).\end{aligned} \tag{A.2}$$

The system of equations formed by (A.1) and (A.2) determine the unknown functions $V(K_t, X_t, A_t)$ and $C(K_t, X_t, A_t)$ that characterize the equilibrium in the economy. The latter can be alternatively represented by the system of equilibrium PDEs associated to the model costate variables. Differentiation of the maximized HJB equation in (A.2) with respect to the state variables and the application of the envelope theorem, yields:

- the optimal costate variable with respect to the aggregate capital stock, V_K, as

$$\begin{aligned}\rho V_K = \Big(\Phi((\exp(A_t)K_t^\alpha - C_t)/K_t) - \delta\Big)K_t V_{KK} \\ + \Big(\Phi((\exp(A_t)K_t^\alpha - C_t)/K_t) + \Phi'((\exp(A_t)K_t^\alpha - C_t)/K_t)((\alpha - 1)\exp(A_t)K_t^{\alpha-1} \\ + C_t/K_t) - \delta\Big)V_K + \Big(bC_t - aX_t\Big)V_{XK} - \rho_A A_t V_{AK} + \tfrac{1}{2}\sigma_A^2 V_{AAK};\end{aligned} \tag{A.3}$$

- the optimal costate variable with respect to the habit formation level, V_X, as

$$\begin{aligned}\rho V_X = -(C_t - X_t)^{-\gamma} + \Big(\Phi((\exp(A_t)K_t^\alpha - C_t)) - \delta\Big)K_t V_{KX,t} \\ + (bC_t - aX_t)V_{XX} - aV_X - \rho_A A_t V_{AX} + \tfrac{1}{2}\sigma_A^2 V_{AAX};\end{aligned} \tag{A.4}$$

- the optimal costate variable with respect to the total factor productivity, V_A, as

$$\begin{aligned}\rho V_A = \Phi'((\exp(A_t)K_t^\alpha - C_t)/K_t)\exp(A_t)K_t^\alpha V_K \\ + (\Phi((\exp(A_t)K_t^\alpha - C_t)/K_t) - \delta)K_t V_{KA} + (bC_t - aX_t)V_{XA} \\ - \rho_A V_A - \rho_A A_t V_{AA} + \tfrac{1}{2}\sigma_A^2 V_{AAA},\end{aligned} \tag{A.5}$$

where $V_{ijl} := \partial^3 V(K_t, X_t, A_t)/(\partial i \partial j \partial l)$ for any $i, j, l = K_t, X_t, A_t$. Together with the first-order condition in (A.1), they form a system of nonlinear functional equations in the unknown (policy) functions $\{V_K, V_X, V_A, C\} := \{V_K(K_t, X_t, A_t), V_X(K_t, X_t, A_t), V_A(K_t, X_t, A_t), C(K_t, X_t, A_t)\}$, where the dynamics of the state variables are given by the system of controlled SDEs (3), (5), and (7).

A.3 Deterministic steady state

The deterministic steady state (DSS) of the economy is given by the values $\{\overline{C}, \overline{I}, \overline{V}_K, \overline{V}_X, \overline{V}_A, \overline{K}, \overline{X}, \overline{A}\}$ that solve the system of equations

$$
\begin{aligned}
\rho - \Phi(\overline{I}/\overline{K}) - \Phi'(\overline{I}/\overline{K})((\alpha-1)\overline{K}^{\alpha-1} + \overline{C}/\overline{K}) + \delta &= 0, && \text{(A.6)}\\
(\rho + a)\overline{V}_X + (\overline{C} - \overline{X})^{-\gamma} &= 0, && \text{(A.7)}\\
-(\rho + \rho_A)\overline{V}_A + \Phi'(\overline{I}/\overline{K})\overline{K}^{\alpha}\overline{V}_K &= 0, && \text{(A.8)}\\
\Phi\left(\overline{I}/\overline{K}\right) - \delta &= 0, && \text{(A.9)}\\
b\overline{C} - a\overline{X} &= 0, && \text{(A.10)}\\
\left(\overline{C} - \overline{X}\right)^{-\gamma} + b\overline{V}_X - \Phi'\left(\overline{I}/\overline{K}\right)\overline{V}_K &= 0, && \text{(A.11)}\\
\overline{A} &= 0, && \text{(A.12)}
\end{aligned}
$$

which results from imposing $\sigma_A = 0$ together the idle condition $\mathrm{d}K_t/\mathrm{d}t = \mathrm{d}X_t/\mathrm{d}t = \mathrm{d}A_t/\mathrm{d}t = 0$ on the equilibrium PDEs (A.3)-(A.5), and where $\overline{I}/\overline{K} = (\overline{K}^{\alpha} - \overline{C})/\overline{K}$.

The solution to this system of nonlinear equations is entirely determined by the steady state value of the investment-capital ratio, $\overline{I}/\overline{K}$. Given the values for a_1 and a_2, it immediately follows that for any value of ξ, the DSS value for the investment-capital ratio satisfies $\overline{I}/\overline{K} = \delta$. Then, $\Phi(\delta) = \delta$, $\Phi'(\delta) = 1$, and $\Phi''(\overline{I}/\overline{K}) = \Phi''(\delta) = -1/(\xi\delta)$. From (A.6), the steady-state value of the capital stock is $\overline{K} = [\alpha/(\rho+\delta)]^{\frac{1}{1-\alpha}}$. Using the definition of the investment-capital ratio, the steady-state value of consumption is $\overline{C} = \overline{K}^{\alpha} - \delta\overline{K}$. From (A.10) we pin down the steady-state value of the habit as $\overline{X} = \frac{b}{a}\overline{C}$. Finally using (A.7), (A.8), and (A.11) we find the steady-state values for the costate variables $\overline{V}_X = -1/(\rho + a)\left(\overline{C} - \overline{X}\right)^{-\gamma}$, $\overline{V}_K = (1 - b/(\rho+a))\left(\overline{C} - \overline{X}\right)^{-\gamma}$, and $\overline{V}_A = \overline{K}^{\alpha}\overline{V}_K/(\rho + \rho_A)$.

