

A comparison of numerical methods for the solution of continuous-time DSGE models

Juan Carlos Parra-Alvarez

CREATES Research Paper 2013-39

A comparison of numerical methods for the solution of continuous-time DSGE models*

Juan Carlos Parra-Alvarez[†]

jparra@creates.au.dk

Department of Economics and Business, and CREATES

Aarhus University

November 2013

Abstract

This paper evaluates the accuracy of a set of techniques that approximate the solution of continuous-time DSGE models. Using the neoclassical growth model I compare linear-quadratic, perturbation and projection methods. All techniques are applied to the HJB equation and the optimality conditions that define the general equilibrium of the economy. Two cases are studied depending on whether a closed form solution is available. I also analyze how different degrees of non-linearities affect the approximated solution. The results encourage the use of perturbations for reasonable values of the structural parameters of the model and suggest the use of projection methods when a high degree of accuracy is required.

Keywords: Continuous-Time DSGE Models, Linear-Quadratic Approximation, Perturbation Method, Projection Method.

JEL classification: C63, C68, E32.

1 Introduction

As described in Blanchard (2009) during the last 20 years macroeconomy has experienced a convergence in methodology to what it is known today as Dynamic Stochastic General Equilibrium models (DSGE). In general, it is not possible to derive analytical solutions for the policy functions that

*I would like to thank the participants of the 18th Annual Conference on Computing in Economics and Finance in Prague, the 6th CSDA International Conference on Computational and Financial Econometrics in Oviedo and the CDMA Workshop on DSGE models in St. Andrews for their helpful comments. I also thank Olaf Posch, Bent Jesper Christensen, Jesús Fernández-Villaverde, Kenneth Judd, Martin Møller Andreasen, Willi Semmler and seminar participants at CREATES for useful suggestions and insights. The author acknowledges support from CREATES - Center for Research in Econometric Analysis of Time Series (DNRF78), funded by the Danish National Research Foundation.

[†]School of Business and Social Sciences, Fuglesangs Allé 4, 8210 Aarhus V, Denmark. Phone Number: (+45) 87 16 55 03

describe the equilibrium outcomes of these models, except under very restrictive and economically uninteresting assumptions. Thus, this convergence has been accompanied with the application and further development of numerical techniques to approximate this unknown solution. Currently, the toolbox of macroeconomists includes a variety of methods that differ in their accuracy and computational cost. These differences were initially studied in Taylor and Uhlig (1990) and later on in Aruoba et al. (2006) for discrete-time economies. More recently Binsbergen et al. (2008) and Caldara et al. (2012) have analyzed the accuracy of numerical methods for the case of discrete-time models with recursive preferences and stochastic volatility. However, little attention has been paid in this regard to structural stochastic dynamic models in continuous-time. This paper attempts to fill this gap in the literature and therefore it complements previous work by Judd (1996), Gaspar and Judd (1997), Judd (1998), Miranda and Fackler (2002), Kompas and Chu (2010) and Posch and Trimborn (2013).

Macroeconomic modeling in continuous-time exhibits some features that make it, in some cases, more attractive than its discrete-time counterpart both from a theoretical and a computational perspective. As stated in Turnovsky (1977): “Both kind of models have their place and the choice between them is often dictated by convenience”. However, continuous-time models are sometimes preferred because its analytical tractability. In fact, it is possible to derive closed form solutions for a wider class of models without the need of strong parametric restrictions. This analytical advantage it is attributable to Ito’s lemma which provides more flexibility for transforming expressions. Furthermore, discrete-time models assume that all decisions made by agents are perfectly synchronized and the distinction between flows and stocks become obscured, situations that are easily overcome in a continuous-time framework.

Setting up DSGE models in continuous-time also allows to link the real side of the economy with all the different devices built in the finance literature which have been useful to explain the behavior of asset prices. As stated in Cochrane (2005): “[...]finance has a lot to say about macroeconomics, and it says that something is desperately wrong with most macroeconomic models” when he refers to the inadequacy of the state of the art discrete-time DSGE models to match financial market facts. Furthermore, the continuous-time framework allows to incorporate parameter uncertainty on the mean and variances of the stochastic processes that drive the optimal decisions in an easier way than the discrete-time setup while retaining a tractable model. Examples of the potential benefits of continuous-time macro finance models include Brennan (1998), Grinols and Turnovsky (1993), Xia (2001), Posch (2011), Wachter (2013), among many others.

In addition, due to the martingale property of the Lévy processes commonly used to model the random shocks that hit the economy, the optimality conditions that describe the equilibrium in a continuous-time setup are deterministic (Chang, 2009). As a byproduct of this result, the computational cost and numerical errors made by approximating the solution can be reduced since there is no need to compute expected values as it is the case in discrete-time. Another advantage of the continuous-time framework is the potential reduction in the curse of dimensionality, as shown in Doraszelski and Judd (2012). In fact, the Bellman equation associated with discrete-time problems

requires the numerical computation of integrals of compound unknown functions. However, the dynamic programming equation in continuous-time does not include any composition of functions or expectation operators speeding up the numerical approximations.

In this paper, I study the performance of different numerical methods to compute an approximated solution of continuous-time DSGE models based on the maximized Hamilton-Jacobi-Bellman (HJB) equation and the first order conditions that describe the general equilibrium of the economy. Despite being well known procedures, little is known about their relative performance, an issue that becomes relevant if we want to proceed with the econometric estimation of the deep parameters of the model. Then, it is necessary to assess which approximation method is more accurate and fast to make the estimation feasible without introducing possible bias due to approximation errors in the solution step. I report the results obtained from three main approaches: first and second order perturbation; projection methods with collocation and least squares residual functions; and linear-quadratic approximations with linear and linearized constraints. Alternative ways to solve the model based on the Euler equations is currently in the research agenda¹.

All the methods are evaluated in terms of accuracy and computing time using as a benchmark the stochastic neoclassical growth model with endogenous labor which is the workhorse model in the DSGE literature. The basic framework used in Merton (1990) is extended with leisure in the household's utility function and with stochastic and stationary variation in the total factor productivity. The model can be thought as a continuous-time representation of the model studied in Aruoba et al. (2006). The model is chosen since under suitable parametrizations it allows for a closed form solution that can be used to assess the different approximation methods². In particular, I compute the relative numerical error made by using the approximations instead of the true solution. For the case where an analytical solution is not available, I use the HJB equation residuals as a measure of accuracy. Robustness checks are done by using different calibrations which differ in the degree of non-linearities introduced into the model. I study the effects of higher volatility both in the stochastic depreciation and/or the productivity shocks and the effects of lower and higher risk aversion.

Four main results are obtained from this exercise. First, I find that despite their conceptual and computational differences, all methods provide an acceptable degree of accuracy around the deterministic steady state. This result is robust to different values of the volatility of the exogenous processes as long as the risk aversion parameter is not unreasonably high. Otherwise perturbation and linear-quadratic methods rapidly deteriorate relative to global methods. Among all the alternatives, projection techniques are the most reliable and stable. However a good initial guess for the value and policy functions is required in order to get an accurate approximation, compromising this way the reliability of the method.

¹The method proposed in Posch and Trimborn (2013) follows this approach to solve for the unknown policy functions.

²The findings reported in this paper are model-dependent but they help to understand the different approximation methods. Hopefully, the main conclusions derived under this benchmark model can be also obtained under alternative setups.

Second, even though perturbation is only locally accurate, the increase in the order of approximation improves substantially the goodness of fit by taking into account possible non-linearities of the functions being approximated. This can be done at a low extra computational cost encouraging its use. This becomes an advantage for global methods since the resulting approximation becomes an adequate initial guess.

Third, linear-quadratic approximations are on average less accurate than perturbation methods despite the fact that both of them use the same local information to build the approximation. In addition when the method is applied to linearized constraints, the policy functions are incorrectly approximated outside the deterministic steady state. However, if the constraints of the original problem are linear then the optimal rules obtained are correctly approximated. On the other hand, the approximation of the value function is not affected considerably by the use of linear or linearized constraints. Similar to the perturbation case, the low implementation and computation cost suggest the use of linear-quadratic approximations as an initial guess for global methods.

Fourth, continuous-time DSGE modeling proves to be a promising area of future policy-oriented research when compared to the discrete-time framework. All the approximation methods use much less computing time since there is no need to approximate composition of unknown functions, neither numerically approximate the integrals associated with expected values.

The first three results are similar to those derived from the numerical solution of DSGE models in discrete-time. As suggested in Aruoba et al. (2006), they should stimulate the use of perturbation methods, preferably with orders of approximation higher than one, and suggest relying on projection techniques whenever high accuracy and stability is needed.

The rest of the paper is organized as follows. Section 2 describes the general stochastic control problem and summarizes the different methods used to approximate its solution. Section 3 presents the benchmark model that will be used to test the different numerical techniques and establishes the conditions under which a closed form solution exists. Section 4 reports the numerical results and accuracy measures obtained under different setups and calibrations, and presents a brief discussion on computing time. Finally, Section 5 concludes.

2 General problem and solution methods

Consider the general dynamic optimization framework in which an agent tries to achieve the maximum possible payoff out of her decisions given the state of the economy at each instant of time. In general, a solution to this problem cannot be found analytically and it is necessary to use numerical methods to approximate it. This section presents a brief overview of the different approximation methods that will be used in the rest of the paper. A detailed explanation of each of them can be found in the references presented in each of the following subsections.

Let (Ω, \mathcal{F}, P) be a filtered probability space with filtration $\{\mathcal{F}_t\}$, $\mathbf{s}_t \in \mathbf{S} \subset \mathbb{R}_+^m$ an m -dimensional vector of state variables at time t with right-continuous sample paths, left-hand limit and initial value $\mathbf{s}_0 = \mathbf{s}$ which include predetermined endogenous and exogenous variables, and $\mathbf{c}_t \in \mathcal{D}(\mathbf{S}) \subset$

\mathbb{R}_+^n an n -dimensional vector of non-negative endogenous-control variables at time t whose coordinates are functions of the state variables. The set $\mathcal{D}(\mathbf{S})$ denotes the set of admissible controls. The evolution of the state vector is given by the autonomous controlled diffusion process:

$$d\mathbf{s}_t = \mu(\mathbf{s}_t, \mathbf{c}_t) dt + \Sigma^{\frac{1}{2}}(\mathbf{s}_t, \mathbf{c}_t) d\mathbf{B}_t \quad (1)$$

where $\mu(\mathbf{s}_t, \mathbf{c}_t) \in \mathbb{R}^m$ and $\Sigma^{\frac{1}{2}}(\mathbf{s}_t, \mathbf{c}_t) \in \mathbb{R}^{m \times p}$ are the real-valued drift vector Lipschitzian function and real-valued square root variance-covariance matrix Lipschitzian function. \mathbf{B}_t is a p -dimensional vector of standard Brownian motions representing stochastic disturbances that affect the state variables. Define $\Sigma(\mathbf{s}_t, \mathbf{c}_t) = \Sigma^{\frac{1}{2}}(\mathbf{s}_t, \mathbf{c}_t) \Sigma^{\frac{1}{2}}(\mathbf{s}_t, \mathbf{c}_t)' \in \mathbb{R}^{m \times m}$ to be the variance-covariance matrix of those disturbances.

The decision problem of the agent corresponds to the autonomous discounted stochastic optimal control problem:

$$V(\mathbf{s}_0) = \max_{\{\mathbf{c}_t\}_{t=0}^{\infty} \in \mathcal{D}(\mathbf{s}) \times \mathcal{D}(\mathbf{s}) \times \dots} \mathbb{E}_0 \int_0^{\infty} e^{-\rho t} \pi(\mathbf{s}_t, \mathbf{c}_t) dt$$

subject to equation (1) and $\mathbf{s}_0 = \mathbf{s}$ given, where \mathbb{E}_0 denotes the expectation operator conditional on the information available at time $t = 0$ and $\pi(\mathbf{s}_t, \mathbf{c}_t) \in \mathbb{R}$ is the time-homogenous continuous and integrable objective function. For simplicity it is assumed that the variance-covariance matrix of the disturbances is control independent. The function $V(\mathbf{s}_0) \in \mathbb{R}$ denotes the maximum expected value, or value function, obtained at $t = 0$ when $\mathbf{s} = \mathbf{s}_0$.

A necessary condition for optimality is given by the HJB equation:

$$0 = \max_{\mathbf{c}_t \in \mathcal{D}(\mathbf{s}_t)} \left\{ \pi(\mathbf{s}_t, \mathbf{c}_t) - \rho V(\mathbf{s}_t) + \mu(\mathbf{s}_t, \mathbf{c}_t) \nabla V(\mathbf{s}_t) + \frac{1}{2} \text{trace}(\Sigma(\mathbf{s}_t) \nabla^2 V(\mathbf{s}_t)) \right\} \quad (2)$$

where $\nabla V(\mathbf{s}_t)$ is the gradient of the value function and $\nabla^2 V(\mathbf{s}_t)$ is the associated Hessian matrix.

The first order conditions for any interior solution are:

$$\pi_c(\mathbf{s}_t, \mathbf{c}_t) + \mu_c(\mathbf{s}_t, \mathbf{c}_t) \nabla V(\mathbf{s}_t) = 0 \quad (3)$$

for each $c_t \in \mathbf{c}_t$, which implicitly makes the vector of controls a function of the state vector, $\mathbf{c}_t = \mathcal{P}(\mathbf{s}_t)$. The vector function $\mathcal{P} : \mathbb{R}^m \rightarrow \mathcal{D}(\mathbf{s}_t)$ maps every possible value of the state vector at time t into the control vector. The implicit maximized (concentrated) HJB equation reads:

$$0 = \pi(\mathbf{s}_t, \mathcal{P}(\mathbf{s}_t)) - \rho V(\mathbf{s}_t) + \mu(\mathbf{s}_t, \mathcal{P}(\mathbf{s}_t)) \nabla V(\mathbf{s}_t) + \frac{1}{2} \text{trace}(\Sigma(\mathbf{s}_t) \nabla^2 V(\mathbf{s}_t)) \quad (4)$$

which together with the first order conditions determine the unknown functions $V(\mathbf{s}_t)$ and $\mathcal{P}(\mathbf{s}_t)$ and form the basis for all the numerical procedures to be introduced later. A solution to the HJB equation amounts to find these unknown functions such that they solve the continuum of problems described by equation (2). The mapping \mathcal{P} is called policy function and represents the optimal response of agents to a given set of values of the possible states.

