View metadata, citation and similar papers at core.ac.uk

brought to you by CORE

provided by Research Papers in Economics

NBER WORKING PAPER SERIES

COMPUTING DSGE MODELS WITH RECURSIVE PREFERENCES

Dario Caldara
Jesús Fernández-Villaverde
Juan F. Rubio-Ramírez
Wen Yao

Working Paper 15026
http://www.nber.org/papers/w15026

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
June 2009

We thank Michel Juillard for his help with computational issues and Larry Christiano, Dirk Krueger, and participants at the Penn Macro lunch for comments. Beyond the usual disclaimer, we must note that any views expressed herein are those of the authors and not necessarily those of the Federal Reserve Bank of Atlanta or the Federal Reserve System. Finally, we also thank the NSF for financial support. The views expressed herein are those of the author(s) and do not necessarily reflect the views of the National Bureau of Economic Research.

NBER working papers are circulated for discussion and comment purposes. They have not been peer-reviewed or been subject to the review by the NBER Board of Directors that accompanies official NBER publications.

© 2009 by Dario Caldara, Jesús Fernández-Villaverde, Juan F. Rubio-Ramírez, and Wen Yao. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

Computing DSGE Models with Recursive Preferences
Dario Caldara, Jesús Fernández-Villaverde, Juan F. Rubio-Ramírez, and Wen Yao
NBER Working Paper No. 15026
June 2009
JEL No. C63,C68,E37

ABSTRACT

This paper compares different solution methods for computing the equilibrium of dynamic stochastic general equilibrium (DSGE) models with recursive preferences such as those in Epstein and Zin (1989 and 1991). Models with these preferences have recently become popular, but we know little about the best ways to implement them numerically. To fill this gap, we solve the stochastic neoclassical growth model with recursive preferences using four different approaches: second- and third-order perturbation, Chebyshev polynomials, and value function iteration. We document the performance of the methods in terms of computing time, implementation complexity, and accuracy. Our main finding is that a third-order perturbation is competitive in terms of accuracy with Chebyshev polynomials and value function iteration, while being an order of magnitude faster to run. Therefore, we conclude that perturbation methods are an attractive approach for computing this class of problems.

Dario Caldara
Institute for International Economic Studies
Stockholm University
SE-106 91 Stockholm
Sweden
dario.caldara@gmail.com

Jesús Fernández-Villaverde
University of Pennsylvania
160 McNeil Building
3718 Locust Walk
Philadelphia, PA 19104
and NBER
jesusfv@econ.upenn.edu

Juan F. Rubio-Ramírez
Duke University
P.O. Box 90097
Durham, NC 27708
juan.rubio-ramirez@duke.edu

Wen Yao
University of Pennsylvania
160 McNeil
Philadelphia, PA 19104
wenyao@econ.upenn.edu

1. Introduction

This paper compares different solution methods for computing the equilibrium of dynamic stochastic general equilibrium (DSGE) models with recursive preferences such as those first proposed by Kreps and Porteus (1978) and later generalized by Epstein and Zin (1989 and 1991) and Weil (1990). This exercise is interesting because recursive preferences have recently become very popular in finance and in macroeconomics. Without any attempt at being exhaustive, and just to show the extent of the literature, a few of those papers include Backus, Routledge, and Zin (2007), Bansal, Dittman, and Kiku (2009), Bansal, Gallant, and Tauchen (2008), Bansal and Shaliastovich (2007), Bansal and Yaron (2004), Campanale, Castro, and Clementi (2007), Campbell (1993 and 1996), Campbell and Viceira (2001), Chen, Favilukis and Ludvigson (2007), Croce (2006), Dolmas (1998 and 2007), Gomes and Michealides (2005), Hansen, Heaton, and Li (2008), Kaltenbrunner and Lochstoer (2007), Kruger and Kubler (2005), Lettau and Uhlig (2002), Piazzesi and Schneider (2006), Restoy and Weil (1998), Rudebusch and Swamson (2008), Tallarini (2000), and Uhlig (2007). All of these economists have been attracted by the extra flexibility of separating risk aversion and intertemporal elasticity of substitution (EIS) and by the intuitive appeal of having preferences for early or later resolution of uncertainty.

Despite this variety of papers, little is known about the numerical properties of the different solution methods that solve equilibrium models with recursive preferences. For example, we do not know how well value function iteration (VFI) performs or how good local approximations are compared with global ones. Similarly, if we want to estimate the model, we need to assess what solution method is sufficiently reliable yet quick enough to make the exercise feasible. More important, the most common solution algorithm in the DSGE literature, (log-) linearization, cannot be applied, since it makes us miss the whole point of recursive preferences: the resulting (log-) linear decision rules are certainty equivalent and do not depend on risk aversion. This paper attempts to fill this gap in the literature, and therefore, it complements previous work by Aruoba, Fernández-Villaverde, and Rubio-Ramírez (2006), in which a similar exercise is performed with the neoclassical growth model with CRRA utility function.

We solve and simulate the model using four main approaches: perturbation (of second- and third-order), Chebyshev polynomials, and VFI. By doing so, we span most of the relevant methods in the literature. Our results provide a strong guess of how some other methods not covered here, such as finite elements, would work (rather similar to Chebyshev polynomials but more computationally intensive). We report results for a benchmark calibration of the model and for alternative calibrations that change the variance of the productivity shock,

the risk aversion, and the intertemporal elasticity of substitution. In that way, we study the performance of the methods both for cases close to the CRRA utility function and for highly non-linear cases far away from the CRRA benchmark. For each method, we compute decision rules and the value function, the ergodic distribution of the economy, business cycle statistics, the welfare costs of aggregate fluctuations, and asset prices. Also, we evaluate the accuracy of the solution by reporting Euler equation errors.

We highlight four main results from our exercise. First, all methods provide a high degree of accuracy. Thus, researchers who stay within our set of solution algorithms can be confident that their quantitative answers are sound.

Second, perturbation methods deliver a surprisingly high level of accuracy with considerable speed. We show how second- and third-order perturbation performs remarkably well in terms of accuracy for the benchmark calibration, being fully competitive with VFI or Chebyshev polynomials. For this calibration, a second-order perturbation that runs in one second does as well in terms of the average Euler equation error as a VFI that takes two hours to run. Even in the extreme calibration with high risk aversion and high volatility of productivity shocks, a second-order perturbation works at an acceptable level and a third-order approximation performs nearly as well as VFI. Since, in practice, perturbation methods are the only computationally feasible method to solve the medium-scale DSGE models used for policy analysis that have dozens of state variables (Christiano, Eichenbaum, and Evans, 2005, and Smets and Wouters, 2007), this finding has an outmost applicability. Moreover, since implementing a second-order perturbation is feasible with off-the-shelf software like `Dynare`, which requires minimum programming knowledge by the user, our findings may induce many researchers to explore recursive preferences in further detail. Two final advantages of perturbation is that, often, the perturbed solution provides insights about the economics of the problem and that it might be an excellent initial guess for VFI or for Chebyshev polynomials.

Third, Chebyshev polynomials provide a terrific level of accuracy with reasonable computational burden. When accuracy is most required and the dimensionality of the state space is not too high, they are the obvious choice. Unfortunately, Chebyshev polynomials suffer from the curse of dimensionality, and for more involved models, we would need to apply some aggressive interpolation scheme as in Kruger and Kubler (2005).

Fourth, we were disappointed by the poor performance of VFI, which could not achieve a high accuracy even with a large grid. This suggests that we should relegate VFI to solving those problems where non-differentiabilities complicate the application of the previous methods.

The rest of the paper is organized as follows. In section 2, we present the stochastic neoclassical growth model with recursive preferences. Section 3 describes the different solu-

tion methods used to approximate the decision rules of the model. Section 4 discusses the calibration of the model. Section 5 reports numerical results and section 6 concludes. An appendix provides some additional details.