A.4 Model class

Let $\mathbf{y} = [V_K, V_X, V_A]^{\top}$ be the vector of costate variables and $\mathbf{x} = [K, X, A]^{\top}$ the vector of state variables. From the first-order condition (11), optimal consumption $C = C(\mathbf{x}, \mathbf{y})$ is implicitly defined by

$$\Delta(\mathbf{x}, \mathbf{y}, C) = (C - X)^{-\gamma} + bV_X - \Phi'(I/K)V_K = 0,$$

where $\Delta : \mathbb{R}^{n+1} \to \mathbb{R}$ is a continuous and differentiable function, and $I(\mathbf{x}, \mathbf{y}) = \exp(A)K^{\alpha} - C(\mathbf{x}, \mathbf{y})$ is the optimal level of investment. Under certain weak conditions, the implicit function theorem guarantees the local existence of a function C for which $\Delta(\mathbf{x}, \mathbf{y}, C(\mathbf{x}, \mathbf{y})) = 0$ in a neighborhood of a given point $(\mathbf{x}_0, \mathbf{y}_0, C_0)$ for which $\Delta(\mathbf{x}_0, \mathbf{y}_0, C_0) = 0$. Then, the equilibrium of the RBC model with capital adjustment costs and habit formation can be represented by the system of quasilinear PDEs

$$\mathcal{H}(\mathbf{x}, \mathbf{y}, \mathbf{y}_{\mathbf{x}}, \mathbf{y}_{\mathbf{xx}}; \eta) := \mathbf{a}(\mathbf{x}, \mathbf{y}) + \mathbf{y}_{\mathbf{x}} \mathbf{b}(\mathbf{x}, \mathbf{y}) + \eta \mathbf{y}_{\mathbf{xx}} \mathbf{c} = \mathbf{0},$$

with

$$\begin{aligned}
\mathbf{a}(\mathbf{x}, \mathbf{y}) &= \begin{bmatrix} -V_K \left(\rho - \Phi(I/K) - \Phi'(I/K)((\alpha - 1)\exp(A)K^{\alpha-1} + C/K) + \delta\right) \\ -(\rho + a) V_X - (C - X)^{-\gamma} \\ \Phi'(I/K)\exp(A)K^{\alpha} V_K - (\rho_A + \rho) V_A \end{bmatrix}, \\
\mathbf{b}(\mathbf{x}, \mathbf{y}) &= \begin{bmatrix} (\Phi(I/K) - \delta)K \\ bC - aX \\ -\rho_A A \end{bmatrix}, \quad \mathbf{c} = \left[0,\, 0,\, 0,\, 0,\, 0,\, 0,\, 0,\, 0,\, \tfrac{1}{2}\sigma_A^2\right]'.
\end{aligned}$$

Moreover, since the prototype economy corresponds to a stochastic optimal discounted control problem it is possible to decompose $\mathbf{a}(\mathbf{x}, \mathbf{y})$ according to (21) with

$$\begin{aligned}
\widetilde{\mathbf{a}}(\mathbf{x}, \mathbf{y}) &= \begin{bmatrix} 0 \\ -(C - X)^{-\gamma} \\ 0 \end{bmatrix}, \quad \text{and} \\
\mathbf{b}_{\mathbf{x}}^{\top}(\mathbf{x}, \mathbf{y}) &= \begin{bmatrix} (\Phi(I/K) + \Phi'(I/K)((\alpha - 1)\exp(A)K^{\alpha-1} + C/K) - \delta) & 0 & 0 \\ 0 & -a & 0 \\ \Phi'(I/K)\exp(A)K^{\alpha} & 0 & -\rho_A \end{bmatrix}.
\end{aligned}$$

A.5 Stochastic discount factor

When habits in consumption are internal, the agent takes into account the effect of today's consumption decisions on the future levels of habits. Following Detemple and Zapatero (1991), the SDF in this case is given by

$$m_t = \kappa e^{-\rho t} \left\{ (C_t - X_t)^{-\gamma} - b\mathbb{E}_t \left[\int_t^{\infty} e^{-(\rho + a)(s - t)} (C_s - X_s)^{-\gamma} \, \mathrm{d}s \right] \right\}, \tag{A.13}$$

for some given constant κ.