The following subsections briefly describe some alternative numerical methods that can be used to approximate the solution of the dynamic optimization problem described above. I do not pretend to cover all the techniques proposed in the stochastic control literature but evaluate the performance of similar methods to the ones usually employed in the discrete-time analysis of DSGE models. In particular I use linear-quadratic approximations, perturbations and projection methods. Alternative techniques can be found in Tapiero and Sulem (1994), Candler (2004) and Posch and Trimborn (2013)

2.1 Linear-Quadratic approximation

Linear-quadratic approximation uses local information to replace the original non-linear problem with a more tractable problem for which a solution can be easily found. In fact, it is known that the resulting policy function is linear on the state variables. Linear-quadratic approximations have been extensively used in economics. First attempts to apply the technique in the field can be found in Magill (1977b) for the deterministic case and in Magill (1977a) for the stochastic case, eventhough the method only became popular after the seminal paper of Kydland and Prescott (1982). Further developments and implementation alternatives can be found in Anderson et al. (1996).

Let $\hat{\mathbf{s}}_t = (\mathbf{s}_t - \mathbf{s}^{ss})$ and $\hat{\mathbf{c}}_t = (\mathbf{c}_t - \mathbf{c}^{ss})$ denote the absolute deviations of the state and control variables from their deterministic steady state values. The linear-quadratic method approximates the original non-linear problem in $(\hat{\mathbf{s}}_t, \hat{\mathbf{c}}_t)$ by constructing a second order Taylor expansion of the objective function and a first order Taylor expansion of the stochastic differential equations that describe the evolution of the state variables around their steady state values. The approximated stochastic control problem is given by:

$$V(\hat{\mathbf{s}}_0) = \max_{\{\hat{\mathbf{c}}_t\}_{t=0}^{\infty}} \mathbb{E}_0 \int_0^{\infty} e^{-\rho t} (\hat{\mathbf{s}}_t' \mathbf{R} \hat{\mathbf{s}}_t + \hat{\mathbf{c}}_t' \mathbf{Q} \hat{\mathbf{c}}_t + 2\hat{\mathbf{c}}_t' \mathbf{W}' \hat{\mathbf{s}}_t) dt$$

subject to

$$d\hat{\mathbf{s}}_t = (\mu_{\mathbf{s}}(\mathbf{s}^{ss}, \mathbf{c}^{ss}) \hat{\mathbf{s}}_t + \mu_{\mathbf{c}}(\mathbf{s}^{ss}, \mathbf{c}^{ss}) \hat{\mathbf{c}}_t) dt + \Sigma^{\frac{1}{2}}(\mathbf{s}^{ss}) d\mathbf{B}_t$$

where now $\hat{\mathbf{s}}_t$ is an $(1+m) \times 1$ vector with the first element being the constant 1 and the matrices $\mu_{\mathbf{s}}(\mathbf{s}^{ss}, \mathbf{c}^{ss})$, $\mu_{\mathbf{c}}(\mathbf{s}^{ss}, \mathbf{c}^{ss})$ and $\Sigma^{\frac{1}{2}}(\mathbf{s}^{ss})$ are of dimensions $(1+m) \times (1+m)$, $(1+m) \times n$ and $(1+m) \times p$ respectively with the elements of their first rows all equal to zero. Furthermore, \mathbf{R} , \mathbf{Q} and \mathbf{W} are $(1+m) \times (1+m)$, $n \times n$ and $(1+m) \times n$ matrices given by:

$$\mathbf{R} = \frac{1}{2} \begin{bmatrix} 2\pi(\mathbf{s}^{ss}, \mathbf{c}^{ss}) & \pi_{\mathbf{s}}(\mathbf{s}^{ss}, \mathbf{c}^{ss}) \\ \pi_{\mathbf{s}}(\mathbf{s}^{ss}, \mathbf{c}^{ss})' & \pi_{\mathbf{ss}}(\mathbf{s}^{ss}, \mathbf{c}^{ss}) \end{bmatrix},$$

$$\mathbf{Q} = \frac{1}{2} \pi_{\mathbf{cc}}(\mathbf{s}^{ss}, \mathbf{c}^{ss}),$$

$$\mathbf{W} = \frac{1}{2} \begin{bmatrix} \pi_{\mathbf{c}}(\mathbf{s}^{ss}, \mathbf{c}^{ss}) \\ \pi_{\mathbf{sc}}(\mathbf{s}^{ss}, \mathbf{c}^{ss}) \end{bmatrix}.$$

Following Ljungqvist and Sargent (2004) I introduce an equivalent problem without cross-products between the states and controls. The transformed stochastic control problem becomes:

$$V(\hat{\mathbf{s}}_0) = \max_{\{\hat{\mathbf{c}}_t\}_{t=0}^{\infty}} \mathbb{E}_0 \int_0^{\infty} e^{-\rho t} \left(\hat{\mathbf{s}}_t' \bar{\mathbf{R}} \hat{\mathbf{s}}_t + \hat{\mathbf{c}}_t' \mathbf{Q} \hat{\mathbf{c}}_t \right) dt$$

subject to

$$d\hat{\mathbf{s}}_t = (\bar{\mu}_{\mathbf{s}}(\mathbf{s}^{ss}, \mathbf{c}^{ss}) \hat{\mathbf{s}}_t + \mu_{\mathbf{c}}(\mathbf{s}^{ss}, \mathbf{c}^{ss}) \hat{\mathbf{c}}_t) dt + \Sigma^{\frac{1}{2}}(\mathbf{s}^{ss}) d\mathbf{B}_t$$

where

$$\bar{\mathbf{R}} = \mathbf{R} - \mathbf{W}\mathbf{Q}^{-1}\mathbf{W}'$$

$$\bar{\mu}_{\mathbf{s}}(\mathbf{s}^{ss}, \mathbf{c}^{ss}) = \mu_{\mathbf{s}}(\mathbf{s}^{ss}, \mathbf{c}^{ss}) - \mu_{\mathbf{c}}(\mathbf{s}^{ss}, \mathbf{c}^{ss}) \mathbf{Q}^{-1}\mathbf{W}'$$

and the new control vector $\hat{\mathbf{c}}_t^*$ is related to the original control $\hat{\mathbf{c}}_t$ by:

$$\hat{\mathbf{c}}_t^* = \mathbf{Q}^{-1}\mathbf{W}'\hat{\mathbf{s}}_t + \hat{\mathbf{c}}_t.$$

The associated HJB equation is:

$$\begin{aligned} \rho V(\hat{\mathbf{s}}_t) = & \max_{\hat{\mathbf{c}}_t^*} \left\{ \left(\hat{\mathbf{s}}_t' \bar{\mathbf{R}} \hat{\mathbf{s}}_t + \hat{\mathbf{c}}_t^{*'} \mathbf{Q} \hat{\mathbf{c}}_t^* \right) + V_{\mathbf{s}}(\hat{\mathbf{s}}_t) (\bar{\mu}_{\mathbf{s}}(\mathbf{s}^{ss}, \mathbf{c}^{ss}) \hat{\mathbf{s}}_t + \mu_{\mathbf{c}}(\mathbf{s}^{ss}, \mathbf{c}^{ss}) \hat{\mathbf{c}}_t^*) \right. \\ & \left. + \frac{1}{2} \text{trace}(\Sigma(\mathbf{s}^{ss}) V_{\mathbf{ss}}(\hat{\mathbf{s}}_t)) \right\} \end{aligned}$$

with first order conditions:

$$2\hat{\mathbf{c}}_t^{*'} \mathbf{Q} + V_{\mathbf{s}}(\hat{\mathbf{s}}_t) \mu_{\mathbf{c}}(\mathbf{s}^{ss}, \mathbf{c}^{ss}) = 0$$

which in turn imply:

$$\hat{\mathbf{c}}_t^* = -\frac{1}{2} \mathbf{Q}^{-1} [\mu_{\mathbf{c}}(\mathbf{s}^{ss}, \mathbf{c}^{ss})' V_{\mathbf{s}}(\hat{\mathbf{s}}_t)']$$

i.e., the controls are linear in the costate variables.

Using a *guess-and-verify* method it can be shown that the optimal policy function for the linear-quadratic problem is given by:

$$\mathbf{c}_t = \mathbf{c}^{ss} - \mathbf{Q}^{-1} [\mathbf{W}' + \mu_{\mathbf{c}} (\mathbf{s}^{ss}, \mathbf{c}^{ss})' \Lambda_1] \hat{\mathbf{s}}_t$$

with associated value function:

$$V(\hat{\mathbf{s}}_t) = \Lambda_0 + \hat{\mathbf{s}}_t' \Lambda_1 \hat{\mathbf{s}}_t$$

where Λ_1 is the solution to the continuous-time algebraic Riccati equation:

$$0 = \left(\bar{\mu}_{\mathbf{s}} (\mathbf{s}^{ss}, \mathbf{c}^{ss})' - \frac{\rho}{2} \mathbf{I}_{m+1} \right) \Lambda_1 + \Lambda_1 \left(\bar{\mu}_{\mathbf{s}} (\mathbf{s}^{ss}, \mathbf{c}^{ss}) - \frac{\rho}{2} \mathbf{I}_{m+1} \right) \\ - [\Lambda_1 \mu_{\mathbf{c}} (\mathbf{s}^{ss}, \mathbf{c}^{ss})] \mathbf{Q}^{-1} [\mu_{\mathbf{c}} (\mathbf{s}^{ss}, \mathbf{c}^{ss})' \Lambda_1] + \bar{\mathbf{R}}$$

with \mathbf{I}_{m+1} and $(m+1)$ identity matrix, and

$$\Lambda_0 = \frac{1}{\rho} \text{trace}(\Sigma (\mathbf{s}^{ss}) \Lambda_1).$$

An undesirable property of the linear-quadratic approximation is the imposition of the certainty equivalence property on the optimal policy function for stochastic models. In particular note that \mathbf{c}_t is unaffected by Λ_0 , the only parameter that depends on the stochastic assumptions made about the exogenous forces that drive the economy.

Furthermore, the linear-quadratic approach is subject to a more general criticism which is independent of whether the model is stochastic or not. Judd (1998) and more recently Benigno and Woodford (2012) have discussed the necessary and sufficient conditions for the method to give a correct policy rule. They have shown that the slope coefficients of the approximated policy function are significantly different from the true coefficients whenever the constraints of the optimization problem are non-linear, and thus linearized by means of a first order Taylor expansion around the deterministic steady state. This approach has been called by Benigno and Woodford (2012) the 'Naive' linear-quadratic approximation. The significant differences in the approximated coefficients come from the fact that no information about the curvature of the state processes is included in the approximation of the objective function. In fact, the correct approximation can be obtained by using the procedure suggested in Magill (1977b) and Magill (1977a) where the approximated objective function incorporates that information through the co-state variables resulting from the associated Hamiltonian of the problem. In other words, it would be possible to obtain the correct policy function by approximating the objective function with:

$$\pi(\mathbf{s}_t, \mathbf{c}_t) \approx \begin{bmatrix} \hat{\mathbf{s}}_t' & \hat{\mathbf{c}}_t' \end{bmatrix} \begin{bmatrix} \mathbf{R} & \mathbf{W} \\ \mathbf{W}' & \mathbf{Q} \end{bmatrix} \begin{bmatrix} \hat{\mathbf{s}}_t \\ \hat{\mathbf{c}}_t \end{bmatrix} + \sum_{i=1}^{m+1} \omega_i^{ss} \begin{pmatrix} 0 & \mathbf{0} \\ \mathbf{0} & \mathbf{F} \end{pmatrix}$$

where ω_i^{ss} is the co-state variable associated with state variable i evaluated in the steady state and

$$\mathbf{F} = \begin{pmatrix} \mu_{ss}^i(\mathbf{s}^{ss}, \mathbf{c}^{ss}) & \mu_{sc}^i(\mathbf{s}^{ss}, \mathbf{c}^{ss}) \\ \mu_{cs}^i(\mathbf{s}^{ss}, \mathbf{c}^{ss}) & \mu_{cc}^i(\mathbf{s}^{ss}, \mathbf{c}^{ss}) \end{pmatrix}.$$

Following Kydland and Prescott (1982) it is possible to reformulate the original non-linear stochastic control problem in such a way that the state processes are linear. Depending on the model at hand, this can be done by substituting out the restrictions of the problem into the objective function and using some linear non-stochastic relations to redefine the drift of the state variables. Under this equivalent representation of the problem all the elements of the matrix \mathbf{F} will vanish, i.e., $\mathbf{F} = \mathbf{0}$ and hence the original linear-quadratic approximation can be used and will deliver the correct policy functions. This approach will be used later to compute the approximated policy function of the benchmark model.