2. The Stochastic Neoclassical Growth Model with Recursive Preferences

We use the stochastic neoclassical growth model with recursive preferences as our test case. We select this model for two reasons. First, it is the workhorse of modern macroeconomics. Even more complicated New Keynesian models with real and nominal rigidities, such as those in Woodford (2003) or Christiano, Eichenbaum, and Evans (2005), are built around the core of the neoclassical growth model. Therefore, any lesson learned in this model is likely to have a wide applicability in a large class of interesting economies. Second, the model is, except for the form of the utility function, the same test case as in Aruoba, Fernández-Villaverde, and Rubio-Ramírez (2006). This provides us with a large set of results to compare to our new findings.

The description of the model is rather straightforward, and we just go through the minimal details required to fix notation. There is a representative household that has preferences over streams of consumption, c_t, and leisure, $1 - l_t$, that are representable by a recursive function of the form:

$$U_t = \max_{c_t, l_t} \left[(1-\beta) \left(c_t^{\upsilon} \left(1 - l_t \right)^{1-\upsilon} \right)^{\frac{1-\gamma}{\theta}} + \beta \left(\mathbb{E}_t U_{t+1}^{1-\gamma} \right)^{\frac{1}{\theta}} \right]^{\frac{\theta}{1-\gamma}} \tag{1}$$

The parameters in these preferences include β, the discount factor, υ, which controls labor supply, γ, which controls risk aversion, and:

$$\theta = \frac{1-\gamma}{1-\frac{1}{\psi}}$$

where ψ is the EIS. The parameter θ is an index of the deviation with respect to the benchmark CRRA utility function (when $\theta = 1$, we are back in that CRRA case where the inverse of the EIS and risk aversion coincide).

The household's budget constraint of the household is given by:

$$c_t + i_t + \frac{b_{t+1}}{R_t^f} = w_t l_t + r_t k_t + b_t$$

where i_t is investment, R_t^f is the risk-free gross interest rate, b_t is the holding of an uncontingent bond that pays 1 unit of consumption good at time $t+1$, w_t is the wage, l_t is labor,

r_t is the rental rate of capital, and k_t is capital. Asset markets are complete and we could have also included in the budget constraint the whole set of Arrow securities. Since we have a representative household, this is not necessary because the net supply of any security must be equal to zero. The uncontingent bond is sufficient to derive the pricing kernel of the economy m_t since, in equilibrium, $\mathbb{E}_t m_{t+1} R_t^f = 1$. Households accumulate capital according to the law of motion $k_{t+1} = (1-\delta)k_t + i_t$ where δ is the depreciation rate.

The final good in the economy is produced by a competitive firm with a Cobb-Douglas production function $y_t = e^{z_t} k_t^{\zeta} l_t^{1-\zeta}$ where z_t is the productivity level that follows an autoregressive process

$$z_{t+1} = \lambda z_t + \sigma \varepsilon_{t+1}$$

with $\lambda < 1$ and $\varepsilon_{t+1} \sim \mathcal{N}(0,1)$. The parameter σ scales the size of the productivity shocks.[1] Finally, the economy must satisfy the aggregate resource constraint $y_t = c_t + i_t$.

The definition of equilibrium is absolutely standard and we skip it in the interest of space. Also, both welfare theorems hold, a fact that we will exploit by jumping back and forth between the solution of the social planner's problem and the competitive equilibrium. Therefore, an alternative way to write this economy is to look at the value function representation of the social planner's problem:

$$V(k_t, z_t) = \max_{c_t, l_t} \left[(1-\beta) \left(c_t^{\upsilon} (1-l_t)^{1-\upsilon} \right)^{\frac{1-\gamma}{\theta}} + \beta \left(\mathbb{E}_t V^{1-\gamma}(k_{t+1}, z_{t+1}) \right)^{\frac{1}{\theta}} \right]^{\frac{\theta}{1-\gamma}}$$
$$\text{s.t. } c_t + k_{t+1} = e^{z_t} k_t^{\zeta} l_t^{1-\zeta} + (1-\delta) k_t$$
$$z_{t+1} = \lambda z_t + \sigma \varepsilon_{t+1},\ \varepsilon_{t+1} \sim \mathcal{N}(0,1)$$

This formulation emphasizes that we have two state variables for the economy, capital k_t and productivity z_t.

The social planner's problem formulation allows us to find the pricing kernel of the economy:

$$m_{t+1} = \frac{\partial V_t / \partial c_{t+1}}{\partial V_t / \partial c_t}$$

Now, note:

$$\frac{\partial V_t}{\partial c_t} = (1-\beta) V_t^{1-\frac{1-\gamma}{\theta}} \upsilon \frac{(c_t^{\upsilon}(1-l_t)^{1-\upsilon})^{\frac{1-\gamma}{\theta}}}{c_t}$$

[1] We use a stationary model to enhance the usefulness of our results. If we had a deterministic trend, we would only need to adjust β in our calibration below (and the results would be nearly identical). If we had a stochastic trend, we would need to rescale the variables by the productivity level and solve the transformed problem. However, in this case, it is well known that the economy fluctuates less than when $\lambda < 1$, and therefore, all solution methods would be closer, limiting our ability to appreciate differences in performance.

and:

$$\frac{\partial V_t}{\partial c_{t+1}} = \beta V_t^{1-\frac{1-\gamma}{\theta}} (\mathbb{E}_t V_{t+1}^{1-\gamma})^{\frac{1}{\theta}-1} \mathbb{E}_t \left(V_{t+1}^{-\gamma} (1-\beta) V_{t+1}^{1-\frac{1-\gamma}{\theta}} \upsilon \frac{(1-\beta) (c_{t+1}^{\upsilon}(1-l_{t+1})^{1-\upsilon})^{\frac{1-\gamma}{\theta}}}{c_{t+1}} \right)$$

where in the last step we have used the result regarding $\partial V_t / \partial c_t$ forwarded by one period. Then, cancelling redundant terms, we get:

$$m_{t+1} = \frac{\partial V_t / \partial c_{t+1}}{\partial V_t / \partial c_t} = \beta \left(\frac{c_{t+1}^{\upsilon}(1-l_{t+1})^{1-\upsilon}}{c_t^{\upsilon}(1-l_t)^{1-\upsilon}} \right)^{\frac{1-\gamma}{\theta}} \frac{c_t}{c_{t+1}} \left(\frac{V_{t+1}^{1-\gamma}}{\mathbb{E}_t V_{t+1}^{1-\gamma}} \right)^{1-\frac{1}{\theta}} \tag{2}$$

This equation shows how the pricing kernel is affected by the presence of recursive preferences. If $\theta = 1$, the last term,

$$\left(\frac{V_{t+1}^{1-\gamma}}{\mathbb{E}_t V_{t+1}^{1-\gamma}} \right)^{1-\frac{1}{\theta}} \tag{3}$$

is equal to 1 and we get back the pricing kernel of the standard stochastic neoclassical growth model. If $\theta \neq 1$, the pricing kernel is twisted by (3).

We identify the net return on equity (conditional on realization of the productivity shock z_{t+1}) with the marginal net return on investment. That is, we posit that:

$$R_{t+1}^k = \zeta e^{z_{t+1}} k_{t+1}^{\zeta-1} l_{t+1}^{1-\zeta} - \delta$$

with expected return $\mathbb{E}_t \left[R_{t+1}^k \right]$.

3. Solution Methods

We are interested in comparing different solution methods to approximate the dynamics of the previous model. Since the literature on computational methods is large, it would be cumbersome to review every proposed method. Instead, we select the solution methods that we find most promising.

The first method we pick is perturbation (introduced by Judd and Guu, 1992 and 1997, and particularly well explained in Schmitt-Grohé and Uribe, 2004). Perturbation algorithms build a Taylor series expansion of the agents' decision rules around an appropriate point (usually the steady state of the economy) and a perturbation parameter (in our case, the volatility of the productivity shocks). In many situations, perturbation methods have proven to be very fast and, despite their local nature, to be highly accurate in a large range of values of the state variables (see, for instance, Aruoba, Fernández-Villaverde, and Rubio-

Ramírez, 2006). This means that, in practice, perturbations are the only method that can handle models with dozens of state variables in any reasonable amount of time. Moreover, perturbation often provides insights into the structure of the solution and on the economics of the model. Finally, linearization and log-linearization, the most common solution methods for DSGE models, are a particular case of a perturbation of first order.