To arrive at equation (56), let us first compute the dynamics of the costate variable with respect to the habit level, $V_X := V_X(K_t, X_t, A_t)$. Using Itô's Lemma, the evolution of the (off-equilibrium) costate variable is given by (see Wälde, 2012)

$$\mathrm{d}V_X = \Big(\Big(\Phi(\exp(A_t)K_t^\alpha - C_t/K_t) - \delta\Big)K_tV_{XK} + (bC_t - aX_t)V_{XX} \\ - \rho_A A_t V_{XA} + \tfrac{1}{2}\sigma_A^2 V_{XAA}\Big)\mathrm{d}t + \sigma_A V_{XA}\mathrm{d}B_{A,t}. \quad \text{(A.14)}$$

Combining equations (A.4) and (A.14) yields the optimal or equilibrium dynamics for V_X as

$$\mathrm{d}V_X = \Big((\rho + a)V_X + (C_t - X_t)^{-\gamma}\Big)\mathrm{d}t + \sigma_A V_{XA}\mathrm{d}B_{A,t}. \quad \text{(A.15)}$$

Multiplying by $e^{-(\rho+a)t}$ on both sides yields

$$e^{-(\rho+a)t}\mathrm{d}V_X = e^{-(\rho+a)t}\left((\rho + a)\,V_X + (C_t - X_t)^{-\gamma}\right)\mathrm{d}t + e^{-(\rho+a)t}V_{XA}\sigma_A\mathrm{d}B_{A,t},$$

or equivalently

$$e^{-(\rho+a)t}(\mathrm{d}V_X - (\rho + a)\,V_X\mathrm{d}t) = e^{-(\rho+a)t}\,(C_t - X_t)^{-\gamma}\,\mathrm{d}t + e^{-(\rho+a)t}V_{XA}\sigma_A\mathrm{d}B_{A,t}.$$

Notice that Itô's formula yields

$$\mathrm{d}(e^{-(\rho+a)t}V_X) \;=\; -(\rho + a)e^{-(\rho+a)t}V_X + e^{-(\rho+a)t}\mathrm{d}V_X$$

such that

$$\mathrm{d}(e^{-(\rho+a)t}V_X) = e^{-(\rho+a)t}\,(C_t - X_t)^{-\gamma}\,\mathrm{d}t + e^{-(\rho+a)t}V_{XA}\sigma_A\mathrm{d}B_{A,t}.$$

Integrating forward in time on both sides yields

$$\begin{aligned}\int_t^T \mathrm{d}(e^{-(\rho+a)s}V_{X,s}) &= \int_t^T e^{-(\rho+a)s}\,(C_s - X_s)^{-\gamma}\,\mathrm{d}s + \int_t^T e^{-(\rho+a)s}V_{XA,s}\sigma_A\mathrm{d}B_{A,s} \\ \Leftrightarrow\; V_{X,t} &= e^{-(\rho+a)(T-t)}V_{X,T} - \int_t^T e^{-(\rho+a)(s-t)}\,(C_s - X_s)^{-\gamma}\,\mathrm{d}s \\ &\quad - \int_t^T e^{-(\rho+a)(s-t)}V_{XA,s}\sigma_A\mathrm{d}B_{A,s}.\end{aligned}$$

Applying the expectation operator (assuming existence of the integrals) implies

$$\mathbb{E}_t\left[V_{X,t}\right] = e^{-(\rho+a)(T-t)}\mathbb{E}_t\left[V_{X,T}\right] - \mathbb{E}_t\left[\int_t^T e^{-(\rho+a)(s-t)}\,(C_s - X_s)^{-\gamma}\,\mathrm{d}s\right].$$

Further, by letting $\lim_{T\to\infty} e^{-(\rho+a)(T-t)}\mathbb{E}\left[V_{X,T}\right]=0$, we may write

$$V_{X,t} := \lim_{T\to\infty} \mathbb{E}_t\left[V_{X,t}\right] \;=\; -\mathbb{E}_t\left[\int_t^\infty e^{-(\rho+a)(s-t)}\left(C_s-X_s\right)^{-\gamma}\mathrm{d}s\right]$$

such that (A.13) can be written as

$$m_t = e^{-\rho t}\Big[\left(C_t-X_t\right)^{-\gamma}+bV_{X_t}\Big], \tag{A.16}$$

and the SDF is

$$m_s/m_t = e^{-\rho(s-t)}\frac{\left(C_s-X_s\right)^{-\gamma}+bV_{X,s}}{\left(C_t-X_t\right)^{-\gamma}+bV_{X,t}}.$$

Notice that the SDF is implicitly determined by the state variables of the economy through its dependence on C_t and $V_{X,t}$. Then, using Itô's lemma, the dynamics of m_t is given by

$$\frac{\mathrm{d}m_t}{m_t} = \mu_m \mathrm{d}t + \sigma_m \mathrm{d}B_{A,t}, \tag{A.17}$$

where the drift and diffusion coefficients are

$$\begin{aligned}
\mu_m \;&=\; -\rho - \frac{\left(C_t-X_t\right)^{-\gamma}}{\left(C_t-X_t\right)^{-\gamma}+bV_{X,t}}\bigg[\gamma\left(C_t-X_t\right)^{-1}\Big(\mu_C-\left(bC_t-aX_t\right)\Big)\\
&\quad -b\left(C_t-X_t\right)^{\gamma}\mu_{V_X}-\tfrac{1}{2}\gamma\left(\gamma+1\right)\left(C_t-X_t\right)^{-2}\sigma_C^2\bigg], \qquad\qquad (A.18)\\
\sigma_m \;&=\; -\frac{\left(C-X\right)^{-\gamma}}{\left(C-X\right)^{-\gamma}+bV_X}\left[\gamma\left(C-X\right)^{-1}\sigma_C-b\left(C-X\right)^{\gamma}\sigma_{V_X}\right]. \qquad\qquad (A.19)
\end{aligned}$$

Note that (A.18) and (A.19) depend on the drift and diffusion coefficients of the policy functions for consumption and the costate variable for habit. By an application of Itô's lemma on $C(K_t,X_t,A_t)$ and $V_X(K_t,X_t,A_t)$, it follows that

$$\begin{aligned}
\mu_C &= C_K\left[\Phi\left(\frac{\exp(A)K^\alpha-C}{K}\right)-\delta\right]K + C_X\left(bC-aX\right) - C_A\rho_A A + \tfrac{1}{2}C_{AA}\sigma_A^2,\\
\mu_{V_X} &= V_{XK}\left[\Phi\left(\frac{\exp(A)K^\alpha-C}{K}\right)-\delta\right]K + V_{XX}\left(bC-aX\right) - V_{XA}\rho_A A + \tfrac{1}{2}V_{XAA}\sigma_A^2,\\
\sigma_C &= C_A\sigma_A,\\
\sigma_{V_X} &= V_{XA}\sigma_A.
\end{aligned}$$

One can alternatively obtain μ_{V_X} directly from (A.15) as $\mu_{V_X}=\left(\rho+a\right)V_X+\left(C-X\right)^{-\gamma}$.