2.2 Perturbation method

The perturbation method approximates the true value and policy functions by means of the implicit function theorem and the Taylor's series expansion theorem. The outcome will be a polynomial that approximates the true solution in a neighborhood of a known solution. Following Judd (1998), the perturbation method can be summarized by the following simple steps:

1. Express the problem (2) and (3) as a continuum of problems parameterized by the added perturbation parameter ϵ with the $\epsilon = 0$ case known.
2. Differentiate the continuum of problems with respect to the control variables, \mathbf{c}_t , the state variables, \mathbf{s}_t , and the perturbation parameter, ϵ . Whenever is needed use the envelope condition to simplify the resulting system of equations.
3. Solve for the implicitly defined derivatives at $\mathbf{s}_t = \mathbf{s}^0$ and $\epsilon = 0$, where \mathbf{s}^0 denotes the vector of approximation points. For the benchmark model studied in Section 3, the approximation is done around the deterministic steady state, $\mathbf{s}^0 = \mathbf{s}^{ss}$, and the associated deterministic model, $\epsilon = 0$.
4. Compute the desired order of approximation by means of Taylor's theorem and set $\epsilon = 1$. In general, the order of approximation should be determined by the first non-trivial term or dominant term, that is, apply a Taylor approximation until the first zero term is reached.

Therefore to apply the perturbation method the autonomous diffusion processes in equation (1) is extended with the new parameter ϵ which measures the amount of variance in the model:

$$d\mathbf{s}_t = \mu(\mathbf{s}_t, \mathbf{c}_t) dt + \sqrt{\epsilon} \Sigma^{\frac{1}{2}}(\mathbf{s}_t) d\mathbf{B}_t$$

The perturbed maximized HJB equation and first order conditions are given by:

$$0 = \pi(\mathbf{s}_t, \mathcal{P}(\mathbf{s}_t, \epsilon)) - \rho V(\mathbf{s}_t, \epsilon) + \mu(\mathbf{s}_t, \mathcal{P}(\mathbf{s}_t, \epsilon)) \nabla V(\mathbf{s}_t, \epsilon) + \frac{1}{2} \epsilon \text{trace}(\Sigma(\mathbf{s}_t) \nabla^2 V(\mathbf{s}_t, \epsilon))$$

and

$$\pi_c(\mathbf{s}_t, \mathbf{c}_t) + \mu_c(\mathbf{s}_t, \mathbf{c}_t) \nabla V(\mathbf{s}_t, \epsilon) = 0.$$

As described in Gaspar and Judd (1997) the perturbation technique is based on the computation of Taylor approximations of the unknowns $V(\mathbf{s}_t, \epsilon)$ and $\mathcal{P}(\mathbf{s}_t, \epsilon)$ that take into account not only deviations of the state variables from their steady state values ($\mathbf{s}_t - \mathbf{s}^{ss}$), but also deviations from the deterministic model as measured by ϵ . The second order perturbation, expressed in tensor form, is given by:

$$\begin{aligned} \hat{V}(\mathbf{s}_t, \epsilon) &= V^{ss,0} + V_i^{ss,0} (\mathbf{s}_t - \mathbf{s}^{ss})^i + V_\epsilon^{ss,0} \epsilon + V_{i\epsilon}^{ss,0} (\mathbf{s}_t - \mathbf{s}^{ss})^i \epsilon \\ &\quad + \frac{1}{2} V_{ij}^{ss,0} (\mathbf{s}_t - \mathbf{s}^{ss})^i (\mathbf{s}_t - \mathbf{s}^{ss})^j + \frac{1}{2} V_{\epsilon\epsilon}^{ss,0} \epsilon^2 \\ \hat{\mathcal{P}}(\mathbf{s}_t, \epsilon) &= \mathcal{P}^{ss,0} + \mathcal{P}_i^{ss,0} (\mathbf{s}_t - \mathbf{s}^{ss})^i + \mathcal{P}_\epsilon^{ss,0} \epsilon + \mathcal{P}_{i\epsilon}^{ss,0} (\mathbf{s}_t - \mathbf{s}^{ss})^i \epsilon \\ &\quad + \frac{1}{2} \mathcal{P}_{ij}^{ss,0} (\mathbf{s}_t - \mathbf{s}^{ss})^i (\mathbf{s}_t - \mathbf{s}^{ss})^j + \frac{1}{2} \mathcal{P}_{\epsilon\epsilon}^{ss,0} \epsilon^2 \end{aligned}$$

where

$$\begin{aligned} V^{ss,0} &\equiv V(\mathbf{s}^{ss}, 0) \\ V_i^{ss,0} (\mathbf{s}_t - \mathbf{s}^{ss})^i &\equiv \sum_i \frac{\partial V(\mathbf{s}_t, \epsilon)}{\partial \mathbf{s}_t^i} \Big|_{\mathbf{s}_t = \mathbf{s}^{ss}, \epsilon = 0} (\mathbf{s}_t - \mathbf{s}^{ss})^i \quad \text{for } i = \{1, \dots, m\} \\ V_\epsilon^{ss,0} \epsilon &\equiv \frac{\partial V(\mathbf{s}_t, \epsilon)}{\partial \epsilon} \Big|_{\mathbf{s}_t = \mathbf{s}^{ss}, \epsilon = 0} \epsilon \\ V_{i\epsilon}^{ss,0} (\mathbf{s}_t - \mathbf{s}^{ss})^i \epsilon &\equiv \sum_i \frac{\partial V(\mathbf{s}_t, \epsilon)}{\partial \mathbf{s}_t^i \partial \epsilon} \Big|_{\mathbf{s}_t = \mathbf{s}^{ss}, \epsilon = 0} (\mathbf{s}_t - \mathbf{s}^{ss})^i \epsilon \quad \text{for } i = \{1, \dots, m\} \\ V_{ij}^{ss,0} (\mathbf{s}_t - \mathbf{s}^{ss})^i (\mathbf{s}_t - \mathbf{s}^{ss})^j &\equiv \sum_i \sum_j \frac{\partial^2 V(\mathbf{s}_t, \epsilon)}{\partial \mathbf{s}_t^i \partial \mathbf{s}_t^j} \Big|_{\mathbf{s}_t = \mathbf{s}^{ss}, \epsilon = 0} (\mathbf{s}_t - \mathbf{s}^{ss})^i (\mathbf{s}_t - \mathbf{s}^{ss})^j \quad \text{for } i, j = \{1, \dots, m\} \\ V_{\epsilon\epsilon}^{ss,0} \epsilon^2 &\equiv \frac{\partial^2 V(\mathbf{s}_t, \epsilon)}{\partial \epsilon^2} \Big|_{\mathbf{s}_t = \mathbf{s}^{ss}, \epsilon = 0} \epsilon^2 \end{aligned}$$

and \mathbf{s}_t^i denotes the i -th component of \mathbf{s}_t . For the policy function approximation a similar notation is used for each of its n components.

The constant terms in the approximations are known once the deterministic steady state of the model is computed. Furthermore, as shown in Gaspar and Judd (1997) and Judd (1998) the

terms associated with the first order approximation of the policy function, i.e., $\mathcal{P}_i^{ss,0}$ for all $i \in \mathbf{s}$ and each n , correspond to the solution of a Ricatti equation with l roots, where l is the number of equilibrium paths. Once the stable path is chosen, the first order approximation is completed and the computation of higher order terms in the approximation becomes relatively easy since they correspond to the solution of linear system of equations³.

One of the main differences between perturbation methods for discrete-time and continuous-time frameworks has to do with the certainty equivalence property. It is well known that in discrete-time the first order terms associated with the perturbation parameter are zero for both the policy and the value functions regardless of the properties of the economic model (See Judd (1998), Binsbergen et al. (2008) and Caldara et al. (2012)). This makes linearization and first order perturbation an equivalent procedure. However, as discussed in Gaspar and Judd (1997) and in Judd (1996), the approximation of continuous-time stochastic problems only exhibits the certainty equivalence property if the economic model exhibits the property itself; for example when the utility function is quadratic and/or the diffusion terms are not only control but state independent. Hence, first order perturbation does not coincide with the commonly used linearization technique. This difference is relevant in the macrofinance literature where it is well established that for discrete-time setups it is necessary to compute third order perturbations around the deterministic steady state in order to obtain variation in the risk premium, whereas a second order perturbation delivers a constant risk premium and a first order perturbation will eliminate it. However, in continuous-time it is only necessary to compute a second order perturbation in order to get a time-varying risk premium, while the first order approximation produces a constant premium.

2.3 Projection methods

The projection method is a widely used technique in applied mathematics and numerical analysis. Two main approaches have been used in the economic literature to solve functional problems: finite element methods and spectral methods. In this paper, I only apply the later since in principle their accuracy is higher and given its structure it is possible to handle a higher number of state variables (Aruoba et al., 2006).

Let $\mathcal{Q} = \{\phi_j\}_{j=0}^{\infty}$ be a given family of polynomials. Then the unknown real-valued function $f : \mathbf{s} \rightarrow \mathbf{c}$ where $\mathbf{s} \subset \mathbb{R}^m$ and $\mathbf{c} \subset \mathbb{R}^n$ can be approximated by a finite linear combination of the first $k + 1$ members of this family. This linear combination is called the approximant \hat{f} of f , and the elements of \mathcal{Q} basis functions:

$$\hat{f}(\mathbf{s}) = \sum_{j=0}^k a_j \phi_j(\mathbf{s}), \quad \mathbf{s} \in \mathbf{S} \subset \mathbb{R}^m$$

³As mentioned in Judd (1998), the stable path corresponds to the root of the Ricatti equation that ensures concavity of the value function at the deterministic steady state. In other words, pick the $\mathcal{P}_i^{ss,0}$ compatible with $V_{ii}^{ss,0} < 0$. If all the roots fulfill this condition then it is not possible to proceed due to the existence of indeterminacies in the economic model.

where $\phi_j(\mathbf{s}) = \phi_j(s^1) \phi_j(s^2) \dots \phi_j(s^m)$ and $k = \prod_{i=1}^m k_i$ with $k_i + 1$ the order of approximation in each dimension of \mathbb{R}^m . In addition let's suppose that the function f is implicitly defined by the functional equation $\mathcal{N}(f) = 0$, where $\mathcal{N} : B_1 \rightarrow B_2$ is a continuous map with B_1 and B_2 spaces of functions.

The simplest form of the projection method applied to the stochastic control problem is the case where we solve for the problem's value function by defining the functional equation:

$$\mathcal{N}_1(V(\mathbf{s}_t)) := \pi(\mathbf{s}_t, \mathbf{c}_t) - \rho V(\mathbf{s}_t) + \mu(\mathbf{s}_t, \mathbf{c}_t) \nabla V(\mathbf{s}_t) + \frac{1}{2} \text{trace}(\Sigma(\mathbf{s}_t) \nabla^2 V(\mathbf{s}_t)) = 0$$

where \mathbf{c}_t , which is a function of the state variables and the derivatives of the unknown value function, is computed from the first order conditions of the problem. The derivatives of the value function can be easily computed once the approximant, $\hat{V}(\mathbf{s}_t)$, is defined. The objective is to approximate the ∞ -dimensional problem associated with the HJB equation with a finite-dimensional problem by using the following parametric approximation of the unknown value function:

$$\hat{V}(\mathbf{a}, \mathbf{s}_t) = \sum_{j=0}^k a_j \phi_j(\mathbf{s}_t) = \Phi(\mathbf{s}_t) \mathbf{a}, \quad \mathbf{s}_t \in \mathbf{S} \subset \mathbb{R}_+^m$$

with $\mathbf{a} = (a_0, \dots, a_k)$ and $\Phi(\mathbf{s}_t) = \Phi_m \otimes \Phi_{m-1} \otimes \dots \otimes \Phi_1$, where

$$\Phi_k = [\Phi_{ij}]_k = [\phi_j(s_t^i)] = \begin{bmatrix} \phi_1(s_t^{i1}) & \phi_2(s_t^{i1}) & \dots & \phi_k(s_t^{i1}) \\ \phi_1(s_t^{i2}) & \phi_2(s_t^{i2}) & \dots & \phi_k(s_t^{i2}) \\ \vdots & & \ddots & \vdots \\ \phi_1(s_t^{ik}) & \phi_2(s_t^{ik}) & \dots & \phi_k(s_t^{ik}) \end{bmatrix}$$

for all $i = 1, \dots, m$ and $j = 1, \dots, k$. Thus, for each $\mathbf{s}_t \in \mathbf{S} \subset \mathbb{R}_+^m$ the associated maximized HJB equation becomes:

$$0 = \pi(\mathbf{s}_t, \mathbf{c}_t) - \rho \Phi(\mathbf{s}_t) \mathbf{a} + \mu(\mathbf{s}_t, \mathbf{c}_t) \nabla \Phi(\mathbf{s}_t) \mathbf{a} + \frac{1}{2} \text{trace}(\Sigma(\mathbf{s}_t) \nabla^2 \Phi(\mathbf{s}_t) \mathbf{a})$$

with first order conditions:

$$\pi_c(\mathbf{s}_t, \mathbf{c}_t) + \mu_c(\mathbf{s}_t, \mathbf{c}_t) \nabla \Phi(\mathbf{s}_t) \mathbf{a} = 0$$

for each $c_t \in \mathbf{c}_t$, where $\nabla \Phi(\mathbf{s}_t)$ and $\nabla^2 \Phi(\mathbf{s}_t)$ denote the gradient and hessian matrix of $\Phi(\mathbf{s}_t)$ respectively. Note that the transformed problem is now finite-dimensional. Instead of looking on the function of spaces for $V(\cdot)$, the task is to compute the $(k+1)$ basis coefficients $\{a_j\}_{j=0}^k$ such that the residual function, $R(\mathbf{a}; \mathbf{s}_t) := \mathcal{N}_1(\hat{V}(\mathbf{a}, \mathbf{s}_t))$, is as close to zero as possible.