We implement a second- and a third-order perturbation of our model. A first-order perturbation is useless for our investigation: the resulting decision rules are certainty equivalent and, therefore, they depend only on the EIS and not at all on risk aversion (that is, the first-order decision rules of the model with recursive preferences coincide with the decision rules of the model with CRRA preferences with the same EIS for *any* value of the risk aversion). In comparison, the second-order decision rules incorporate a constant term that depends on risk aversion (Binsbergen *et al.*, 2009) and, hence, allows recursive preferences to play a role. Also, a second-order perturbation can be run with standard software, such as `Dynare`, which opens the door for performing perturbations to many applied researchers who fear to tread through the sandbars of coding analytic derivatives. The third-order approximation provides additional terms to increase accuracy and, in the case of those functions pricing assets, a time-varying risk-premium. For our purposes, a third-order will provide enough accuracy without unnecessary complications.

The second method is a projection algorithm with Chebyshev polynomials (Judd, 1992). Projection algorithms build approximated decision rules that minimize a residual function that measures the distance between the left- and right-hand side of the equilibrium conditions of the model. Projection methods are attractive because they offer a global solution over the whole range of the state space. Their main drawback is that they suffer from an acute curse of dimensionality that makes it quite challenging to extend then to models with many state variables. Among the many different types of projection methods, Aruoba, Fernández-Villaverde, and Rubio-Ramírez (2006) show that Chebyshev polynomials are particularly efficient. Other projection methods, such as finite elements or parameterized expectations, tend to perform somewhat worse than Chebyshev polynomials, and therefore, in the interest of space, we do not consider them.

Finally, we compute the model using VFI. VFI is slow and it suffers as well from the curse of dimensionality, but it is safe, reliable, and we know its convergence properties. Thus, it is a natural default algorithm for the solution of DSGE models.[2]

We describe now each of the different methods in more detail. Then, we calibrate the model and present numerical results.

[2]Epstein and Zin (1989) show that a version of the contraction mapping theorem holds in the value function of the problem with recursive preferences.

3.1. Perturbation

We start by explaining how to use a perturbation approach to solve DSGE models using the value function of the household. We are not the first, of course, to explore the perturbation of value function problems. Judd (1998) already presents the idea of perturbing the value function instead of the equilibrium conditions of a model. Unfortunately, he does not elaborate much on the topic. Schmitt-Grohé and Uribe (2005) employ a perturbation approach to find a second-order approximation to the value function that allows them to rank different fiscal and monetary policies in terms of welfare. However, we follow Binsbergen *et al.* (2009) in their emphasis on the generality of the approach, and we discuss some of its theoretical and numerical advantages.

The perturbation method is linked with Benigno and Woodford (2006) and Hansen and Sargent (1995). Benigno and Woodford present a new linear-quadratic approximation to solve optimal policy problems that avoids some problems of the traditional linear-quadratic approximation when the constraints of the problem are non-linear.[3] In particular, Benigno and Woodford find the correct local welfare ranking of different policies. The method in this paper, as in theirs, can deal with non-linear constraints and obtains the correct local approximation to welfare and policies. One advantage of the method presented here is that it is easily generalizable to higher-order approximations without adding further complications. Hansen and Sargent (1995) modify the linear-quadratic regulator problem to include an adjustment for risk. In that way, they can handle some versions of recursive utilities, such as the ones that motivate our investigation. Hansen and Sargent's method, however, requires imposing a tight functional form for future utility and to surrender the assumption that risk-adjusted utility is separable across states of the world. The perturbation we have presented does not suffer from those limitations.

To illustrate the procedure, we limit our exposition to deriving the second-order approximation to the value function and the decision rules of the agents. Higher-order terms are derived in similar ways, but the algebra becomes too cumbersome to be developed explicitly in this paper (in our application, the symbolic algebra is undertaken by a computer employing `Mathematica`, which automatically generates Fortran 95 code that we can evaluate numerically). Hopefully, our steps will be enough to allow the reader to understand the main thrust of the procedure and to let the reader obtain higher-order approximations by herself.

The first step is to rewrite the productivity process in terms of a perturbation parameter χ,

$$z_{t+1} = \lambda z_t + \chi \sigma \varepsilon_{t+1}.$$

[3] See also Levine, Pearlman, and Pierse (2007) for a similar treatment of the problem.

When $\chi = 1$, which is just a normalization of the perturbation parameter implied by the standard deviation of the shock σ, we are dealing with the stochastic version of the model. When $\chi = 0$, we are dealing with the deterministic case with steady state k_{ss} and $z_{ss} = 0$. Then, we can write the social planner's value function, $V(k_t, z_t; \chi)$, and the decision rules for consumption, $c(k_t, z_t; \chi)$, investment, $i(k_t, z_t; \chi)$, capital, $k(k_t, z_t; \chi)$, and labor, $l(k_t, z_t; \chi)$, as a function of the two states, k_t and z_t, and the perturbation parameter χ.

The second step is to note that, under differentiability conditions, the second-order Taylor approximation of the value function around the deterministic steady state $(k_{ss}, 0; 0)$ is:

$$\begin{aligned} V(k_t, z_t; \chi) \simeq\ & V_{ss} + V_{1,ss}(k_t - k_{ss}) + V_{2,ss} z_t + V_{3,ss}\chi \\ & + \frac{1}{2}V_{11,ss}(k_t - k_{ss})^2 + \frac{1}{2}V_{12,ss}(k_t - k_{ss}) z_t + \frac{1}{2}V_{13,ss}(k_t - k_{ss})\chi \\ & + \frac{1}{2}V_{21,ss} z_t (k_t - k_{ss}) + \frac{1}{2}V_{22,ss} z_t^2 + \frac{1}{2}V_{23,ss} z_t \chi \\ & + \frac{1}{2}V_{31,ss}\chi(k_t - k_{ss}) + \frac{1}{2}V_{32,ss}\chi z_t + \frac{1}{2}V_{33,ss}\chi^2 \end{aligned}$$

where $V_{ss} = V(k_{ss}, 0; 0)$, $V_{i,ss} = V_i(k_{ss}, 0; 0)$ for $i = \{1, 2, 3\}$, and $V_{ij,ss} = V_{ij}(k_{ss}, 0; 0)$ for $i, j = \{1, 2, 3\}$. We can extend this notation to higher-order derivatives of the value function. This expansion could also be performed around a different point of the state space, such as the mode of the ergodic distribution of the state variables. In section 5, we discuss this point further.

By certainty equivalence, we will have that $V_{3,ss} = V_{13,ss} = V_{23,ss} = 0$. Below, we will argue that this is indeed the case (in fact, all the terms in odd powers of χ have coefficient values equal to zero). Moreover, taking advantage of the equality of cross-derivatives, and setting $\chi = 1$, the approximation we look for has the simpler form:

$$\begin{aligned} V(k_t, z_t; 1) \simeq\ & V_{ss} + V_{1,ss}(k_t - k_{ss}) + V_{2,ss} z_t \\ & + \frac{1}{2}V_{11,ss}(k_t - k_{ss})^2 + \frac{1}{2}V_{22,ss} z_t^2 + V_{12,ss}(k_t - k_{ss}) z_t + \frac{1}{2}V_{33,ss} \end{aligned}$$

Binsbergen *et al.* (2009) demonstrate that γ does not affect the values of any of the coefficients except $V_{33,ss}$ and also that $V_{33,ss} \neq 0$. Hence, we have a different approximation from the one resulting from the standard linear-quadratic approximation to the utility functions, where all the constants disappear. However, this result is intuitive, since the value function of a risk-adverse agent is in general affected by uncertainty and we want to have an approximation with terms that capture this effect and allow for the appropriate welfare ranking of decision rules.