B Proofs

B.1 Proof of Proposition 1

Differentiating (24) with respect to $\mathbf{x}$ yields the $n_x \times n_x$ system of equations

$$\begin{aligned} \mathbf{F}_{\mathbf{x}}(\mathbf{x};\eta) &= \mathbf{a}_{\mathbf{x}} + \mathbf{a}_{\mathbf{y}}\mathbf{g}_{\mathbf{x}} + \boldsymbol{\Omega}^{\mathbf{a}}_{\mathbf{x}} + \mathbf{g}_{\mathbf{x}}\left(\mathbf{b}_{\mathbf{x}} + \mathbf{b}_{\mathbf{y}}\mathbf{g}_{\mathbf{x}} + \boldsymbol{\Omega}^{\mathbf{b}}_{\mathbf{x}}\right) \\ &\quad + \mathbf{g}_{\mathbf{xx}}\left(\mathbf{I}_{n_x}\otimes\mathbf{b}\right) + \eta\mathbf{g}_{\mathbf{xx}}\left(\mathbf{c}_{\mathbf{x}} + \mathbf{c}_{\mathbf{y}}\mathbf{g}_{\mathbf{x}} + \boldsymbol{\Omega}^{\mathbf{c}}_{\mathbf{x}}\right) + \eta\mathbf{g}_{\mathbf{xxx}}\left(\mathbf{I}_{n_x}\otimes\mathbf{c}\right) = \mathbf{0}, \end{aligned} \tag{B.1}$$

where $\boldsymbol{\Omega}^{\mathbf{m}}_{\mathbf{x}} = \mathcal{D}_{\mathbf{x}^\top}\{\mathbf{m}(\cdot,\cdot,\eta\mathbf{y}_{\mathbf{x}})\}$ is given by

$$\begin{aligned} \boldsymbol{\Omega}^{\mathbf{m}}_{\mathbf{x}} &= \eta\left(\left(\operatorname{vec}\mathbf{I}_{n_x}\right)^\top\left(\mathbf{I}_{n_x}\otimes\mathbf{g}_{\mathbf{xx}}\right)\left(\mathbf{K}_{n_x,n_x}\otimes\left(\mathbf{I}_{n_x}\mathbb{1}_{\{\mathbf{m}=\mathbf{a}\vee\mathbf{b}\}} + \mathbf{I}_{n_x^2}\mathbb{1}_{\{\mathbf{m}=\mathbf{c}\}}\right)\right)\otimes\mathbf{I}_{n_x}\right) \\ &\quad\times\left(\mathbf{I}_{n_x}\otimes\mathbf{m}_{(\operatorname{vec}\mathbf{y}_{\mathbf{x}})^\top}\right), \end{aligned}$$

for $\mathbf{m} = \{\mathbf{a},\mathbf{b},\mathbf{c}\}$, $\mathbb{1}_{\{\cdot\}}$ is an indicator function, and $\mathbf{K}_{n_x,n_x}$ is an $n_x^2\times n_x^2$ commutation matrix (see Magnus and Neudecker, 2019). Evaluating (B.1) at the DSS $(\mathbf{x},\mathbf{y},\eta) = (\overline{\mathbf{x}},\overline{\mathbf{y}},0)$ yields

$$\overline{\mathbf{a}}_{\mathbf{x}} + \overline{\mathbf{a}}_{\mathbf{y}}\overline{\mathbf{g}}_{\mathbf{x}} + \overline{\mathbf{g}}_{\mathbf{x}}\overline{\mathbf{b}}_{\mathbf{x}} + \overline{\mathbf{g}}_{\mathbf{x}}\overline{\mathbf{b}}_{\mathbf{y}}\overline{\mathbf{g}}_{\mathbf{x}} = \mathbf{0}.$$

From (21) it follows that

$$\begin{aligned} \overline{\mathbf{a}}_{\mathbf{x}} + \overline{\mathbf{a}}_{\mathbf{y}}\overline{\mathbf{g}}_{\mathbf{x}} &= \overline{\overline{\mathbf{a}}}_{\mathbf{x}} + \overline{\mathbf{b}}_{\mathbf{xx}}\left(\mathbf{I}_{n_x}\otimes\overline{\mathbf{y}}\right) + \overline{\mathbf{b}}_{\mathbf{x}}^\top\overline{\mathbf{g}}_{\mathbf{x}} - \rho\overline{\mathbf{g}}_{\mathbf{x}} \\ &= \overline{\overline{\mathbf{a}}}_{\mathbf{x}} + \overline{\mathbf{b}}_{\mathbf{xx}}\left(\mathbf{I}_{n_x}\otimes\overline{\mathbf{y}}\right) + \overline{\mathbf{b}}_{\mathbf{x}}^\top\overline{\mathbf{g}}_{\mathbf{x}} - \tfrac{1}{2}\rho\overline{\mathbf{g}}_{\mathbf{x}} - \overline{\mathbf{g}}_{\mathbf{x}}\tfrac{1}{2}\rho. \end{aligned}$$