Following Heer and Maussner (2009), I assume there is a set of test functions $\{g_j(\mathbf{s}_t)\}_{j=0}^k$ and a weighting function $w(\mathbf{s}_t) \geq 0$ which combined with the residual function define an inner product of the form:

$$\langle R(\mathbf{a}; \mathbf{s}_t), g(\mathbf{s}_t) \rangle \equiv \int_{\mathbf{S}} w(\mathbf{s}_t) R(\mathbf{a}; \mathbf{s}_t) g_j(\mathbf{s}_t) d\mathbf{s}_t.$$

Since this inner product induces a norm on the function space, it is possible to choose the set of parameters (a_0, \dots, a_k) such that

$$\int_{\mathbf{S}} w(\mathbf{s}_t) R(\mathbf{a}; \mathbf{s}_t) g_j(\mathbf{s}_t) d\mathbf{s}_t = 0, \quad \forall j = 0, \dots, k.$$

Different definitions of the test functions and the weighting function give rise to different projection techniques. In particular I am interested in two particular cases: the least squares projection method which uses $g_j^{LS}(\mathbf{s}_t) \equiv \partial R(\mathbf{a}; \mathbf{s}_t) / \partial a_j$ and $w^{LS}(\mathbf{s}_t) \equiv 1$, and the collocation method which uses $g_j^{COL}(\mathbf{s}_t) \equiv 1$ and for the weighting function the Dirac delta function $w^{COL}(\mathbf{s}) \equiv \delta(\mathbf{s}_t - \mathbf{s}_t^j)$, where \mathbf{s}_t^j is the j -th element of the state-vector usually called node. The weighting function $w^{COL}(\mathbf{s}_t)$ implies that $R(\mathbf{a}; \mathbf{s}_t^j) = 0$ for $j = 1, \dots, k$, where the n nodes, $\mathbf{s}_t^1, \dots, \mathbf{s}_t^k$, are called collocation points. In other words, the first method searches for the vector of parameters that minimizes the sum of squared residuals along all the nodes, \mathbf{a}^{LS} , while the second one searches for the vector of parameters that make the residual function zero at all collocation points, \mathbf{a}^{COL} . The implementation relies on a numerical optimization algorithm to find \mathbf{a}^{LS} , and on a root-finding algorithm to find \mathbf{a}^{COL} .

When the problem at hand is highly non-linear it is useful to make a simple extension of the method just described. In order to give more flexibility to the approximation and increase the efficiency of the algorithms it is recommendable to apply the projection technique to more than one functional equation. In particular, it is optimal to approximate not only the HJB equation but also the policy functions resulting from the non-linear first order conditions, that is:

$$\begin{bmatrix} \mathcal{N}_1(V(\mathbf{s}_t)) \\ \mathcal{N}_2(\mathcal{P}(\mathbf{s}_t)) \end{bmatrix} = \mathbf{0}$$

where the second set of equations comes from the first order conditions of the stochastic control problem:

$$\mathcal{N}_2(\mathcal{P}(\mathbf{s}_t)) := \pi_c(\mathbf{s}_t, \mathcal{P}(\mathbf{s}_t)) + \mu_c(\mathbf{s}_t, \mathcal{P}(\mathbf{s}_t)) \nabla V(\mathbf{s}_t) = 0.$$

Under this extension the approximants are given by:

$$\begin{aligned} \hat{V}(\mathbf{a}, \mathbf{s}_t) &= \sum_{j=0}^{k_a} a_j \phi_{a,j}(\mathbf{s}_t) = \Phi_a(\mathbf{s}_t) \mathbf{a}, \quad \mathbf{s}_t \in \mathbf{S} \subset \mathbb{R}_+^m \\ \hat{\mathcal{P}}(\mathbf{b}, \mathbf{s}_t) &= \sum_{j=0}^{k_b} b_j \phi_{b,j}(\mathbf{s}_t) = \Phi_b(\mathbf{s}_t) \mathbf{b}, \quad \mathbf{s}_t \in \mathbf{S} \subset \mathbb{R}_+^m \end{aligned}$$

which allow for more flexibility in the approximation through different basis function and basis points for each of the functional equations. The problem now becomes to find two different set of parameters \mathbf{a} and \mathbf{b} that make the residual function $R(\mathbf{a}, \mathbf{b}; \mathbf{s}_t) := \left(\mathcal{N}_1 \left(\hat{V}(\mathbf{a}, \mathbf{s}_t) \right), \mathcal{N}_2 \left(\hat{\mathcal{P}}(\mathbf{b}, \mathbf{s}_t) \right) \right)^\top$ as close to zero as possible.

Regarding the selection of the basis functions, I choose the set of orthogonal polynomials generated by the Chebyshev function. The nodal points of the state vector correspond to the zeros of the Chebyshev polynomials. The combination of Chebyshev basis and nodes yields an extremely well behaved projection equation that can be solved accurately and efficiently even for high degrees of approximation (Judd, 1998). For this choice, the state variables in $\mathbf{S} \subset \mathbb{R}_+^m$ must be transformed from their original domain since the Chebyshev polynomials are only defined in $[-1, 1]$. This is achieved by using the map $\Gamma \left(s_t^j \right) = \left(2s_t^j - a - b \right) / (b - a)$, where $s_t^j \in [a, b]$ and a, b are the predefined bounds of the j -th state variable.

3 Benchmark model

I consider as a test case a continuous-time version of the RBC model described in Aruoba et al. (2006). The one-good in this economy is produced according to a constant return to scale technology:

$$Y_t = A_t K_t^\alpha L_t^{1-\alpha}, \quad \alpha \in (0, 1) \quad (5)$$

where A_t denotes the total factor productivity (TFP), K_t the aggregate capital stock and L_t the fraction of hours worked. I assume that the TFP is driven by a Cox-Ingersoll-Ross (CIR) mean reverting stochastic process of the form:

$$dA_t = \kappa(\omega - A_t)dt + \eta\sqrt{A_t}dB_{A,t}, \quad \kappa, \omega > 0 \text{ and } A_0 > 0 \text{ given} \quad (6)$$

where $B_{A,t}$ is a standard Brownian motion and η denotes the volatility of the TFP. This process ensures that the random variable A_t only takes positive values. In fact, if $A_t \rightarrow 0$, then the drift term $\kappa\omega > 0$ and the variance $\eta^2 A_t \rightarrow 0$, implying that the process will remain positive w.p.1. Furthermore, adding the restriction $\eta^2 < 2\kappa\omega$ ensures that the process never reaches zero: in this case, the drift will be very high relative to the variance, shooting up the process even before it reaches zero.

On the other hand, the aggregate capital stock increases if the gross investment I_t exceeds the capital depreciation, δK_t , where δ is the gross depreciation rate. However, following Wälde (2011) it is assumed that the rate at which the capital stock depreciates is stochastic. That is, the accumulation of capital is a risky activity. The intuition behind this assumption is similar to that in Furlanetto and Seneca (2013). In what follows, I assume that the aggregate capital stock follows the stochastic differential equation:

$$dK_t = (I_t - \delta K_t) dt + \sigma K_t dB_{K,t}, \quad K_0 > 0 \text{ given}, \quad (7)$$

where $B_{K,t}$ is a standard Brownian and σ denotes the volatility of the stochastic depreciation rate. As mentioned in Wälde (2011) introducing uncertainty in the production factors instead of doing so directly in the resource constraint avoids the problems faced by the common differential representation of technologies introduced in Merton (1969) and Merton (1971).

The firms producing the only good in the economy are owned by the households and assumed to operate in a competitive environment. Hence, their demands for capital and labor are given respectively by:

$$r_t = \alpha A_t K_t^{\alpha-1} L_t^{1-\alpha} \quad (8)$$

$$w_t = (1 - \alpha) A_t K_t^\alpha L_t^{-\alpha} \quad (9)$$

where r_t is the rental rate of capital and w_t is the real wage.

Following Turnovsky and Smith (2006) and Posch (2011), the representative household maximizes the expected discounted life-time utility derived from consumption, C_t , and leisure, $(1 - L_t)$:

$$\max_{\{C_t, L_t\}_{t=0}^{\infty}} \mathbb{E}_0 \int_0^{\infty} e^{-\rho t} \frac{(C_t (1 - L_t)^\psi)^{1-\gamma}}{1 - \gamma} dt, \quad \psi \geq 0, \gamma > 0$$

subject to the intertemporal budget constraint

$$dK_t = ((r_t - \delta) K_t + w_t L_t - C_t) dt + \sigma K_t dB_{K,t} \quad (10)$$

where $\rho > 0$ is the rate of time preference, ψ measures the preference for leisure and γ denotes the inverse of the elasticity of intertemporal substitution (EIS). To ensure concavity of the utility function the consumption-leisure measure of relative risk aversion has to be greater or equal to zero, i.e., $\gamma - (1 - \gamma) \psi \geq 0$ (Swanson, 2012).

Under the assumption of competitive markets both welfare theorems hold. Thus, it is possible to solve the social planner's problem where the benevolent planner maximizes the utility of the representative household subject to the production function (Equation 5), the evolution of the TFP (Equation 6), the evolution of the capital stock (Equation 7), the market clearing condition $Y_t = C_t + I_t$ and some initial values $K_0 > 0$ and $A_0 > 0$.

The Hamilton-Jacobi-Bellman (HJB) equation for the planner's problem is given by:

$$\rho V(K_t, A_t) = \max_{\{C_t, L_t\} \in \mathbb{R}_+^2} \left\{ \frac{(C_t (1 - L_t)^\psi)^{1-\gamma}}{1 - \gamma} + (A_t K_t^\alpha L_t^{1-\alpha} - \delta K_t - C_t) V_K(K_t, A_t) \right. \\ \left. + \kappa (\omega - A_t) V_A(K_t, A_t) + \frac{1}{2} \sigma^2 K_t^2 V_{KK}(K_t, A_t) + \frac{1}{2} \eta^2 A_t V_{AA}(K_t, A_t) \right\} \quad (11)$$

for any $t \in [0, \infty)$, where $V(K_t, A_t)$ is the value function and denotes the value at instant t of the planner's expected utility along the optimal program. The first order conditions for any interior solution are:

$$\frac{\left(C_t(1-L_t)^\psi\right)^{1-\gamma}}{C_t} = V_K(K_t, A_t) \quad (12)$$

$$\psi \frac{\left(C_t(1-L_t)^\psi\right)^{1-\gamma}}{(1-L_t)} = (1-\alpha)A_t K_t^\alpha L_t^{-\alpha} V_K(K_t, A_t) \quad (13)$$

making optimal consumption and the optimal fraction of hours worked implicit functions of the state variables, i.e., $C_t = C(K_t, A_t)$ and $L_t = L(K_t, A_t)$. The solution to the planner's problem is fully characterized by the maximized (concentrated) HJB equation:

$$\rho V(K_t, A_t) = \frac{\left(C(K_t, A_t)(1-L(K_t, A_t))^\psi\right)^{1-\gamma}}{1-\gamma} + \left(A_t K_t^\alpha L(K_t, A_t)^{1-\alpha} - \delta K_t - C(K_t, A_t)\right) V_K(K_t, A_t) + \kappa(\omega - A_t) V_A(K_t, A_t) + \frac{1}{2}\sigma^2 K_t^2 V_{KK}(K_t, A_t) + \frac{1}{2}\eta^2 A_t V_{AA}(K_t, A_t) \quad (14)$$

which corresponds to a functional equation in the unknown value and policy functions. Solving for the equilibrium of this economy amounts to find in the space of functions for $V(K_t, A_t)$, $C(K_t, A_t)$ and $L(K_t, A_t)$ such that given a random path for the exogenous process, $\{A_t\}_{t=0}^\infty$, and an initial condition $K_0 > 0$:

1. The planner solves his problem. That is, Equation (14) is satisfied at every instant of time,
2. The accumulation constraints (6) and (7) are satisfied at every instant of time, and
3. The goods market clears, $Y_t = C_t + I_t$.

In general, the planner's problem does not have a closed form solution. However, under some parametric restrictions it is possible to derive analytical forms for the value and policy functions. The parametrizations under which it is possible to derive a closed form solution provide a benchmark against to which is possible to assess the accuracy of the numerical approximations⁴.

Proposition 3.1 (Linear consumption rule). *Suppose the household does not exhibit disutility from labor, i.e., $\psi = 0$, and the output elasticity of capital equals the degree of risk aversion, $\alpha = \gamma$. Then the value function is given by:*

$$V(K_t, A_t) = \Theta^{-\gamma} \left(\frac{K_t^{1-\gamma}}{1-\gamma} + \frac{\rho A_t + \kappa \omega}{\rho(\rho + \kappa)} \right) \quad (15)$$

⁴Posch (2011) finds a closed form solution for the case of endogenous labor supply and non-stationary TFP. Assuming that A_t evolves according to a geometric Brownian motion and that the capital accumulation is subject to jumps, a similar solution can be derived.

and the optimal policy functions are:

$$C(K_t, A_t) = \Theta K_t, \quad L(K_t, A_t) = 1, \quad (16)$$

where

$$\Theta = \left[\frac{\rho}{\gamma} + \frac{1-\gamma}{\gamma} \left(\delta + \frac{1}{2} \sigma^2 \gamma \right) \right]$$

Proof. See Posch (2009). □

Proposition 3.2 (Constant savings function). *Suppose that total factor productivity is constant, i.e., $A_t = \bar{A}$. Let $\psi > 0$, $\gamma \geq 1$ and $\rho = \bar{\rho}$ with*

$$\bar{\rho} = -(1 - \alpha\gamma) \left(\delta + \frac{1}{2} \alpha\gamma\sigma^2 \right) > 0$$

then, the value function is given by:

$$V(K_t, \bar{A}) = \Gamma \frac{K_t^{1-\alpha\gamma}}{1 - \alpha\gamma} \quad (17)$$

where

$$\Gamma = \left(- \frac{(1-s)^{(1-\gamma)} \bar{A}^{(1-\gamma)} L^{(1-\alpha)(1-\gamma)} (1-L)^{\psi(1-\gamma)}}{(1-\gamma) \bar{A} L^{(1-\alpha)} - (1-\gamma)(1-s) \bar{A} L^{(1-\alpha)}} \right)$$

and

$$1-s \equiv \frac{(1-\alpha)(1-L)}{\psi L}$$

denotes the constant propensity to consume out of income. On the other hand, the optimal consumption will be a constant fraction of income:

$$C(K_t, \bar{A}) = (1-s) \bar{A} K_t^\alpha L^{1-\alpha} \quad (18)$$

and the optimal fraction of hours supplied will be constant:

$$L(K_t, \bar{A}) = \frac{\gamma(1-\alpha)}{\gamma(1-\alpha) - \psi(1-\gamma)} \in [0, 1]. \quad (19)$$

Proof. See Posch (2011). □

Using equations (6), (7), (8), (9), (12), (13) together with the first order derivatives of equation (14) with respect to the state variables it is possible to derive the non-stochastic steady state of the planner's problem. This is given by the quantities: $r^{ss} = \rho + \delta$, $A^{ss} = \omega$, $L^{ss} = (1-\alpha) / \left(1 + \psi - \alpha - \frac{\alpha\psi\delta}{\rho+\delta} \right)$, $K^{ss} = (\alpha A^{ss} / (\rho + \delta))^{\frac{1}{1-\alpha}} L^{ss}$, $w^{ss} = (1-\alpha) A^{ss} (K^{ss})^\alpha (L^{ss})^{-\alpha}$, $C^{ss} =$

$(r^{ss} - \delta) K^{ss} + w^{ss} L^{ss}$, $V^{ss} = \left(C^{ss} (1 - L^{ss})^\psi \right)^{1-\gamma} / \rho (1 - \gamma)$ which will be used for the implementation of the perturbation method and the linear-quadratic approximation.