Indeed, $V_{33,ss}$ has a straightforward interpretation. At the deterministic steady state, we have:

$$V\left(k_{ss}, 0; 1\right) \simeq V_{ss} + \frac{1}{2} V_{33,ss}$$

Hence

$$\frac{1}{2} V_{33,ss}$$

is a measure of the welfare cost of the business cycle, that is, of how much utility changes when the variance of the productivity shocks is σ^2 instead of zero.[4] This term is an accurate evaluation of the third-order of the welfare cost of business cycle fluctuations because all of the third-order terms in the approximation of the value function either have zero coefficient values (for example, $V_{333,ss} = 0$) or drop when evaluated at the deterministic steady state.

This cost of the business cycle can easily be transformed into consumption equivalent units. We can compute the decrease in consumption τ that will make the household indifferent between consuming $(1 - \tau) c_{ss}$ units per period with certainty or c_t units with uncertainty. To do so, note that the steady-state value function is just $V_{ss} = c_{ss}^{\upsilon} \left(1 - l_{ss}\right)^{1-\upsilon}$, which implies that:

$$c_{ss}^{\upsilon} \left(1 - l_{ss}\right)^{1-\upsilon} + \frac{1}{2} V_{33,ss} = \left(\left(1 - \tau\right) c_{ss}\right)^{\upsilon} \left(1 - l_{ss}\right)^{1-\upsilon}$$

or:

$$V_{ss} + \frac{1}{2} V_{33,ss} = \left(1 - \tau\right)^{\upsilon} V_{ss}$$

Then:

$$\tau = 1 - \left[1 + \frac{1}{2} \frac{V_{33,ss}}{V_{ss}}\right]^{\frac{1}{\upsilon}}$$

Also, notice that we are perturbing the value function in levels of the variables. However, there is nothing special about levels and we could have done the same in logs (a common practice when linearizing DSGE models) or in any other function of the states. These changes of variables may improve the performance of perturbation (Fernández-Villaverde and Rubio-Ramírez, 2006). By doing the perturbation in levels, we are picking the most conservative case for perturbation. Since one of the conclusions that we will reach from our numerical results is that perturbation works surprisingly well in terms of accuracy, that conclusion will only be reinforced by an appropriate change of variables.[5]

[4] This quantity is not necessarily negative. In fact, in some of our computations below, it will be positive.

[5] This comment begets the question, nevertheless, of why we did not perform a perturbation in logs, since many economists will find it more natural than in levels. Our experience with the CRRA utility case (Aruoba, Fernández-Villaverde, and Rubio-Ramírez, 2006) and some computations with recursive preferences not included in the paper suggest that a perturbation in logs does worse than a perturbation in levels. Thus, we continue in the paper with a perturbation in levels.

The decision rules can be expanded in exactly the same way. For example, the second-order approximation of the decision rule for consumption is:

$$\begin{aligned} c\left(k_t, z_t; \chi\right) \simeq\ & c_{ss} + c_{1,ss}\left(k_t - k_{ss}\right) + c_{2,ss} z_t + c_{3,ss}\chi \\ & + \frac{1}{2} c_{11,ss}\left(k_t - k_{ss}\right)^2 + \frac{1}{2} c_{12,ss}\left(k_t - k_{ss}\right) z_t + \frac{1}{2} c_{13,ss}\left(k_t - k_{ss}\right)\chi \\ & + \frac{1}{2} c_{21,ss} z_t \left(k_t - k_{ss}\right) + \frac{1}{2} c_{22,ss} z_t^2 + \frac{1}{2} c_{23,ss} z_t \chi \\ & + \frac{1}{2} c_{31,ss}\chi\left(k_t - k_{ss}\right) + \frac{1}{2} c_{32,ss}\chi z_t + \frac{1}{2} c_{33,ss}\chi^2 \end{aligned}$$

where $c_{ss} = c\left(k_{ss}, 0; 0\right)$, $c_{i,ss} = c_i\left(k_{ss}, 0; 0\right)$ for $i = \{1, 2, 3\}$, $c_{ij,ss} = c_{ij}\left(k_{ss}, 0; 0\right)$ for $i, j = \{1, 2, 3\}$. In a similar way to the approximation of the value function, Binsbergen *et al.* (2009) show that γ does not affect the values of any of the coefficients except $c_{33,ss}$. This term is a constant that captures precautionary behavior caused by the presence of uncertainty. This observation tells us two important facts. First, a linear approximation to the decision rule does not depend on γ (it is certainty equivalent), and therefore, if we are interested in recursive preferences, we need to go at least to a second-order approximation. Second, the difference between the second-order approximation to the decision rules of a model with CRRA preferences and a model with recursive preferences is a constant.[6] This constant generates a second, indirect effect, because it changes the ergodic distribution of the state variables and, hence, the points where we evaluate the decision rules along the equilibrium path. These arguments demonstrate how perturbation methods can provide analytic insights beyond computational advantages and help in understanding the numerical results in Tallarini (2000). In the third-order approximation, all of the terms that depend on functions of χ^2 depend on γ.

Similarly, we can derive the decision rules for labor, investment, and capital. In addition we have functions that give us the evolution of other variables of interest, such as the pricing kernel or the risk-free gross interest rate R_t^f. All of these functions have the same structure and properties regarding γ as the decision rule for consumption. In the case of functions pricing assets, the second-order approximation generates a constant risk premium, while the third-order approximation creates a time-varying risk premium.

Once we have reached this point, there are two paths we can follow to solve for the coefficients of the perturbation. The first procedure is to write down the equilibrium conditions of the model plus the definition of the value function. Then, we take successive derivatives

[6] When all of the parameters of the two versions of the model are the same, except risk aversion, which in the CRRA case is restricted to be equal to the inverse of the EIS while in the recursive preference case it is not.

in this augmented set of equilibrium conditions and solve for the unknown coefficients. This approach, which we call equilibrium conditions perturbation (ECP), allows us to get, after n iterations, the n-th-order approximation to the value function and to the decision rules.

A second procedure is to take derivatives of the value function with respect to states and controls and use those derivatives to find the unknown coefficient. This approach, which we call value function perturbation (VFP), delivers after $(n+1)$ steps, the $(n+1)$-th order approximation to the value function and the $n-$th order approximation to the decision rules.[7] Loosely speaking, ECP undertakes the first step of VFP by hand by forcing the researcher to derive the equilibrium conditions.

The ECP approach is simpler but it relies on our ability to find equilibrium conditions that do not depend on derivatives of the value function. Otherwise, we need to modify the equilibrium conditions to include the definitions of the derivatives of the value function. Even if this is possible to do (and not particularly difficult), it amounts to solving a problem that is equivalent to VFP. This observation is important because it is easy to postulate models that have equilibrium conditions where we cannot get rid of all the derivatives of the value function (for example, in problems of optimal policy design). ECP is also faster from a computational perspective. However, VFP is only trivially more involved because finding the $(n+1)$-th-order approximation to the value function on top of the n-th order approximation requires nearly no additional effort.

The algorithm presented here is based on the system of equilibrium equations derived using the ECP. In the appendix, we show how to derive a system of equations using the VFP. We take the first-order conditions of the social planner. First, with respect to consumption today:

$$\frac{\partial V_t}{\partial c_t} - \lambda_t = 0$$

where λ_t is the Lagrangian multiplier associated with the resource constraint. Second, with respect to capital:

$$-\lambda_t + \mathbb{E}_t \lambda_{t+1} \left(\zeta e^{z_{t+1}} k_{t+1}^{\zeta-1} l_{t+1}^{1-\zeta} + 1 - \delta \right) = 0$$

Third, with respect to labor:

$$\frac{1-\upsilon}{\upsilon} \frac{c_t}{(1-l_t)} = (1-\zeta) e^{z_t} k_t^{\zeta} l_t^{-\zeta}$$

[7] The classical strategy of finding a quadratic approximation of the utility function to derive a linear decision rule is a second-order example of VFP (Anderson *et al.*, 1996). A standard linearization of the equilibrium conditions of a DSGE model when we add the value function to those equilibrium conditions is a simple case of ECP. This is, for instance, the route followed by Schmitt-Grohé and Uribe (2005).