Substituting back yields the desired result

$$\mathbf{A}^\top\overline{\mathbf{g}}_{\mathbf{x}} + \overline{\mathbf{g}}_{\mathbf{x}}\mathbf{A} + \overline{\mathbf{g}}_{\mathbf{x}}\mathbf{C}\overline{\mathbf{g}}_{\mathbf{x}} + \mathbf{B} = \mathbf{0},$$

where $\mathbf{C} = \overline{\mathbf{b}}_{\mathbf{y}}$, $\mathbf{B} = \overline{\overline{\mathbf{a}}}_{\mathbf{x}} + \overline{\mathbf{b}}_{\mathbf{xx}}\left(\mathbf{I}_{n_x}\otimes\overline{\mathbf{y}}\right)$, and $\mathbf{A} = \overline{\mathbf{b}}_{\mathbf{x}} - \frac{\rho}{2}\mathbf{I}_{n_x}$.

B.2 Proof of Theorem 1

Differentiating (24) with respect to η yields the $n_x\times 1$ system of inhomogeneous linear equations

$$\begin{aligned} \mathbf{F}_{\eta}(\mathbf{x};\eta) &= \mathbf{a}_{\mathbf{y}}\mathbf{g}_{\eta} + \boldsymbol{\Omega}^{\mathbf{a}}_{\eta} + \mathbf{g}_{\mathbf{x}}\left(\mathbf{b}_{\mathbf{y}}\mathbf{g}_{\eta} + \boldsymbol{\Omega}^{\mathbf{b}}_{\eta} + \mathbf{b}_{\eta}\right) \\ &\quad + \mathbf{g}_{\mathbf{x}\eta}\mathbf{b} + \mathbf{g}_{\mathbf{xx}}\mathbf{c} + \eta\left(\mathbf{g}_{\mathbf{xx}}\left(\mathbf{c}_{\mathbf{y}}\mathbf{g}_{\eta} + \boldsymbol{\Omega}^{\mathbf{c}}_{\eta}\right) + \mathbf{g}_{\mathbf{xx}\eta}\mathbf{c}\right) = \mathbf{0}, \end{aligned} \tag{B.2}$$

where $\boldsymbol{\Omega}^{\mathbf{m}}_{\eta} = \mathcal{D}_{\eta}\{\mathbf{m}(\cdot,\cdot,\eta\mathbf{y}_{\mathbf{x}})\}$ is given by

$$\boldsymbol{\Omega}^{\mathbf{m}}_{\eta} = \Big[\Big((\operatorname{vec}\mathbf{I}_{n_x})^{\top} (\mathbf{I}_{n_x}\otimes \mathbf{g_x}) + \eta\, (\operatorname{vec}\mathbf{I}_{n_x})^{\top} (\mathbf{I}_{n_x}\otimes \mathbf{g_{x\eta}}) (\mathbf{K}_{n_x,1}\otimes \mathbf{I}_{n_x})\Big) \otimes \big(\mathbf{I}_{n_x}\mathbb{1}_{\{\mathbf{m}=\mathbf{a}\vee\mathbf{b}\}} + \mathbf{I}_{n_x^2}\mathbb{1}_{\{\mathbf{m}=\mathbf{c}\}}\big)\Big]\mathbf{m}_{(\operatorname{vec}\eta\mathbf{y_x})^{\top}},$$

for $\mathbf{m} = \{\mathbf{a},\mathbf{b},\mathbf{c}\}$, $\mathbb{1}_{\{\cdot\}}$ is an indicator function, and $\mathbf{K}_{n_x,1}$ is an $n_x \times n_x$ commutation matrix (see Magnus and Neudecker, 2019). Evaluating (B.2) at the DSS $(\mathbf{x},\mathbf{y},\eta) = (\overline{\mathbf{x}},\overline{\mathbf{y}},0)$, the system of equations reduces to

$$\left(\overline{\mathbf{a}}_{\mathbf{y}} + \overline{\mathbf{g}}_{\mathbf{x}}\overline{\mathbf{b}}_{\mathbf{y}}\right)\overline{\mathbf{g}}_{\eta} + \overline{\boldsymbol{\Omega}}^{\mathbf{a}}_{\eta} + \overline{\mathbf{g}}_{\mathbf{x}}\left(\overline{\boldsymbol{\Omega}}^{\mathbf{b}}_{\eta} + \overline{\mathbf{b}}_{\eta}\right) + \overline{\mathbf{g}}_{\mathbf{xx}}\overline{\mathbf{c}} = \mathbf{0}. \tag{B.3}$$

The desired result follows from solving for $\overline{\mathbf{g}}_{\eta}$ in (B.3).