The model just outlined has an equivalent representation with investment and leisure as control variables. Under this alternative setup the drift of stochastic differential equation describing the capital stock is linear in the control and state variables. This formulation will be used later to compute an alternative linear-quadratic approximation of the policy function of consumption with correct coefficients as mentioned in Subsection 2.1. This representation can be found in A.

4 Numerical results

4.1 Calibration

In order to evaluate the performance of the different numerical methods I carry out two type of exercises⁵. In the first one I use the analytical solution derived under Proposition 3.2 to assess the methods by comparing deviations of the numerical approximations from the true solution. In a second exercise I use the model without any parametric restriction and rely on a different set of measures to evaluate the relative performance of the numerical procedures.

Both exercises are based on the benchmark parametrization presented in Table 1. For the case where no closed form solution is available I use the calibration for the U.S. economy described in Aruoba et al. (2006). The rate of time preference is set to $\rho = 0.0105$ to match a 4% annual interest rate. The risk aversion parameter γ is fixed to 2.0, a value widely used in the literature. The share of labor in aggregate output and the depreciation rate are set to $\alpha = 0.4$ and $\delta = 0.0196$ respectively, while the leisure preference parameter is fixed to $\psi = 1.8011$ to match a labor supply of 31% in the steady state. Regarding the stochastic components of the model, I set the volatility of the capital stock accumulation equal to zero and choose κ , ω and η to match the properties of the Solow residual in the U.S. economy. For the constant savings function of proposition 3.2 the risk aversion parameter is increased to 3.85 in order to obtain a “knife-edge” value of the rate of time preference, $\bar{\rho}$, close to 0.0105, while the TFP is set to a constant $\bar{A} = 1$ and the standard deviation of the stochastic depreciation rate is equal to 0.001 to allow for some stochastics in the model.

Parameter	ρ	ψ	γ	α	δ	σ	κ	ω	η	\bar{A}
Analytical Sol.	$\bar{\rho}$	1.8011	3.85	0.4	0.0196	0.001	N.A.	N.A.	N.A.	1.00
No Analytical Sol.	0.0105	1.8011	2.00	0.4	0.0196	0.000	0.05	1.00	0.007	N.A.

Table 1: Benchmark calibration

In Section 4.3 the second exercise is repeated using alternative calibrations to check the robustness of the solution methods to different degrees of non-linearities of model arising from the

⁵Regarding the implementation of the algorithms, global methods and the linear-quadratic approximation are coded in Matlab[®] 7.11.0. For projection methods I complement my codes with the CompEcon Toolbox developed by Miranda and Fackler (2002). On the other hand, perturbations are coded in Mathematica[®] 7.0.

risk aversion parameter and the volatility of the disturbances. Table 2 summarizes the alternative scenarios.

Case	$\eta = 0.007$		$\eta = 0.035$	
	$\sigma = 0.000$	$\sigma = 0.001$	$\sigma = 0.000$	$\sigma = 0.001$
$\gamma = 0.65$	M1	M3	M6	M8
$\gamma = 2$	Benchmark	M4	M7	M9
$\gamma = 10$	M2	M5	Extreme I	Extreme II

Table 2: Sensitivity analysis. The value $\gamma = 0.65$ was chosen in such a way that the concavity condition for the utility function was fulfilled given the calibrated value of ψ .

4.2 Value function, policy functions and HJB equation residuals

Figure 1 plots the approximated value and policy functions and their corresponding true values derived under Proposition 3.2. From the analytical solution it is evident that $V(K_t, \bar{A}) \rightarrow -\infty$ as $K_t \rightarrow 0$, i.e., the value function exhibits a very steep slope for low values of the capital stock. This result suggests to include values of the capital stock away from the deterministic steady state in order to compare the performance of the approximation methods around highly non-linear regions of the state space. The unknown functions are therefore approximated over the interval $[0.5K^{ss}, 1.5K^{ss}]$ using 1000 equally spaced grid points. For projection methods I first compute a Chebyshev approximation with 17 Chebyshev basis functions using the true values as a starting point. Given the width of the state-space a fewer number of nodes results in non-accurate and non-smooth policy functions. Once the coefficients of the approximation are found, I use linear interpolation on a set of 1000 equally spaced points. Finally, the tolerance level for the iterative algorithms is fixed at $\mathcal{E} = 1.4901e^{-8}$.

As expected from theory, global approximations (projection methods) outperform the local approximations (perturbation method and linear-quadratic approximations) of both the value function and policy functions for levels of the capital stock away from the steady state. However, it is remarkable the improvement obtained by moving from a first to a second order perturbation. In particular, it is worth emphasizing the correction of about 100% in the curvature of the value function at the lower bound of the grid. This encourages the computation of higher order perturbations which as shown later come at no significant extra cost in terms of implementation.

Regarding the accuracy of the linear-quadratic approximation two points are worth to mention. First, when the 'Naive' linear-quadratic approximation is used the slopes of the policy functions are miscalculated as discussed in Subsection 2.1. This can be easily identified in the middle panel of Figure 1 under the label *LQ:consumption*. However, if the equivalent representation of the problem, as described in A, is used the slopes of the approximated functions are correctly computed. Since the restrictions of the optimization problem are linear no information is lost in the approximation step of the objective function. The resulting policy functions can be found under the label *LQ:investment*. Second, the approximated value function is not dramatically affected by the model's representation

being used. Nevertheless some improvement can be obtained by using linear instead of linearized constraints.

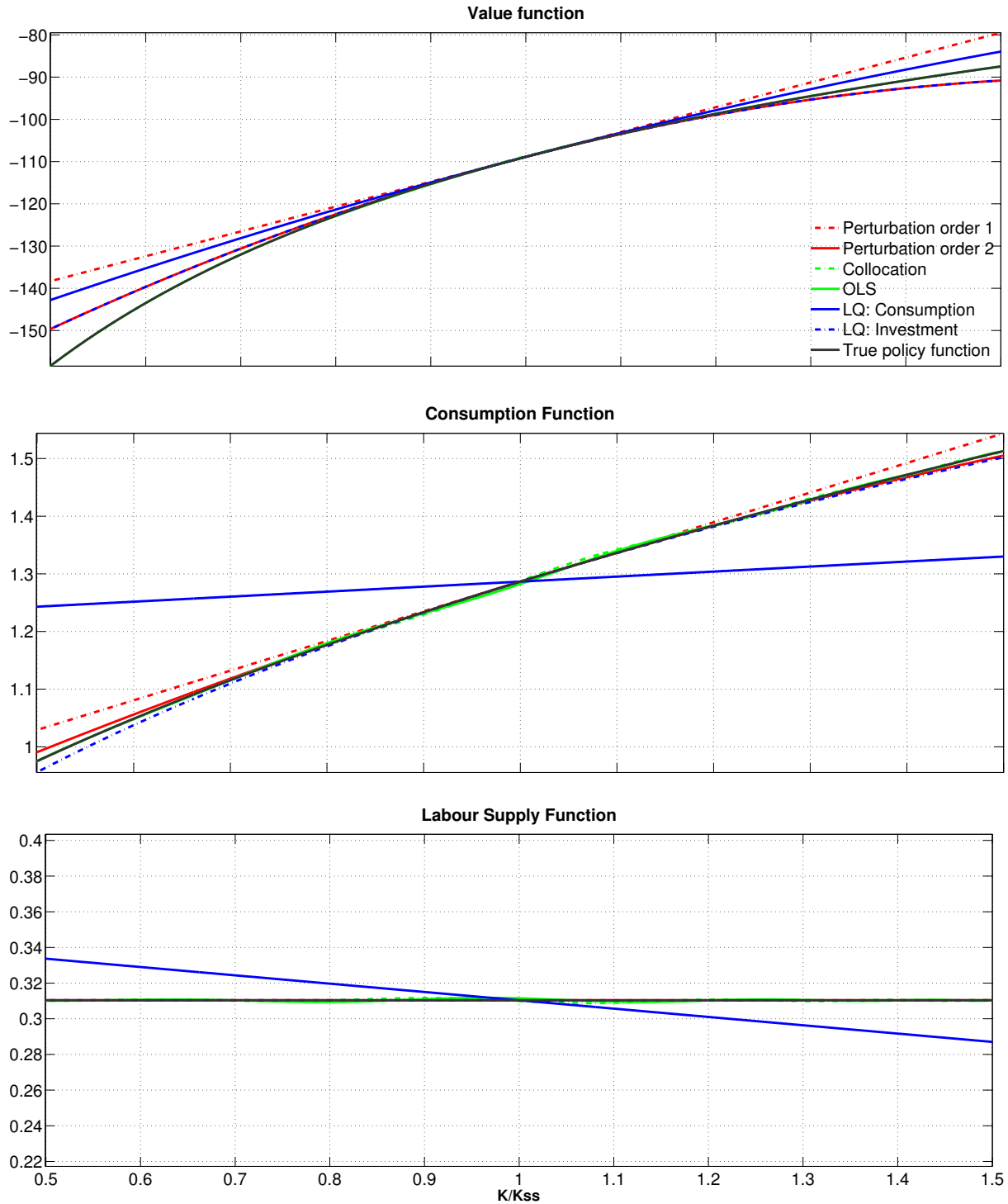


Figure 1: Approximated vs. true value and policy functions along $[0.5K^{ss}, 1.5K^{ss}]$

To complement the previous findings I compute the following measures introduced in Collard

and Juillard (2001) designed to evaluate the quality of the approximations:

$$E_1^l = 100 \times \frac{1}{n_K} \sum_{i=1}^{n_K} \left| \frac{f(K_i, \bar{A}) - \hat{f}^l(K_i, \bar{A})}{f(K_i, \bar{A})} \right| \quad \text{and} \quad E_\infty^l = 100 \times \max_i \left\{ \left| \frac{f(K_i, \bar{A}) - \hat{f}^l(K_i, \bar{A})}{f(K_i, \bar{A})} \right| \right\}$$

where $f(K_i, \bar{A})$ denotes the true solution of either the value or the policy functions evaluated at grid point i , $\hat{f}^l(K_i, \bar{A})$ is the approximation obtained by method l evaluated at grid point i , and n_K is the number of grid points in the state-space. The statistic E_1^l measures the average relative error of using the approximation instead the true solution, while E_∞^l measures the maximum relative error. Sometimes, the latter is preferred since it bounds the mistake made by using the approximation instead of the true function (Aruoba et al., 2006). The results are shown in Table 3.

Overall, the results suggest the use of projection methods over the alternatives for the solution of continuous-time DSGE models: a maximum error of 0.16% when collocation is used for the case of the value function. However, as it will be shown later, a greater level of accuracy requires considerable more computing time relative to perturbation or linear-quadratic methods which might relegate it from an econometric perspective where the solution of the model is needed at each iteration of the estimation procedure, e.g., the maximum likelihood estimator. On the other hand, one alternative to improve the accuracy of perturbation for any given degree of approximation is to use rational functions. This parametric method, known as Padé approximations, has been proven to perform better away from the steady state. Judd and Guu (1993) discuss its application in the context of a one state - one control variable continuous-time stochastic growth model.

Method	Value function		Consumption		Labor supply	
	E_1	E_∞	E_1	E_∞	E_1	E_∞
Perturbation 1	11.2008	57.1436	6.8988	60.7601	0.0000	0.0001
Perturbation 2	6.6285	47.4301	3.1341	36.3450	0.0000	0.0000
Collocation	0.0275	0.1592	0.0948	0.4635	0.1201	0.5224
Least Squares	0.0229	0.0904	0.0962	0.3863	0.1186	0.4430
Linear-Quadratic (cons.)	9.1221	53.3116	24.2086	135.8970	5.6952	13.5280
Linear-Quadratic (inv.)	6.6287	47.4303	2.4204	21.3193	0.0003	0.0004

Table 3: Accuracy check for benchmark model under Proposition 3.2 (%).

Figure 2 plots the log10 magnitude of the relative numerical error obtained from the approximation of the value function along all the state-space. A value of -6 indicates that for every million of units of welfare, the agent makes an error of 1 unit by using the approximation instead of the true value function. Values close to zero are not desirable. The plot depicts the global nature of projection methods as well as the local nature of perturbation and linear-quadratic methods where the approximation deteriorates as the capital stock moves away from the steady state. However, the use of a second order approximation significantly reduces this error.