Then, we can write the Euler equation:

$$\mathbb{E}_t m_{t+1}\left(\zeta e^{z_t} k_{t+1}^{\zeta-1} l_{t+1}^{1-\zeta} + 1 - \delta\right) = 1$$

where m_{t+1} was derived above in equation (2). Note that, as we explained above, the derivatives of the value function in (2) are eliminated.

Once we substitute for the pricing kernel, the augmented equilibrium conditions are:

$$V_t - \left[(1-\beta)\left(c_t^{\upsilon}\left(1-l_t\right)^{1-\upsilon}\right)^{\frac{1-\gamma}{\theta}} + \beta\left(\mathbb{E}_t V^{1-\gamma}\left(k_{t+1}, z_{t+1}\right)\right)^{\frac{1}{\theta}}\right]^{\frac{\theta}{1-\gamma}} = 0$$

$$\mathbb{E}_t\left[\beta\left(\frac{c_{t+1}}{c_t}\right)^{\frac{1-\gamma}{\theta}-1}\left(\frac{V_{t+1}^{1-\gamma}}{\mathbb{E}_t V_{t+1}^{1-\gamma}}\right)^{1-\frac{1}{\theta}}\left(\zeta e^{z_{t+1}} k_{t+1}^{\zeta-1} l_{t+1}^{1-\zeta} + 1 - \delta\right)\right] - 1 = 0$$

$$\frac{1-\upsilon}{\upsilon}\frac{c_t}{(1-l_t)} = (1-\zeta) e^{z_t} k_t^{\zeta} l_t^{-\zeta} = 0$$

$$\mathbb{E}_t \beta\left(\frac{c_{t+1}}{c_t}\right)^{\frac{1-\gamma}{\theta}-1}\left(\frac{V_{t+1}^{1-\gamma}}{\mathbb{E}_t V_{t+1}^{1-\gamma}}\right)^{1-\frac{1}{\theta}} R_t^f - 1 = 0$$

$$c_t + i_t - e^{z_t} k_t^{\zeta} l_t^{1-\zeta} = 0$$

$$k_{t+1} - i_t - (1-\delta) k_t$$

plus the law of motion for productivity $z_{t+1} = \lambda z_t + \chi\sigma\varepsilon_{t+1}$ and where we have dropped the max operator in front of the value function because we are already evaluating it at the optimum. In more compact notation,

$$F\left(k_t, z_t, \chi\right) = \mathbf{0}$$

where F is a 6-dimensional function (and where all the endogenous variables in the previous equation are not represented explicitly because they are functions themselves of k_t, z_t, and χ) and $\mathbf{0}$ is the vectorial zero.

In steady state, $m_{ss} = \beta$ and the set of equilibrium conditions simplifies to:

$$V_{ss} = c_{ss}^{\upsilon}\left(1-l_{ss}\right)^{1-\upsilon}$$

$$\left(\zeta k_{ss}^{\zeta-1} l_{ss}^{1-\zeta} + 1 - \delta\right) = 1/\beta$$

$$\frac{1-\upsilon}{\upsilon}\frac{c_{ss}}{(1-l_{ss})} = (1-\zeta) k_{ss}^{\zeta} l_{ss}^{-\zeta}$$

$$\begin{aligned} R^f_{ss} &= 1/\beta \\ c_{ss} + i_{ss} &= k^{\zeta}_{ss} l^{1-\zeta}_{ss} \\ i_{ss} &= \delta k_{ss} \end{aligned}$$

a system of 6 equations on 6 unknowns, V_{ss}, c_{ss}, k_{ss}, i_{ss}, l_{ss}, and R^f_{ss} that can be easily solved (see the appendix for the derivations). This steady state is identical to the steady state of the real business cycle model with a standard CRRA utility function.

To find the first-order approximation to the value function and the decision rules, we take first derivatives of the function F with respect to the states (k_t, z_t) and to the perturbation parameter χ and evaluate them at the deterministic steady state $(k_{ss}, 0; 0)$ that we just found:

$$F_i\left(k_{ss}, 0; 0\right) = \mathbf{0} \text{ for } i = \{1, 2, 3\}$$

This step gives us 18 different first derivatives (6 equilibrium conditions times the 3 variables of F). Since each dimension of F is equal to zero for all possible values of k_t, z_t, and χ, their derivatives must also be equal to zero. Therefore, once we substitute in the values of the steady state, we have a quadratic system of 18 equations on 18 unknowns: $V_{1,ss}$, $V_{2,ss}$, $V_{3,ss}$, $c_{1,ss}$, $c_{2,ss}$, $c_{3,ss}$, $i_{1,ss}$, $i_{2,ss}$, $i_{3,ss}$, $k_{1,ss}$, $k_{2,ss}$, $k_{3,ss}$, $l_{1,ss}$, $l_{2,ss}$, $l_{3,ss}$, $R^f_{1,ss}$, $R^f_{2,ss}$, and $R^f_{3,ss}$. One of the solutions is an unstable root of the system that violates the transversality condition of the problem and we eliminate it. Thus, we keep the solution that implies stability. In the solution, it is easy to see that $V_{3,ss} = c_{3,ss} = k_{3,ss} = i_{3,ss} = l_{3,ss} = R^f_{3,ss} = 0$, that is, we have certainty equivalence. This result is not a surprise and it could have been guessed from a reading of Schmitt-Grohé and Uribe (2004).

To find the second-order approximation, we take derivatives on the first derivatives of the function F, again with respect to the states (k_t, z_t), and the perturbation parameter χ:

$$F_{ij}\left(k_{ss}, 0; 0\right) = \mathbf{0} \text{ for } i, j = \{1, 2, 3\}$$

This step gives us a new system of equations. Then, we plug in the terms that we already know from the steady state and from the first-order approximation and we get that the only unknowns left are the second-order terms of the value function and other functions of interest. Quite conveniently, this system of equations is linear and it can be solved quickly. Repeating these steps (taking higher-order derivatives, plugging in the terms already known, and solving for the remaining unknowns), we can get any arbitrary order approximation. For simplicity and since we were already obtaining a high accuracy, we decided to stop at a third-order approximation.

3.2. Projection

Projection methods take basis functions to build an approximated value function and decision rules that minimize a residual function defined by the augmented equilibrium conditions of the model. There are two popular methods for choosing basis functions: finite elements and the spectral method. We will present only the spectral method below. There are several reasons for this: first, in the neoclassical growth model the decision rules and value function are always smooth and spectral methods provide an excellent approximation (Aruoba, Fernández-Villaverde, and Rubio-Ramírez, 2006). Second, spectral methods allow us to use a large number of basis functions, with the consequent promise of high accuracy. Third, spectral methods are easier to implement. Their main drawback is that since they approximate the solution globally, if the decision rules display a rapidly changing local behavior or kinks, it may be extremely difficult for this scheme to capture those local properties.

Our target is to solve the decision rule for labor and the value function $\{l_t, V_t\}$ from the two conditions:

$$\mathcal{H}(l_t, V_t) = \begin{bmatrix} u_{c,t} - \beta \left(\mathbb{E}_t V_{t+1}^{1-\gamma}\right)^{\frac{1}{\theta}-1} \mathbb{E}_t \left[V_{t+1}^{\frac{(1-\gamma)(\theta-1)}{\theta}} u_{c,t+1} \left(\zeta e^{z_{t+1}} k_{t+1}^{\zeta-1} l_{t+1}^{1-\zeta} + 1 - \delta \right) \right] \\ V_t - \left[(1-\beta)(c_t^{\upsilon}(1-l_t^{\upsilon}))^{\frac{1-\gamma}{\theta}} - \beta \mathbb{E}_t (V_{t+1}^{1-\gamma})^{\frac{1}{\theta}} \right]^{\frac{\theta}{1-\gamma}} \end{bmatrix} = \mathbf{0}$$

where, to save on notation, we define $V_t = V(k_t, z_t)$ and:

$$u_{c,t} = \frac{1-\gamma}{\theta} \upsilon \frac{\left(c_t^{\upsilon}(1-l_t)^{1-\upsilon}\right)^{\frac{1-\gamma}{\theta}}}{c_t}$$

Then, from the static condition

$$c_t = \frac{\upsilon}{1-\upsilon}(1-\zeta) e^{z_t} k_t^{\zeta} l_t^{-\zeta} (1-l_t)$$

and the resource constraint, we can find c_t and k_{t+1}.