B.3 Proof of Proposition 2

Define $\mathbf{q}^{\mathbf{m}} := (\operatorname{vec}\mathbf{I}_{n_x})^{\top}(\mathbf{I}_{n_x}\otimes\mathbf{g_{xx}})\left(\mathbf{K}_{n_x,n_x}\otimes\left(\mathbf{I}_{n_x}\mathbb{1}_{\{\mathbf{m}=\mathbf{a}\vee\mathbf{b}\}} + \mathbf{I}_{n_x^2}\mathbb{1}_{\{\mathbf{m}=\mathbf{c}\}}\right)\right)\otimes\mathbf{I}_{n_x}$ and $\mathbf{r}^{\mathbf{m}} := \mathbf{I}_{n_x}\otimes\mathbf{a}_{(\operatorname{vec}\mathbf{y_x})^{\top}}$ for $\mathbf{m} = \{\mathbf{a},\mathbf{b},\mathbf{c}\}$, where $\mathbb{1}_{\{\cdot\}}$ is an indicator function. Then, differentiating (B.1) with respect to $\mathbf{x}$ yields the $n_x \times n_x^2$ system of linear equations

$$\begin{aligned}
\mathbf{F_{xx}}(\mathbf{x};\eta) \;=\;& \mathbf{a_{xx}} + \mathbf{a_{xy}}(\mathbf{g_x}\otimes\mathbf{I}_{n_x}) + \boldsymbol{\Omega}^{\mathbf{a_x}}_{\mathbf{x}} + (\mathbf{a_{yx}} + \mathbf{a_{yy}}(\mathbf{g_x}\otimes\mathbf{I}_{n_x}) + \boldsymbol{\Omega}^{\mathbf{a_y}}_{\mathbf{x}})(\mathbf{I}_{n_x}\otimes\mathbf{g_x})\\
&+\mathbf{a_y}\mathbf{g_{xx}} + \boldsymbol{\Omega}^{\mathbf{a}}_{\mathbf{xx}} + \mathbf{g_{xx}}\left(\mathbf{I}_{n_x}\otimes\left(\mathbf{b_x} + \mathbf{b_y}\mathbf{g_x} + \boldsymbol{\Omega}^{\mathbf{b}}_{\mathbf{x}}\right)\right)\\
&+\mathbf{g_x}\left(\mathbf{b_{xx}} + \mathbf{b_{xy}}(\mathbf{g_x}\otimes\mathbf{I}_{n_x}) + \boldsymbol{\Omega}^{\mathbf{b_x}}_{\mathbf{x}} + \left(\mathbf{b_{yx}} + \mathbf{b_{yy}}(\mathbf{g_x}\otimes\mathbf{I}_{n_x}) + \boldsymbol{\Omega}^{\mathbf{b_y}}_{\mathbf{x}}\right)(\mathbf{I}_{n_x}\otimes\mathbf{g_x})\right.\\
&\left.+\mathbf{b_y}\mathbf{g_{xx}} + \boldsymbol{\Omega}^{\mathbf{b}}_{\mathbf{xx}}\right) + \mathbf{g_{xxx}}(\mathbf{I}_{n_x}\otimes(\mathbf{I}_{n_x}\otimes\mathbf{b}))\\
&+\mathbf{g_{xx}}\left(\mathbf{I}_{n_x}\otimes\left(\mathbf{b_x} + \mathbf{b_y}\mathbf{g_x} + \boldsymbol{\Omega}^{\mathbf{b}}_{\mathbf{x}}\right)\right)\mathbf{K}_{n_x,n_x}\\
&+\eta\mathbf{g_{xxx}}(\mathbf{I}_{n_x}\otimes(\mathbf{c_x} + \mathbf{c_y}\mathbf{g_x} + \boldsymbol{\Omega}^{\mathbf{c}}_{\mathbf{x}})) + \eta\mathbf{g_{xx}}\left(\mathbf{c_{xx}} + \mathbf{c_{xy}}(\mathbf{g_x}\otimes\mathbf{I}_{n_x}) + \boldsymbol{\Omega}^{\mathbf{c_x}}_{\mathbf{x}}\right.\\
&\left.+\left(\mathbf{c_{yx}} + \mathbf{c_{yy}}(\mathbf{g_x}\otimes\mathbf{I}_{n_y}) + \boldsymbol{\Omega}^{\mathbf{c_y}}_{\mathbf{x}}\right)(\mathbf{I}_{n_x}\otimes\mathbf{g_x}) + \mathbf{c_y}\mathbf{g_{xx}} + \boldsymbol{\Omega}^{\mathbf{c}}_{\mathbf{xx}}\right)\\
&+\eta\mathbf{g_{xxxx}}(\mathbf{I}_{n_x}\otimes(\mathbf{I}_{n_x}\otimes\mathbf{c})) + \eta\mathbf{g_{xxx}}(\mathbf{I}_{n_x}\otimes(\mathbf{c_x} + \mathbf{c_y}\mathbf{g_x} + \boldsymbol{\Omega}^{\mathbf{c}}_{\mathbf{x}}))\mathbf{K}_{n_x,n_x} = \mathbf{0},
\end{aligned} \tag{B.4}$$

where $\boldsymbol{\Omega}^{\mathbf{m}}_{\mathbf{x}} = \mathcal{D}_{\mathbf{x}^{\top}}\{\mathbf{m}(\cdot,\cdot,\eta\mathbf{y_x})\}$ is defined in the proof to Proposition 1. Moreover, $\boldsymbol{\Omega}^{\mathbf{m}}_{\mathbf{xx}} = \boldsymbol{\Omega}^{\mathbf{m_x}}_{\mathbf{x}} := \mathcal{D}_{(\mathbf{x}^{\top})^2}\{\mathbf{m}(\cdot,\cdot,\eta\mathbf{y_x})\}$ is given by $\boldsymbol{\Omega}^{\mathbf{m}}_{\mathbf{xx}} = \eta\left(\mathbf{q}^{\mathbf{m}}_{\mathbf{x}}(\mathbf{I}_{n_x}\otimes\mathbf{r}^{\mathbf{m}}) + \mathbf{q}^{\mathbf{m}}\mathbf{r}^{\mathbf{m}}_{\mathbf{x}}\right)$, with