Next I consider the general case where no closed form solution is available. Under this scenario the state-space becomes \mathbb{R}_+^2 . The approximations are computed over $[0.5K^{ss}, 1.5K^{ss}] \times [\bar{A}, \underline{A}]$ where

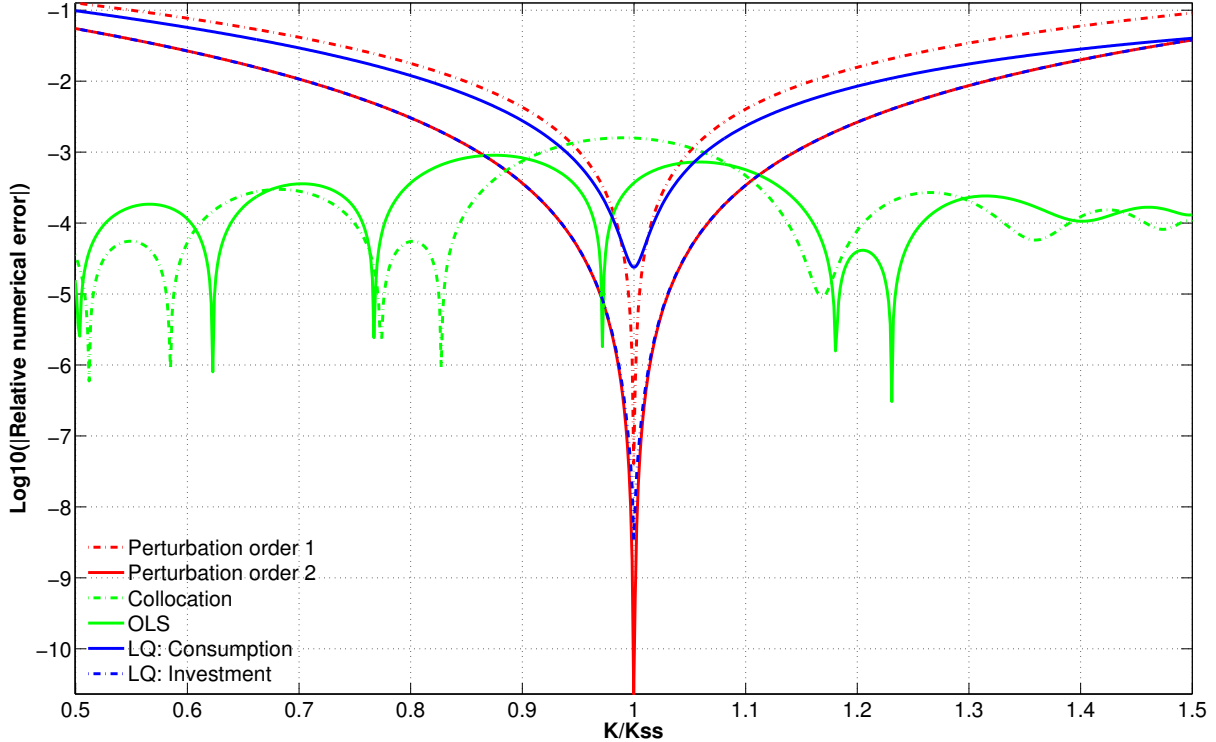


Figure 2: Numerical error for benchmark model under Proposition 3.2

the grid along the TFP lattice is designed following Heer and Maussner (2009): I set $(\underline{A} - \bar{A})$ equal to a multiple of the limiting standard deviation of Equation (6) which is given by⁶:

$$\sigma_A = \sqrt{\omega \left(\frac{\eta^2}{2\kappa} \right)}.$$

In particular, the grid for the TFP is defined as $[A^{ss} - 5\sigma_A, A^{ss} + 5\sigma_A]$ which for the AR(1) discrete-time representation of the Ornstein-Uhlenbeck process it has been shown to be a reasonably good approximation. For projection methods I use 9 Chebyshev basis functions. Regarding the initial values for the global methods I use the following rule-of-thumb: First obtain a local approximation by means of a first order perturbation. Use this as an initial guess for the least squares approximation. Once a reasonable approximation has been found (usually before 100.000 iterations of the *trust-region-dogleg* numerical procedure) use it as an initial guess for the collocation method. When the later converges (usually after just a few iterations using the *trust-region-dogleg* numerical procedure) use it as a starting value for the least squares algorithm again to check for robustness. I use the least squares residual function first since it is less demanding in terms of the projection criterion. While least squares seeks to minimize the residual function, collocation calls for exact zeros in each of the approximating nodes.

Figure 3 plots the value and policy functions for each of the numerical methods using 301 grid

⁶See Cox et al. (1985) for a derivation of this formula.

points in each lattice and fixing the TFP is at its steady state value. Plots for different values of the TFP can be easily derived and are available upon request. A similar pattern to that described for the closed form solution is obtained. All the approximations are almost indistinguishable around the steady state value of the capital stock. For linear-quadratic approximations and perturbation some deviations are found when moving away from that point. However, a correction of almost 100% in the boundaries of the later can be obtained for the approximation of value function when we move from a first to a second order perturbation.

The statistics E_1 and E_∞ defined previously cannot be used anymore to evaluate the performance of the different approximation methods since no closed form solution is available⁷. Therefore I follow Judd and Guu (1993) and compute the unit free vector of HJB equation residuals for each of the procedures. A similar measure, usually called Euler equation errors, is used in discrete-time analysis of DSGE models⁸. The idea is to check how much $\mathcal{N}_1 \left(\hat{V}^l(K_i, A_j) \right)$ for $i = 1, \dots, n_K$, $j = 1, \dots, n_A$, as a fraction of the discounted steady state value function, differs from the zero function⁹. I define the HJB equation residuals as:

$$R_{HJB}^l(K_i, A_j) \equiv \left(\frac{\mathcal{N}_1 \left(\hat{V}^l(K_i, A_j) \right)}{\rho \hat{V}^l(K^{ss}, A^{ss})} \right)$$

for all i, j in the state-space and approximation method l ¹⁰. This measure can be interpreted as the relative optimization error incurred by the use of the approximated value and policy functions. It is an optimization error since the residuals are computed from the maximized HJB equation. The results are summarized in Table 4 where I have computed a similar set of statistics to those presented for the closed form solution case, i.e.:

$$\tilde{E}_1^l = \frac{1}{n_A n_K} \sum_{j=1}^{n_A} \sum_{i=1}^{n_K} R_{HJB}^l(K_i, A_j) \text{ and } \tilde{E}_\infty^l = \max_{i,j} \left\{ R_{HJB}^l(K_i, A_j) \right\}$$

where \tilde{E}_1^l is the average HJB equation residual and \tilde{E}_∞^l is the maximum HJB equation residual. As before, the maximum HJB equation residual bounds the error made by using a particular approximation method.

Once again global methods outperform local methods. The rule-of-thumb described previously

⁷In the discrete-time literature it is common to use as “true solution” the approximation obtained by Value Function iteration due to its convergence properties.

⁸Alternative accuracy checks like the one proposed in Den Haan and Marcet (1994), or the simulation based assessments used in Taylor and Uhlig (1990) and Aruoba et al. (2006) are not used in this paper due to their discrete-time nature. Their implementation require a discretization procedure for the partial/stochastic differential equations that describe the model adding more errors to those already obtained in the approximation stage of the value and policy functions.

⁹For projection methods this measure is computed for the interpolated value and policy functions and not for the functions at the approximation nodes.

¹⁰To prevent for under- and overflow problems, the measure is corrected by computing its base 10 logarithm and adding a small number respectively.

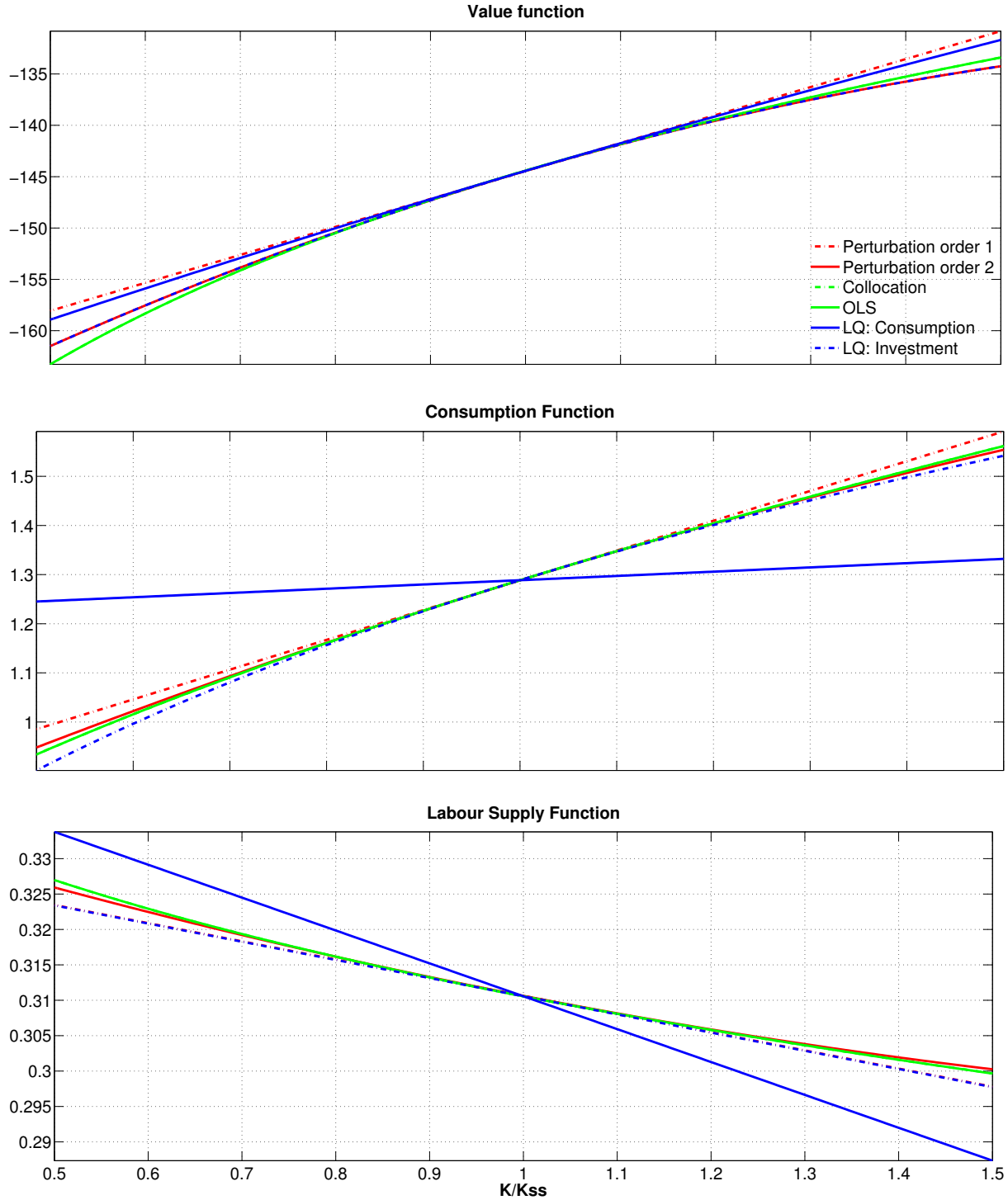


Figure 3: Computed value and policy functions for benchmark calibration at $A_t = A^{ss}$

seems to aminorate the trade-off between accuracy and computing time that compromises the reliability of projection methods. Aruoba et al. (2006) follow an alternative rule-of-thumb which is related to the choice of the approximating nodes: solve the model for a small number of basis

Method	\bar{E}_1	\bar{E}_∞
Perturbation 1	-1.6243	-0.5889
Perturbation 2	-2.3272	-0.8374
Collocation	-5.4546	-4.8822
Least Squares	-5.3024	-4.5718
Linear-Quadratic (cons.)	-0.8691	-0.1122
Linear-Quadratic (inv.)	-0.6340	-0.1147

Table 4: HJB residuals for benchmark calibration

functions and then use that solution as an initial guess in a system with more basis functions after using a proper interpolation scheme.

Figure 4 plots a transversal cut of the HJB equation residuals for $A_t = A^{ss}$. Figure 5 in B presents the HJB equation residuals on both dimensions of the state-space for each of the numerical procedures. The superiority along all the state space of the global methods relative to perturbation it is clear from the results.