Spectral methods solve this problem by specifying the decision rule for labor and the value function $\{l_t, V_t\}$ as linear combinations of weighted basis functions:

$$\begin{aligned} l(k_t, z_j; \rho) &= \Sigma_i \rho_{ij}^l \psi_i(k_t) \\ V(k_t, z_j; \rho) &= \Sigma_i \rho_{ij}^V \psi_i(k_t) \end{aligned}$$

where $\{\psi_i(k)\}_{i=1,\ldots,n_k}$ are the n_k basis functions that we will use for our approximation along the capital dimension and $\rho = \{\rho_{ij}^l, \rho_{ij}^V\}_{i=1,\ldots,n_k; j=1,\ldots,N}$ are unknown coefficients to be deter-

mined. In this expression, we have discretized the stochastic process z_t for productivity using Tauchen's (1986) method with N points $G_z = \{z_1, z_2, ..., z_N\}$ and a transition matrix Π^N with generic element $\pi^N_{i,j} = Prob\,(z_{t+1} = z_j | z_t = z_i)$. Values for the decision rule outside the grid G_z can be approximated by interpolation. Since we set $N = 41$, the approximation is quite accurate along the productivity axis.

A common choice for the basis functions are Chebyshev polynomials because of their flexibility and convenience. Since their domain is [-1,1], we need to bound capital to the set $[\underline{k}, \overline{k}]$, where $\underline{k}$ ($\overline{k}$) is chosen sufficiently low (high) so that it will bind with an extremely low probability, and define a linear map from those bounds into [-1,1]. Then, we set $\psi_i(k_t) = \widetilde{\psi}_i(\phi_k(k_t))$ where $\widetilde{\psi}_i(\cdot)$ are Chebyshev polynomials and $\phi_k(k_t)$ is our linear mapping from $[\underline{k}, \overline{k}]$ to [-1,1].

By plugging $l(k_t, z_j; \rho)$ and $V(k_t, z_j; \rho)$ into $\mathcal{H}(l_t, V_t)$, we find the residual function:

$$\mathcal{R}(k_t, z_j; \rho) = \mathcal{H}(l(k_t, z_j; \rho), V(k_t, z_j; \rho))$$

We determine the coefficients ρ to get the residual function as close to $\mathbf{0}$ as possible. However, to do so, we need to choose a weight of the residual function over the space (k_t, z_j). Numerical analysts have determined that a collocation point criterion delivers the best trade-off between speed and accuracy (Fornberg, 1998).[8] Collocation simply makes the residual function exactly equal to zero in $\{k_i\}_{i=1}^{n_k}$ roots of the n_k-th order Chebyshev polynomial and in the Tauchen points $\{z_i\}_{i=1}^{Z}$. Therefore, we just need to solve the following system of $n_k \times N \times 2$ equations:

$$\mathcal{R}(k_i, z_j; \rho) = \mathbf{0} \text{ for any } i, j \text{ collocation points}$$

on $n_k \times N \times 2$ unknowns ρ. We solve this system with a Newton method and an iteration based on the increment of the number of basis functions. First, we solve a system with only three collocation points for capital and 41 points for technology. Then, we use that solution as a guess for a system with one more collocation point for capital (with the new coefficients being guessed to be equal to zero). We find a new solution and continue the procedure until we use up to 28 polynomials in the capital dimension (therefore, in the last step we solve for $2,296 = 28 \times 41 \times 2$ coefficients). The iteration is needed because otherwise the residual function is too cumbersome to allow for direct solution of the $2,296$ final coefficients.

[8] Also, the Chebyshev interpolation theorem tells us that if an approximating function is exact at the roots of the n_k-th order Chebyshev polynomial, then, as $n_k \to \infty$, the approximation error becomes arbitrarily small.

3.3. Value Function Iteration

Our final solution method is VFI. Since the dynamic algorithm is well known, our presentation is most brief. Consider the following Bellman operator:

$$TV\left(k_t,z_t\right)=\max_{c_t>0,0<l_t<1,k_{t+1}>0}\left[\left(1-\beta\right)\left(c_t^{\upsilon}\left(1-l_t\right)^{1-\upsilon}\right)^{\frac{1-\gamma}{\theta}}+\beta\left(\mathbb{E}_tV^{1-\gamma}\left(k_{t+1},z_{t+1}\right)\right)^{\frac{1}{\theta}}\right]^{\frac{\theta}{1-\gamma}}$$
$$\text{s.t. } c_t+k_{t+1}=e^{z_t}k_t^{\zeta}l_t^{1-\zeta}+(1-\delta)k_t$$
$$z_{t+1}=\lambda z_t+\sigma\varepsilon_{t+1}$$

To solve for this Bellman operator, we define a grid on capital, $G_k=\{k_1,k_2,...,k_M\}$ and a grid on the productivity level. The grid on capital is just a uniform distribution of points over the capital dimension. As we did for projection, we set a grid $G_z=\{z_1,z_2,...,z_N\}$ for productivity and a transition matrix Π^N using Tauchen's (1986) procedure. The algorithm to iterate on the value function for this grid is:

1. Set $n=0$ and $V^0(k_t,z_t)=c_{ss}^{\upsilon}\left(1-l_{ss}\right)^{1-\upsilon}$ for all $k_t\in G_k$ and all $z_t\in G_z$.

2. For every $\{k_t,z_t\}$, use Newton method to find c_t^*, l_t^*, k_{t+1}^* that solve:

$$c_t=\frac{\upsilon}{1-\upsilon}(1-\zeta)e^{z_t}k_t^{\zeta}l_t^{-\zeta}(1-l_t)$$
$$(1-\beta)\,\upsilon\frac{\left(c_t^{\upsilon}\left(1-l_t\right)^{1-\upsilon}\right)^{\frac{1-\gamma}{\theta}}}{c_t}=\beta\left(\mathbb{E}_t\left(V_{t+1}^n\right)^{1-\gamma}\right)^{\frac{1}{\theta}-1}\mathbb{E}_t\left[\left(V_{t+1}^n\right)^{-\gamma}V_{1,t+1}^n\right]$$
$$c_t+k_{t+1}=e^{z_t}k_t^{\zeta}l_t^{1-\zeta}+(1-\delta)k_t$$

3. Construct V^{n+1} from the Bellman equation:

$$V^{n+1}=((1-\beta)(c_t^{*\upsilon}(1-l_t^*)^{1-\upsilon})^{\frac{1-\gamma}{\theta}}+\beta(\mathbb{E}_t(V(k_{t+1}^*,z_{t+1})^{1-\gamma}))^{\frac{1}{\theta}})^{\frac{\theta}{1-\gamma}}$$

4. If $\frac{|V^{n+1}-V^n|}{|V^n|}\geq 1.0e^{-7}$, then $n=n+1$ and go to 2. Otherwise, stop.

To accelerate convergence and give VFI a fair chance, we implement a multigrid scheme as described by Chow and Tsitsiklis (1991). We start by iterating on a small grid. Then, after convergence, we add more points to the grid and recompute the Bellman operator using the previously found value function as an initial guess (with linear interpolation to fill the unknown values in the new grid points). Since the previous value function is an excellent grid, we quickly converge in the new grid. Repeating these steps several times, we move from

an initial 8,200-point grid into a final one with 123,000 points (3,000 points for capital and 41 for the productivity level).