$$\begin{aligned}
\mathbf{q}^{\mathbf{m}}_{\mathbf{x}} \;=\;& (\operatorname{vec}\mathbf{I}_{n_x})^{\top}(\mathbf{I}_{n_x}\otimes\mathbf{g_{xxx}})\left(\mathbf{K}_{n_x,n_x}\otimes\mathbf{I}_{n_x^2}\right)\\
&\times\left(\mathbf{I}_{n_x}\otimes\left(\mathbf{K}_{n_x,n_x}\otimes\left(\mathbf{I}_{n_x}\mathbb{1}_{\{\mathbf{m}=\mathbf{a}\vee\mathbf{b}\}} + \mathbf{I}_{n_x^2}\mathbb{1}_{\{\mathbf{m}=\mathbf{c}\}}\right)\right)\right)\otimes\mathbf{I}_{n_x},
\end{aligned}$$

and

$$\mathbf{r}^{\mathbf{m}}_{\mathbf{x}} \;=\; \left(\mathbf{I}_{n_x}\otimes\left(\mathbf{m}_{(\operatorname{vec}\mathbf{y_x})^{\top}\mathbf{x}} + \mathbf{m}_{(\operatorname{vec}\mathbf{y_x})^{\top}\mathbf{y}}\mathbf{g_x} + \eta\mathbf{m}_{(\operatorname{vec}\mathbf{y_x})^{\top}(\operatorname{vec}\mathbf{y_x})^{\top}}\operatorname{vec}\mathbf{g_{xx}}\right)\right)\mathbf{K}_{n_x,n_x}.$$

Evaluating (B.4) at the DSS $(\mathbf{x},\mathbf{y},\eta) = (\overline{\mathbf{x}},\overline{\mathbf{y}},0)$ yields the desired result.

References

Anderson, E. W., E. R. McGrattan, L. P. Hansen, and T. J. Sargent (1996): "Mechanics of Forming and Estimating Dynamic Linear Economies," in *Handbook of Computational Economics*, ed. by H. M. Amman, D. A. Kendrick, and J. Rust, Elsevier, vol. 1 of *Handbook of Computational Economics*, chap. 4, 171–252.

Andreasen, M. (2012): "On the Effects of Rare Disasters and Uncertainty Shocks for Risk Premia in Non-Linear DSGE Models," *Review of Economic Dynamics*, 15, 295–316.

Andreasen, M. M. and P. Zabczyk (2015): "Efficient Bond Price Approximations in Non-linear Equilibrium-based Term Structure Models," *Studies in Nonlinear Dynamics & Econometrics*, 19, 1–33.

Aruoba, S. B., J. Fernández-Villaverde, and J. Rubio-Ramírez (2006): "Comparing Solution Methods for Dynamic Equilibrium Economies," *Journal of Economic Dynamics and Control*, 30, 2477–2508.

Binning, A. (2013a): "Solving second and Third-order Approximations to DSGE Models: A Recursive Sylvester Equation Solution," Working Paper 2013/18, Norges Bank.

——— (2013b): "Third-order Approximation of Dynamic Models Without the Use of Tensors," Working Paper 2013/13, Norges Bank.

Boldrin, M., L. J. Christiano, and J. D. M. Fisher (2001): "Habit Persistence, Asset Returns, and the Business Cycle," *American Economic Review*, 91, 149–166.

Brunnermeier, M. K. and Y. Sannikov (2014): "A Macroeconomic Model with a Financial Sector," *American Economic Review*, 104, 379–421.

Buiter, W. H. (1984): "Saddlepoint Problems in Continuous Time Rational Expectations Models: A General Method and Some Macroeconomic Examples," *Econometrica*, 52, 665–680.

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

Chang, F.-R. (2009): *Stochastic Optimization in Continuous Time*, Cambridge University Press, second ed.

Christensen, B. J., O. Posch, and M. van der Wel (2016): "Estimating Dynamic Equilibrium Models Using Mixed Frequency Macro and Financial Data," *Journal of Econometrics*, 194, 116 – 137.

Cochrane, J. H. (2005): *Asset Pricing*, Princeton University Press, 2 ed.

Coeurdacier, N., H. Rey, and P. Winant (2011): "The Risky Steady State," *American Economic Review*, 101, 398–401.

Collard, F. and M. Juillard (2001): "A Higher-Order Taylor Expansion Approach to Simulation of Stochastic Forward-Looking Models with an Application to a Nonlinear Phillips Curve Model," *Computational Economics*, 17, 125–139.

Constantinides, G. M. (1990): "Habit Formation: A Resolution of the Equity Premium Puzzle," *Journal of Political Economy*, 98, 519–543.

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

Detemple, J. B. and F. Zapatero (1991): "Asset Prices in an Exchange Economy with Habit Formation," *Econometrica*, 59, 1633–1657.

Fernández-Villaverde, J., J. Rubio-Ramírez, and F. Schorfheide (2016): "Solution and Estimation Methods for DSGE Models," in *Handbook of Macroeconomics*, ed. by J. B. Taylor and H. Uhlig, Elsevier, vol. 2 of *Handbook of Macroeconomics*, chap. 9, 527 – 724.

Fernández-Villaverde, J. and J. F. Rubio-Ramírez (2007): "Estimating Macroeconomic Models: A Likelihood Approach," *The Review of Economic Studies*, 74, 1059–1087.

Fleming, W. H. (1971): "Stochastic Control for Small Noise Intensities," *SIAM Journal on Control*, 9, 473–517.