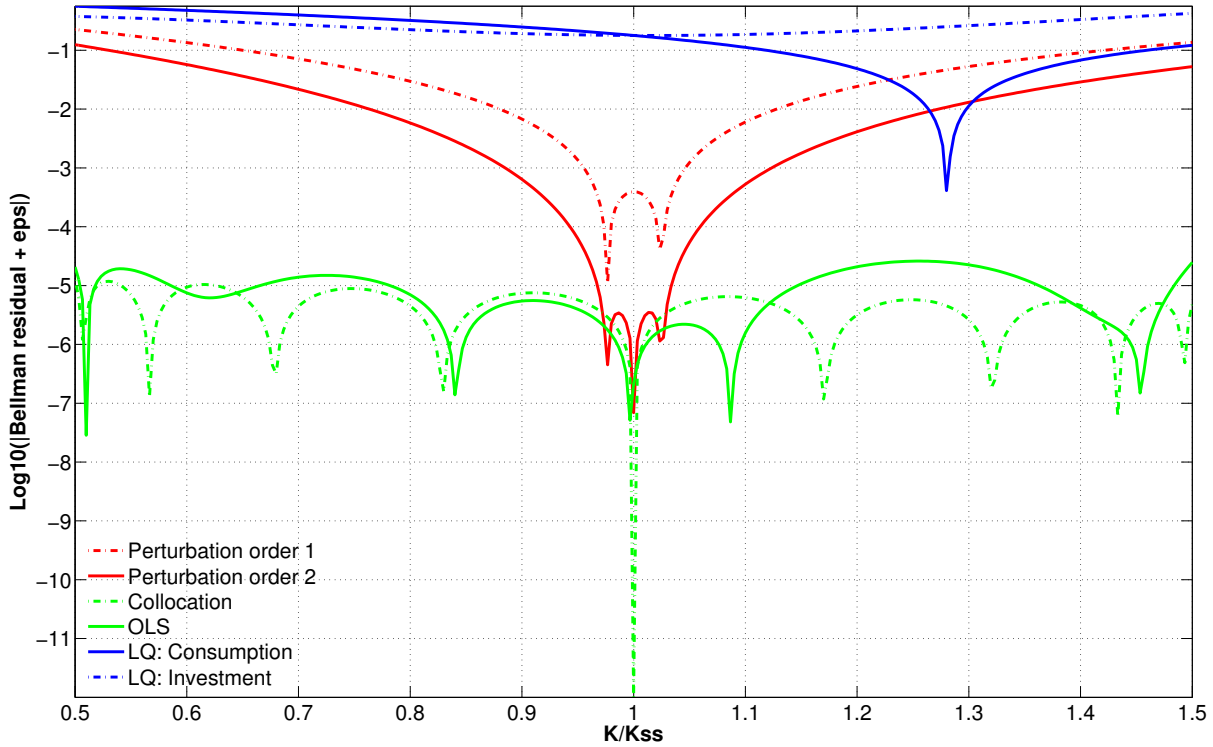


Figure 4: HJB residuals

4.3 Robustness check

Tables 5 and 6 presents the average and maximum HJB equation residuals obtained under the alternative calibrations in Table 2. In particular I analyze the effects of different values of the risk aversion parameter and of increased volatility in both the shocks to productivity and the capital

stock. The value and policy functions as well as the transversal cuts of the residuals are not plotted due to space considerations but are available upon request.

	Pert. 1	Pert. 2	Coll.	LS	LQ (cons.)	LQ (inv.)
M1	-1.7534	-2.5994	-3.5808	-3.5808	-1.1115	-0.9203
M2	-0.0404	-0.4921	-5.4004	-4.4462	0.4303	0.9677
M3	-1.7534	-2.5997	-3.5808	-3.5807	-1.1115	-0.9203
M4	-1.6241	-2.3272	-5.4546	-5.4406	-0.8691	-0.6339
M5	-0.0408	-0.4921	-5.4006	-5.1345	0.4306	0.9677
M6	-1.3447	-1.8854	-3.1957	-3.1957	-0.8112	-0.5140
M7	-1.4026	-1.7357	-4.8053	-4.7996	-0.5698	-0.3513
M8	-1.3417	-1.9640	-3.1957	-3.1952	-0.8112	-0.4495
M9	-1.4025	-1.7357	-4.8155	-4.8128	-0.5698	-0.3513
Extreme I	0.4966	0.2155	-3.0293	-3.0293	0.8952	1.3156
Extreme II	0.4967	0.2157	-3.0295	-3.0294	0.8952	1.3156

Table 5: Average HJB residuals for alternative calibrations (\tilde{E}_1)

	Pert. 1	Pert. 2	Coll.	LS	LQ (cons.)	LQ (inv.)
M1	-0.7104	-1.0813	-2.9793	-2.9793	-0.4211	-0.3793
M2	1.2829	1.2766	-4.4311	-3.7181	1.5801	1.6142
M3	-0.7104	-1.0813	-2.9793	-2.9801	-0.4211	-0.3793
M4	-0.5890	-0.8375	-4.8822	-4.8511	-0.1122	-0.1146
M5	1.2830	1.2766	-4.4313	-4.2401	1.5801	1.6142
M6	-0.5497	-0.7261	-2.4306	-2.4306	-0.0643	0.0441
M7	-0.3938	-0.7039	-3.1427	-3.1254	0.2949	0.4192
M8	-0.5536	-0.7252	-2.4306	-2.4305	-0.0643	0.1070
M9	-0.3984	-0.7039	-3.1456	-3.1476	0.2949	0.4192
Extreme 1	2.8473	2.5840	-2.3470	-2.3740	2.0828	2.2600
Extreme 2	2.8474	2.5841	-2.3742	-2.3742	2.0828	2.2600

Table 6: Maximum HJB residuals for alternative calibrations (\tilde{E}_∞)

In terms of the relative performance of the numerical methods the results are robust to changes in the degree of non-linearities. Nevertheless, global methods require additional computing time to converge relative to the benchmark case especially when the coefficient of risk aversion is high. In other words changes in the concavity of the utility function, and hence, in the concavity of the value function reduce the computational efficiency of projection methods. Furthermore this effect dominates when compared to the additional computing time induced by changes in the variance of the shocks. Despite this result, global methods exhibit a superior level of accuracy relative to perturbations and linear-quadratic approximations. The performance of the local methods deeply deteriorates for high levels of the risk aversion parameter as can be seen from HJB equation residuals crossing the zero line under the extreme calibrations and scenarios M2 and M5.

To reduce the impact of higher non-linearities on the implementation of projection methods the

rule-of-thumb suggested in Section 4.2 is complemented in the following recursive way: approximate the unknown functions on a narrow grid for the capital stock, e.g. between $[0.95K^{ss}, 1.05K^{ss}]$, using as an initial guess the solution from perturbation. Once a solution has been found increase the size of this lattice by a small amount, e.g. $\pm 0.05K^{ss}$, while keeping fixed the number of nodes in both dimensions and use as an initial value the previous solution. Continue in this fashion until the solution for the desired grid has been found.

A puzzling result regarding the approximated policy functions is found. Contrary to the results in Aruoba et al. (2006), when the risk aversion is high (and probably unrealistic) and the volatility of shocks are either low or high, the levels of the consumption and labor supply functions derived by perturbation and the linear-quadratic method differ substantially from the levels obtained by global methods: on average 51% more labor is supplied and consumption decreases 23%. Aruoba et al. (2006) argue that the difference in levels of the linear and log-linear approximations relative to the other methods (including second order perturbation) is due to the certainty equivalence property that holds for discrete-time models. However, as mentioned before, the first order perturbation for the continuous-time case does not exhibit this property and still produces different levels of the policy functions. Thus the question regarding what explains the different levels obtained through linear or quasi-linear methods remains unanswered and should be addressed in future research¹¹.

Another difference with respect to the results reported in Aruoba et al. (2006) has to do with the threshold level of the risk aversion parameter for which the slope of the labor supply changes from negative to positive. They conclude from their numerical exercises that for levels of $\gamma > 40$ highly risk averse agents will work hard whenever the capital stock is high in order to accumulate even more capital in order to insure themselves against unforeseen bad times. However, I find that for the model of Section 3 a value of $\gamma = 10$ already exhibit this precautionary behaviour of agents.

4.4 A note on computing time

As argued in Aruoba et al. (2006) the computing time can be considered of minor relevance when compared with the programming and debugging time used in coding. Specially because the first indicator does not include the time spent by the researcher in the formulation of the problem and its implementation on the computer. This is particularly relevant for the perturbation method where it is necessary to code the derivatives of both the implicit value and policy functions¹².

Nonetheless Table 7 presents a comparison of the computing times for each of the methods implemented for the model with closed form solution as well as the model without analytical solution using the benchmark calibration. The former gives a clearer idea of the computational efficiency of the each individual methods since I use the true solution of the model as an initial value for the global methods. To compute the approximations when no closed form is available I

¹¹Doubts about the accuracy of second order perturbations has been previously addressed in Becker et al. (2007).

¹²Recall that for perturbation methods I have used Mathematica[®] 7.0 instead of Matlab[®] 7.11.0 (see footnote 10). The main reason for choosing the former is the need of computing derivatives of implicit functions with respect to the state variables and then evaluate this unknown functions at a particular point. To my knowledge this task cannot be performed in the Symbolic Toolbox of Matlab[®].

follow the rule-of-thumb described earlier. However it is important to keep in mind that this rule favors the collocation method over least squares since more time is used in the later in order to find a good initial candidate for the former. For all the computations I have used a 2.53GHz Intel® Core™2 Duo running Windows 7.

From the first two columns it is evident how local approximation methods, which do not need an initial guess, outperform global methods. However, as discussed in previous sections, this comes with a cost in terms of accuracy, especially for highly non-linear models. In contrast, conditional on a correct initial value, the collocation method provides the best balance between speed and accuracy offering an even lower computing time than second order perturbation and at the same time a global precision in the approximation of the unknown functions.

The last column of Table 7 reports the ratio of computational time used by the different methods in the discrete-time DSGE model to the computational time used by the same methods applied to the continuous-time version of the same model. For the former, I use the companion codes of Aruoba et al. (2006)¹³. For the calculations is use the model with no closed form solution.

The results indicate that perturbation methods use much less time when applied to continuous-time models. In fact, for first order approximations discrete-time models take approximately 40 times more time to deliver the value and policy functions, while for second order approximations the computational time is 5 times that of continuous-time models. Mixed results are obtained when we compare projection methods. Collocation methods are more efficient in continuous-time setups but least square methods perform better in discrete-time models.

Method	Analytical solution	No analytical solution	Computing time ratio (DT/CT)
Perturbation 1	2.900E-2	4.900E-2	40.816
Perturbation 2	8.500E-2	5.770E-1	5,746
Collocation	5.928E-2	2.976E+0	572,043
Least Squares	1.078E+1	9.648E+2	-97.927
Linear-Quadratic (cons.)	2.448E-1	6.511E-1	N.A.
Linear-Quadratic (inv.)	1.613E+0	2.167E+0	N.A.

Table 7: Computing times for alternative approximation methods using the benchmark calibration (in seconds)

5 Concluding remarks

This paper describes and compares different numerical methods to approximate the solution of continuous-time DSGE models in the spirit of Taylor and Uhlig (1990), Aruoba et al. (2006) and Caldara et al. (2012). The stochastic Ramsey model described in Merton (1990) with endogenous labor and stationary stochastic variation in total factor productivity is used as a test case. Under

¹³It is important to mention that in the case of projection methods, Aruoba et al. (2006) use Fortran and not Matlab.

plausible parametrizations this version of the standard neoclassical growth model admits a closed form solution that can be used to check the accuracy of the different solution methods by measuring how far the approximations are from the true solution. When the model is left unrestricted alternative measures based on the Hamilton-Jacobi-Bellman (HJB) equation residuals are used to evaluate the performance of the approximations. I study both global and local methods to approximate the value and policy functions that characterize the optimal behaviour of a benevolent planner.

Regarding global methods I implement collocation and least squares projections. For the local methods I focus on first and second order perturbation and linear-quadratic approximations with linear and linearized constraints. Regarding the latter, it is important to remember that: “subject to applicability, other linearization methods like the Eigenvalue decomposition [...], Generalized Schur decomposition [...] or the QZ decomposition among many others should deliver the same results. The linear approximation of a differentiable function is unique and invariant to differentiable parameters transformation.” (working paper version of Aruoba et al. (2006)).

Similarly to the discrete-time case I find projection methods to be overwhelmingly more accurate and robust than perturbation for a wide range of values of the state-space centered around the deterministic steady state. Their accuracy carries over for higher degrees of non-linearities. In particular I study the effect of different values of the risk aversion parameter and higher values of the volatility of the shocks that hit the economy. This becomes a relevant issue not only for its qualitative and quantitative economic implications but also from an econometric perspective. When concerned about the estimation of the structural parameters of the model we will be interested in studying the global shape of the approximated likelihood function. This will not be possible if the solution of the model is built from a local approximation. Furthermore, as shown in Rubio-Ramirez and Fernandez-Villaverde (2005) it is possible to obtain a better fit of the model to the data as well as more accurate point estimates of the moments of the model by exploiting the non-linear structure of the economic model which can only be achieved through the use of global methods. However, all these advantages come at a non-negligible cost. A good initial guess for the value function is required in order to obtain a good approximation which increases the computing time of the solution and hence compromises the feasibility of any econometric procedure.

On the other hand, the fit of perturbation and linear-quadratic approximations deteriorate when the degree of non-linearities is increased. In particular, they deliver approximated value and policy functions with a different level to that obtained by global methods. It has been argued previously in the literature that this could be explained by the lack of volatility correction that the certainty equivalence property implies. However, perturbation methods do not exhibit this property in the continuous-time framework and still the same results are obtained. A future line of research should include the study of this result more deeply.

From an econometric point of view, the use of perturbations is usually preferred since the estimation of the deep parameters of the model is simple and fast given that the implementation of the solution method is easy and quick. However, as described in Fernandez-Villaverde et al. (2006) the approximation error arising from the solution modify the likelihood function at least up to a

first-order degree. In addition, the period by period error made in the approximation accumulates at a rate exactly equal to the rate at which the sample size grows making the approximated likelihood diverge from the exact likelihood of the economic model. One way to circumvent these problems is to increase the order of the approximation in the solution step which comes at a very low computational cost.

Finally, after gathering the computing times of each of the numerical methods I find that continuous-time DSGE modeling prove to be a promising area of future research when compared to the discrete-time framework. The approximation methods use much less computing time in both perturbation and projection methods since there is no need to approximate composition of unknown function neither to numerically approximate the integrals associated with expected values.