4. Calibration

We now select a benchmark calibration for our numerical computations. We follow the literature as closely as possible. We set the discount factor $\beta = 0.991$ to generate an annual interest rate of around 3.6 percent. We set the parameter that governs labor supply, θ= 0.357, to get the representative household to work one-third of its time. The elasticity of output to capital, $\zeta = 0.4$, matches the labor share of national income. A value of the depreciation rate $\delta = 0.0196$ matches the ratio of investment-output. Finally, $\lambda = 0.95$ and σ= 0.007 are standard values for the stochastic properties of the Solow residual of the U.S. economy.

Table 1: Calibrated Parameters						
Parameter	β	υ	ζ	δ	λ	σ
Value	0.991	0.357	0.4	0.0196	0.95	0.007

Since we do not have a tight prior on the values for γ and ψ and we want to explore the dynamics of the model for a reasonable range of values, we select four values for the parameter that controls risk aversion, γ, 2, 5, 10, and 40, and two values for EIS ψ, 0.5, and 1.5, which bracket most of the values used in the literature (although many authors prefer smaller values for ψ, we found that the simulation results for smaller ψ's do not change much from the case when ψ= 0.5). We then compute the model for all eight combinations of values of γ and ψ, that is $\{2, 0.5\}$, $\{5, 0.5\}$, $\{10, 0.5\}$, and so on. When ψ= 0.5 and $\gamma = 2$, we are back in the standard CRRA case. However, in the interest of space, we will report only a limited subset of results, although we have selected those that we find the most interesting ones.

We pick as the benchmark case the calibration $\{\gamma, \psi, \sigma\} = \{5, 0.5, 0.007\}$. These values reflect an EIS centered around the median of the estimates in the literature, a reasonably high level of risk aversion, and the observed volatility of productivity shocks. To check robustness, we increase, in the extreme case, the risk aversion and standard deviation of the productivity shock to $\{\gamma, \psi, \sigma\} = \{40, 0.5, 0.035\}$. This case combines levels of risk aversion that are in the upper bound of all estimates in the literature with huge productivity shocks. Therefore, it pushes all solution methods to their limits, in particular, making life hard for perturbation since the interaction of large precautionary behavior induced by γ and large shocks will move the economy far away from the deterministic steady state. We leave the discussion of the effects of $\psi = 1.5$ for the robustness analysis at the end of the next section.

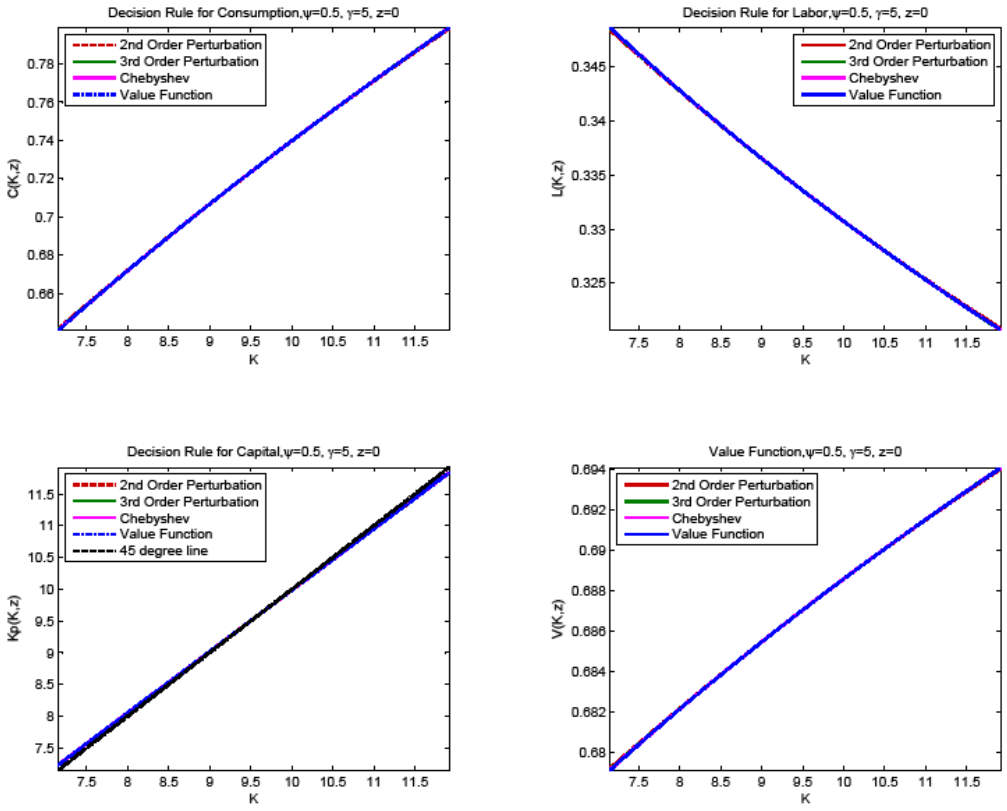

Figure 1: Decision Rules and Value Function, benchmark case

5. Numerical Results

In this section we report our numerical findings. First, we present and discuss the computed decision rules. Second, we show the results of simulating the model. Third, we report the Euler equation errors as proposed by Judd (1992) and Judd and Guu (1997). Fourth, we discuss the effects of changing the EIS and the perturbation point. Finally, we discuss implementation and computing time.

5.1. Decision Rules

One of our first results is the decision rules and the value function of the agent. Figure 1 plots the decision rules for consumption, labor supply, capital, and the value function in the benchmark case when $z = 0$ over a capital interval centered on the steady-state level of capital of 9.54 with a width of $\pm 25\%$, [7.16,11.93]. We selected an interval for capital big enough to encompass all the simulations in our sample. Similar figures could be plotted for other values of z. We omit them because of space considerations.

Since all methods provide nearly indistinguishable answers, we observe only one line in all figures. It is possible to appreciate very tiny differences in labor supply between second-order perturbation and the other methods only when capital is far from its steady-state level. Monotonicity of the decision rules is preserved by all methods. We must be cautious, however, mapping differences in choices into differences in utility. The Euler error function below provides a better view of the welfare consequences of different approximations.

We see bigger differences in the decision rules and value functions as we increase the risk aversion and variance of the shock. Figure 2 plots the decision rules and value functions for the extreme calibration. In this figure, we change the interval reported because, owing to the high variance of the calibration, the equilibrium paths fluctuate through much wider ranges of capital.

We highlight several results. First, all the methods deliver similar results in the range of [7.16,11.93], our original interval for the benchmark calibration. Second, as we go far away from the steady state, VFI and the Chebyshev polynomial still overlap with each other but second- and third-order approximations start to deviate. Third, the decision rule for consumption and the value function approximated by the third-order perturbation changes from concavity into convexity for values of capital bigger than 15. This phenomenon is in line with the evidence documented in Aruoba, Fernández-Villaverde, and Rubio-Ramírez (2006) and it is due to the poor performance of local approximation when we move too far away from the expansion point and the polynomials begin to behave wildly. In any case, this issue is irrelevant because, as we will show below, the problematic region is visited with nearly zero probability

5.2. Simulations

Applied economists often characterize the behavior of the model through statistics from simulated paths of the economy. We simulate the model, starting from the deterministic steady state, for 10,000 periods, using the decision rules for each of the eight combinations of risk aversion and EIS discussed above. To make the comparison meaningful, the shocks are common across all paths. We discard the first 1,000 periods as a burn-in. The burn-in is important because it eliminates the transition from the deterministic steady state of the model to the middle regions of the ergodic distribution of capital. This is usually achieved in many fewer periods than the ones in our burn-in, but we want to be conservative in our results. The remaining observations constitute a sample from the ergodic distribution of the economy.