Gaspar, J. and K. Judd (1997): "Solving Large-Scale Rational-Expectations Models," *Macroeconomic Dynamics*, 1, 45–75.

Gomme, P. and P. Klein (2011): "Second-order Approximation of Dynamic Models Without the Use of Tensors," *Journal of Economic Dynamics and Control*, 35, 604–615.

Hansen, L. P. and T. Sargent (2014): *Recursive Models of Dynamic Linear Economies*, The Gorman Lectures in Economics, Princeton University Press.

Harvey, A. (2006): "Forecasting with Unobserved Components Time Series Models," Elsevier, vol. 1, chap. 07, 327–412, 1 ed.

Harvey, A. C. and J. H. Stock (1985): "The Estimation of Higher-Order Continuous Time Autoregressive Models," *Econometric Theory*, 1, 97–117.

Jermann, U. J. (1998): "Asset Pricing in Production Economies," *Journal of Monetary Economics*, 41, 257–275.

Jin, H.-H. and K. L. Judd (2002): "Perturbation Methods for General Dynamic Stochastic Models," Tech. rep.

Judd, K. L. (1996): "Approximation, Perturbation, and Projection Methods in Economic Analysis," in *Handbook of Computational Economics*, ed. by H. M. Amman, D. A. Kendrick, and J. Rust, Elsevier, vol. 1 of *Handbook of Computational Economics*, chap. 12, 509–585.

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

Judd, K. L. and S.-M. Guu (1993): "Perturbation Methods for Economic Growth Models," in *Economic and financial modeling with Mathematica*, ed. by H. R. Varian, Springer-Verlag, chap. 2, 80–103.

Kim, J., S. Kim, E. Schaumburg, and C. A. Sims (2008): "Calculating and Using Second-order Accurate Solutions of Discrete Time Dynamic Equilibrium Models," *Journal of Economic Dynamics and Control*, 32, 3397–3414.

Kimball, M. S. (1990a): "Precautionary Saving and the Marginal Propensity to Consume," NBER Working Papers 3403, National Bureau of Economic Research, Inc.

——— (1990b): "Precautionary Saving in the Small and in the Large," *Econometrica*, 58, 53–73.

Kydland, F. E. and E. C. Prescott (1982): "Time to Build and Aggregate Fluctuations," *Econometrica*, 50, 1345–70.

Lan, H. and A. Meyer-Gohde (2013): "Solving DSGE Models with a Nonlinear Moving Average," *Journal of Economic Dynamics and Control*, 37, 2643–2667.

——— (2014): "Solvability of Perturbation Solutions in DSGE Models," *Journal of Economic Dynamics and Control*, 45, 366–388.

Lettau, M. and S. Ludvigson (2009): "Euler Equation Errors," *Review of Economic Dynamics*, 12, 255–283.

Levintal, O. (2017): "Fifth-order Perturbation Solution to DSGE Models," *Journal of Economic Dynamics and Control*, 80, 1–16.

Lopez, P., D. Lopez-Salido, and F. Vazquez-Grande (2018): "Risk-Adjusted Linearizations of Dynamic Equilibrium Models," Tech. rep.

Magnus, J. R. and H. Neudecker (2019): *Matrix Differential Calculus with Applications in Statistics and Econometrics*, John Wiley, third ed.

Merton, R. C. (1971): "Optimum Consumption and Portfolio Rules in a Continuous-Time Model," *Journal of Economic Theory*, 3, 373–413.

Meyer-Gohde, A. (2015): "Risk-Sensitive Linear Approximations," Annual Conference 2015 (Muenster): Economic Development - Theory and Policy 113057, Verein für Socialpolitik / German Economic Association.

Parra-Alvarez, J. C. (2018): "A Comparison of Numerical Methods for the Solution of Continuous-time DSGE Models," *Macroeconomic Dynamics*, 22, 1555 – 1583.

Posch, O. (2020): "Resurrecting the New-Keynesian Model: (Un)conventional Policy and the Taylor rule," CESifo Working Paper Series 6925.

Ruge-Murcia, F. (2012): "Estimating Nonlinear DSGE Models by the Simulated Method of Moments: With an Application to Business Bycles," *Journal of Economic Dynamics and Control*, 36, 914 – 938.

Schmitt-Grohe, S. and M. Uribe (2004): "Solving Dynamic General Equilibrium Models Using a Second-order Approximation to the Policy Function," *Journal of Economic Dynamics and Control*, 28, 755–775.

Simon, H. A. (1956): "Dynamic Programming Under Uncertainty with a Quadratic Criterion Function," *Econometrica*, 24, 74–81.

Singer, H. (1998): "Continuous Panel Models with Time Dependent Parameters," *The Journal of Mathematical Sociology*, 23, 77–98.

Steger, T. (2005): "Stochastic Growth under Wiener and Poisson Uncertainty," *Economics Letters*, 86, 311–316.

Theil, H. (1957): "A Note on Certainty Equivalence in Dynamic Planning," *Econometrica*, 25, 346–349.

Vetter, W. J. (1973): "Matrix Calculus Operations and Taylor Expansions," *SIAM Review*, 15, 352–369.

Wälde, K. (2011): "Production Technologies in Stochastic Continuous Time Models," *Journal of Economic Dynamics and Control*, 35, 616–622.

——— (2012): *Applied Intertemporal Optimization*, Gutenberg University Mainz.

Zeldes, S. P. (1989): "On the Concavity of the Consumption Function," *The Quarterly Journal of Economics*, 104, 275–298.

Zhou, K., J. Doyle, and K. Glover (1996): *Robust and Optimal Control*, Prentice Hall.