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.
- ARUOBA, S. B., J. FERNANDEZ-VILLAVERDE, AND J. F. RUBIO-RAMIREZ (2006): “Comparing solution methods for dynamic equilibrium economies,” *Journal of Economic Dynamics and Control*, 30, 2477–2508.
- BECKER, S., L. GRNE, AND W. SEMMLER (2007): “Comparing accuracy of second-order approximation and dynamic programming,” *Computational Economics*, 30, 65–91.
- BENIGNO, P. AND M. WOODFORD (2012): “Linear-quadratic approximation of optimal policy problems,” *Journal of Economic Theory*, 147, 1–42.
- BINSBERGEN, J. V., J. FERNANDEZ-VILLAVERDE, R. S. KOIJEN, AND J. RUBIO-RAMIREZ (2008): “Likelihood estimation of DSGE models with Epstein-Zin preferences,” 2008 meeting papers, Society for Economic Dynamics.
- BLANCHARD, O. (2009): “The state of macro,” *Annual Review of Economics*, 1, 209–228.
- BRENNAN, M. J. (1998): “The Role of Learning in Dynamic Portfolio Decisions,” *European Finance Review*, 1, 295–306.
- CALDARA, D., J. FERNANDEZ-VILLAVERDE, J. RUBIO-RAMIREZ, AND W. YAO (2012): “Computing DSGE models with recursive preferences and stochastic volatility,” *Review of Economic Dynamics*, 15, 188–206.
- CANDLER, G. V. (2004): “Finite difference methods for continuous time dynamic programming,” in *Computational Methods for the Study of Dynamic Economies*, ed. by R. Marimon and A. Scott, Oxford University Press, chap. 8, 172–194.
- CHANG, F.-R. (2009): *Stochastic Optimization in Continuous Time*, Cambridge University Press, second ed.
- COCHRANE, J. (2005): “Financial Markets and the Real Economy,” NBER Working Papers 11193, National Bureau of Economic Research, Inc.
- COLLARD, F. AND M. JUILLARD (2001): “Accuracy of stochastic perturbation methods: The case of asset pricing models,” *Journal of Economic Dynamics and Control*, 25, 979–999.
- DEN HAAN, W. J. AND A. MARCET (1994): “Accuracy in simulations,” *Review of Economic Studies*, 61, 3–17.

- DORASZELSKI, U. AND K. L. JUDD (2012): “Avoiding the curse of dimensionality in dynamic stochastic games,” *Quantitative Economics*, 3, 53–93.
- FERNANDEZ-VILLAYERDE, J., J. F. RUBIO-RAMIREZ, AND M. S. SANTOS (2006): “Convergence Properties of the Likelihood of Computed Dynamic Models,” *Econometrica*, 74, 93–119.
- FURLANETTO, F. AND M. SENECA (2013): “New Perspectives on Depreciation Shocks as a Source of Business Cycle Fluctuations,” *Macroeconomic Dynamics*, In press.
- GASPAR, J. AND K. JUDD (1997): “Solving large-scale rational-expectations models,” *Macroeconomic Dynamics*, 1, 45–75.
- GRINOLS, E. L. AND S. J. TURNOVSKY (1993): “Risk, the financial market, and macroeconomic equilibrium,” *Journal of Economic Dynamics and Control*, 17, 1–36.
- HEER, B. AND A. MAUSSNER (2009): *Dynamic General Equilibrium Modeling: Computational Methods and Applications*, Springer-Verlag, second ed.
- 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 Mathematics*, ed. by H. R. Varian, Springer-Verlag, chap. 2, 80–103.
- KOMPAS, T. AND L. CHU (2010): “A comparison of parametric approximation techniques to continuous-time stochastic dynamic programming problems,” Environmental Economics Research Hub Research Reports 1071, Environmental Economics Research Hub, Crawford School of Economics and Government, The Australian National University.
- KYDLAND, F. E. AND E. C. PRESCOTT (1982): “Time to Build and Aggregate Fluctuations,” *Econometrica*, 50, 1345–70.
- LJUNGQVIST, L. AND T. J. SARGENT (2004): *Recursive Macroeconomic Theory*, vol. 1, The MIT Press, 2 ed.
- MAGILL, M. J. P. (1977a): “A local analysis of N-sector capital accumulation under uncertainty,” *Journal of Economic Theory*, 15, 211–219.
- (1977b): “Some new results on the local stability of the process of capital accumulation,” *Journal of Economic Theory*, 15, 174–210.
- MERTON, R. C. (1969): “Lifetime portfolio selection under uncertainty: The continuous-time case,” *The Review of Economics and Statistics*, 51, 247–57.

- (1971): “Optimum consumption and portfolio rules in a continuous-time model,” *Journal of Economic Theory*, 3, 373–413.
- (1990): *Continuous Time Finance*, Blackwell Publishers.
- MIRANDA, M. J. AND P. L. FACKLER (2002): *Applied Computational Economics and Finance*, The MIT Press.
- POSCH, O. (2009): “Structural estimation of jump-diffusion processes in macroeconomics,” *Journal of Econometrics*, 153, 196–210.
- (2011): “Risk premia in general equilibrium,” *Journal of Economic Dynamics and Control*, 35, 1557–1576.
- POSCH, O. AND T. TRIMBORN (2013): “Numerical solution of dynamic equilibrium models under Poisson uncertainty,” *Journal of Economic Dynamics and Control*.
- RUBIO-RAMIREZ, J. F. AND J. FERNANDEZ-VILLAVARDE (2005): “Estimating dynamic equilibrium economies: linear versus nonlinear likelihood,” *Journal of Applied Econometrics*, 20, 891–910.
- SWANSON, E. T. (2012): “Risk Aversion and the Labor Margin in Dynamic Equilibrium Models,” *American Economic Review*, 102, 1663–91.
- TAPIERO, C. S. AND A. SULEM (1994): “Computational aspects in applied stochastic control,” *Computational Economics*, 7, 109–46.
- TAYLOR, J. B. AND H. UHLIG (1990): “Solving nonlinear stochastic growth models: A comparison of alternative solution methods,” *Journal of Business & Economic Statistics*, 8, 1–17.
- TURNOVSKY, S. J. (1977): *Macroeconomic Analysis and Stabilization Policies*, Cambridge University Press.
- TURNOVSKY, S. J. AND W. T. SMITH (2006): “Equilibrium consumption and precautionary savings in a stochastically growing economy,” *Journal of Economic Dynamics and Control*, 30, 243–278.
- WACHTER, J. (2013): “Can Time-Varying Risk of Rare Disasters Explain Aggregate Stock Market Volatility?” *The Journal of Finance*, LXVIII.
- WÄLDE, K. (2011): “Production technologies in stochastic continuous time models,” *Journal of Economic Dynamics and Control*, 35, 616–622.
- XIA, Y. (2001): “Learning about Predictability: The Effects of Parameter Uncertainty on Dynamic Asset Allocation,” *The Journal of Finance*, 56, 205–246.

Appendix

A Equivalent benchmark model

By using the aggregate resource constraint of the economy, $Y_t = C_t + I_t$, Equation (5) and Equation (7) the model described in Section (3) can be written in the following equivalent way:

$$V(K_0, A_0) \equiv \max_{\{I_t, L_t\}_{t=0}^{\infty}} \mathbb{E}_0 \int_0^{\infty} e^{-\rho t} \frac{\left((A_t K_t^\alpha L_t^{1-\alpha} - I_t) (1 - L_t)^\psi \right)^{1-\gamma}}{1 - \gamma} dt$$

subject to

$$\begin{aligned} dK_t &= (I_t - \delta K_t) dt + \sigma K_t dB_{K,t} \\ dA_t &= \kappa (\omega - A_t) dt + \eta \sqrt{A_t} dB_{A,t} \end{aligned}$$

where now the control variables are the level of investment, I_t , and the fraction of hours worked, L_t . Note that under this alternative representation, the drift function of both stochastic state processes are linear in the control and state variables.

Once the model has been solved for their respective policy functions it is possible to recover the level of optimal consumption by using the aggregate resource constraint:

$$C_t = A_t K_t^\alpha L_t (K_t, A_t)^{1-\alpha} - I(K_t, A_t).$$

B HJB residuals

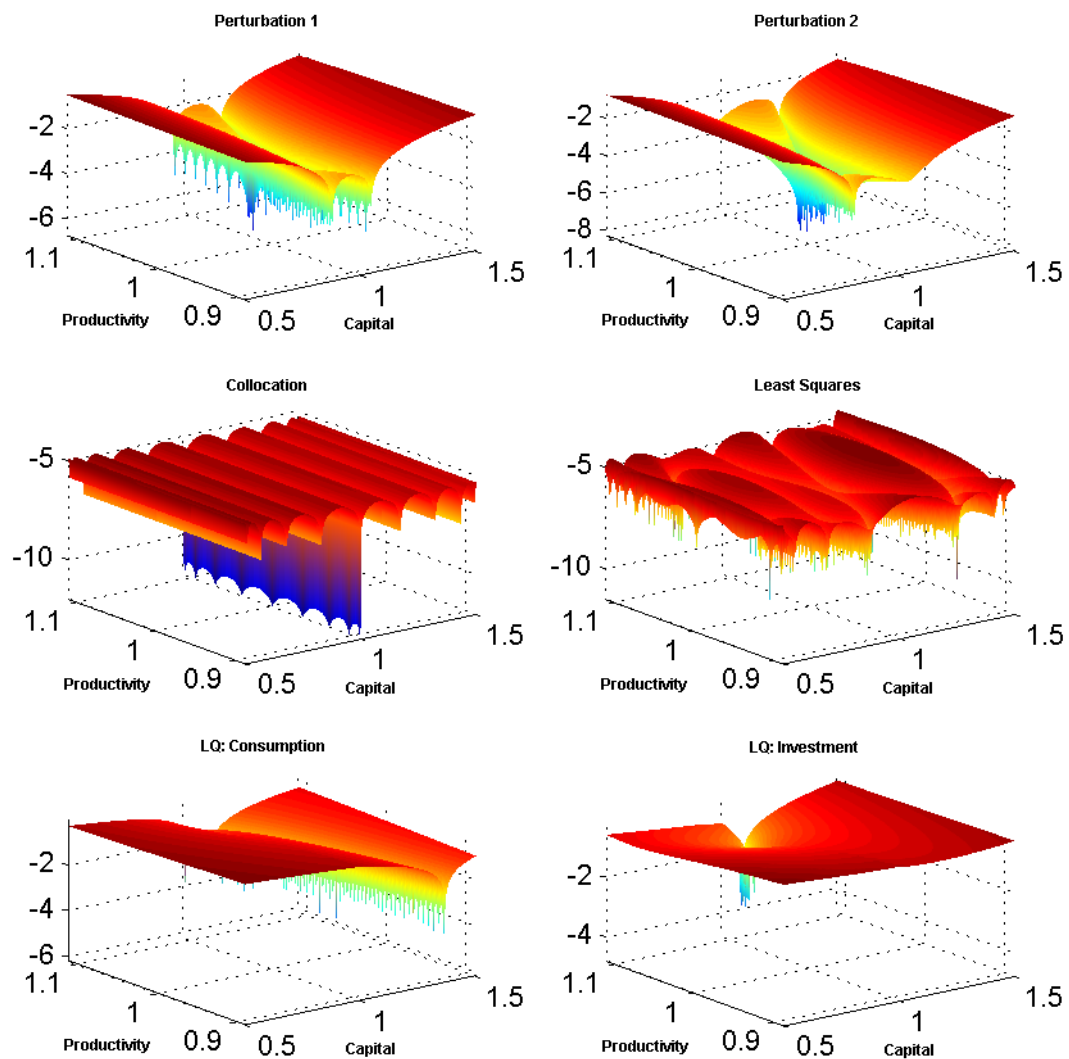


Figure 5: HJB residuals under benchmark calibration and all grid points in state-space

Research Papers 2013



- 2013-21: Malene Kallestrup-Lamb, Anders Bredahl Kock and Johannes Tang Kristensen: Lassoing the Determinants of Retirement
- 2013-22: Johannes Tang Kristensen: Diffusion Indexes with Sparse Loadings
- 2013-23: Asger Lunde and Anne Floor Brix: Estimating Stochastic Volatility Models using Prediction-based Estimating Functions
- 2013-24: Nima Nonejad: A Mixture Innovation Heterogeneous Autoregressive Model for Structural Breaks and Long Memory
- 2013-25: Nima Nonejad: Time-Consistency Problem and the Behavior of US Inflation from 1970 to 2008
- 2013-26: Nima Nonejad: Long Memory and Structural Breaks in Realized Volatility: An Irreversible Markov Switching Approach
- 2013-27: Nima Nonejad: Particle Markov Chain Monte Carlo Techniques of Unobserved Component Time Series Models Using Ox
- 2013-28: Ulrich Hounyo, Sílvia Goncalves and Nour Meddahi: Bootstrapping pre-averaged realized volatility under market microstructure noise
- 2013-29: Jiti Gao, Shin Kanaya, Degui Li and Dag Tjøstheim: Uniform Consistency for Nonparametric Estimators in Null Recurrent Time Series
- 2013-30: Ulrich Hounyo: Bootstrapping realized volatility and realized beta under a local Gaussianity assumption
- 2013-31: Nektarios Aslanidis, Charlotte Christiansen and Christos S. Savva: Risk-Return Trade-Off for European Stock Markets
- 2013-32: Emilio Zanetti Chini: Generalizing smooth transition autoregressions
- 2013-33: Mark Podolskij and Nakahiro Yoshida: Edgeworth expansion for functionals of continuous diffusion processes
- 2013-34: Tommaso Proietti and Alessandra Luati: The Exponential Model for the Spectrum of a Time Series: Extensions and Applications
- 2013-35: Bent Jesper Christensen, Robinson Kruse and Philipp Sibbertsen: A unified framework for testing in the linear regression model under unknown order of fractional integration
- 2013-36: Niels S. Hansen and Asger Lunde: Analyzing Oil Futures with a Dynamic Nelson-Siegel Model
- 2013-37: Charlotte Christiansen: Classifying Returns as Extreme: European Stock and Bond Markets
- 2013-38: Christian Bender, Mikko S. Pakkanen and Hasanjan Sayit: Sticky continuous processes have consistent price systems
- 2013-39: Juan Carlos Parra-Alvarez: A comparison of numerical methods for the solution of continuous-time DSGE models