For the benchmark calibration, the simulations from all of the solution methods generate

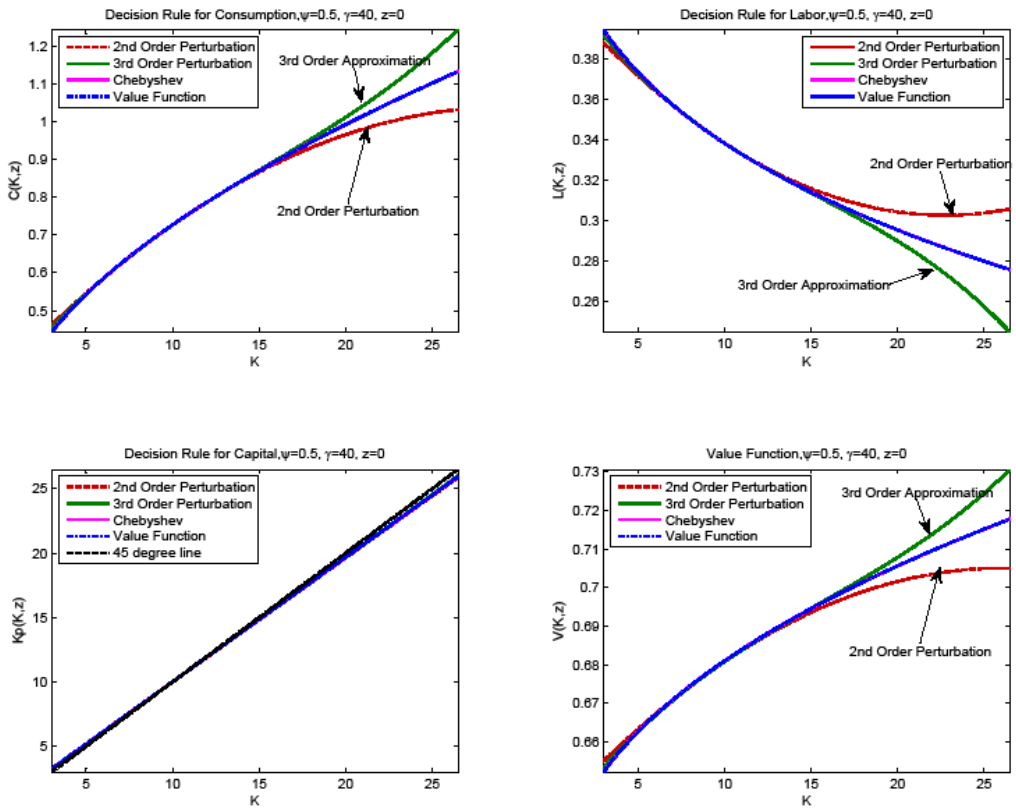

Figure 2: Decision Rules and Value Function, extreme case

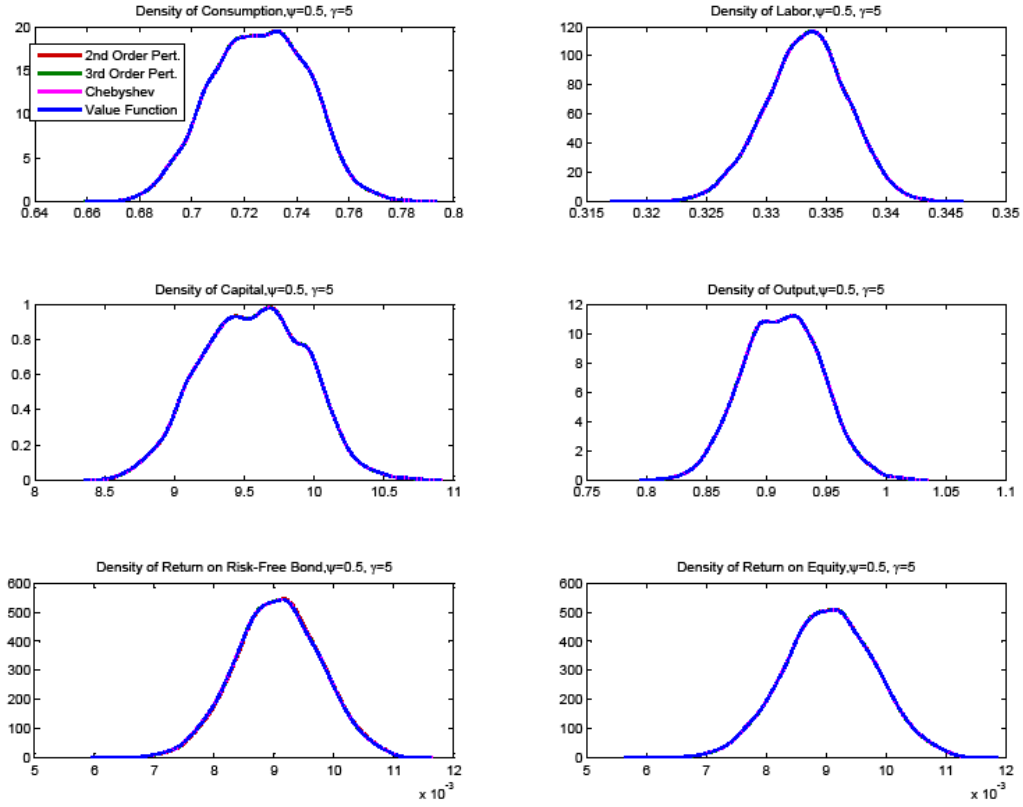

Figure 3: Densities, benchmark case

almost identical equilibrium paths (and therefore we do not report them). We focus instead on the densities of the endogenous variables as shown in figure 3. Given the relatively low level of risk aversion and variance of the productivity shocks, all densities are roughly centered around the deterministic steady state value of the variable. For example, the mean of the distribution of capital is only 0.3 percent higher than the deterministic value. Note that capital is nearly always between 8.5 and 10.5. This range will be important below to judge the accuracy of our approximations.

Table 2 reports business cycle statistics and, because DSGE models with recursive preferences are often used for asset pricing, the average and variance of the (quarterly) risk-free rate and return on capital. Again, we see that nearly all values are the same, a simple consequence of the similarity of the decision rules.

Table 2: Business Cycle Statistics - Benchmark Calibration					
	c	y	i	$R^f(\%)$	$R^k(\%)$
	Mean				
Second-Order Perturbation	0.7259	0.9134	0.1875	0.9069	0.9060
Third-Order Perturbation	0.7259	0.9134	0.1875	0.9057	0.9060
Chebyshev Polynomial	0.7259	0.9134	0.1875	0.9058	0.9060
Value Function Iteration	0.7259	0.9134	0.1875	0.9058	0.9061
	Variance (%)				
Second-Order Perturbation	0.0326	0.1036	0.0284	0.0001	0.0001
Third-Order Perturbation	0.0327	0.1036	0.0285	0.0001	0.0001
Chebyshev Polynomial	0.0327	0.1039	0.0286	0.0001	0.0001
Value Function Iteration	0.0327	0.1039	0.0286	0.0001	0.0001

The welfare cost of the business cycle is reported in Table 3 in consumption equivalent terms. The computed costs are actually negative. This is due to the fact that when we have leisure in the utility function, the indirect utility function may be convex in input prices (agents change their behavior over time by a large amount to take advantage of changing productivity). Cho and Cooley (2000) present a similar example and offer further discussion. Welfare costs are comparable across methods. Remember that the welfare cost of the business cycle for the second- and third-order perturbations is the same because the third-order terms all drop or are zero when evaluated at the steady state.

Table 3: Welfare Costs of Business Cycle - Benchmark Calibration			
2nd-Order Pert.	3rd-Order Pert.	Chebyshev	Value Function
-2.0864(e-5)	-2.0864(e-5)	-2.4406e(-5)	-2.8803e(-6)

When we move to the extreme calibration, we see more differences. Figure 4 plots the histograms of the simulated series for each solution method. Looking at quantities, the histograms of consumption, output, and labor are the same across all of the methods. The histogram of capital produced using second-order perturbation is more centered around the mean of the distribution. The ergodic distribution of capital puts nearly all the mass between values of 6 and 15. This considerable move to the right in comparison with figure 3 is due to the effect of precautionary behavior in the presence of large productivity shocks and high risk aversion. Capital also goes down more than in the benchmark calibration because of large, persistent productivity shocks, but not nearly as much as it increases when shocks are positive. With respect to prices, as expected, the distribution implied by the second-order approximation is more concentrated than the other three.

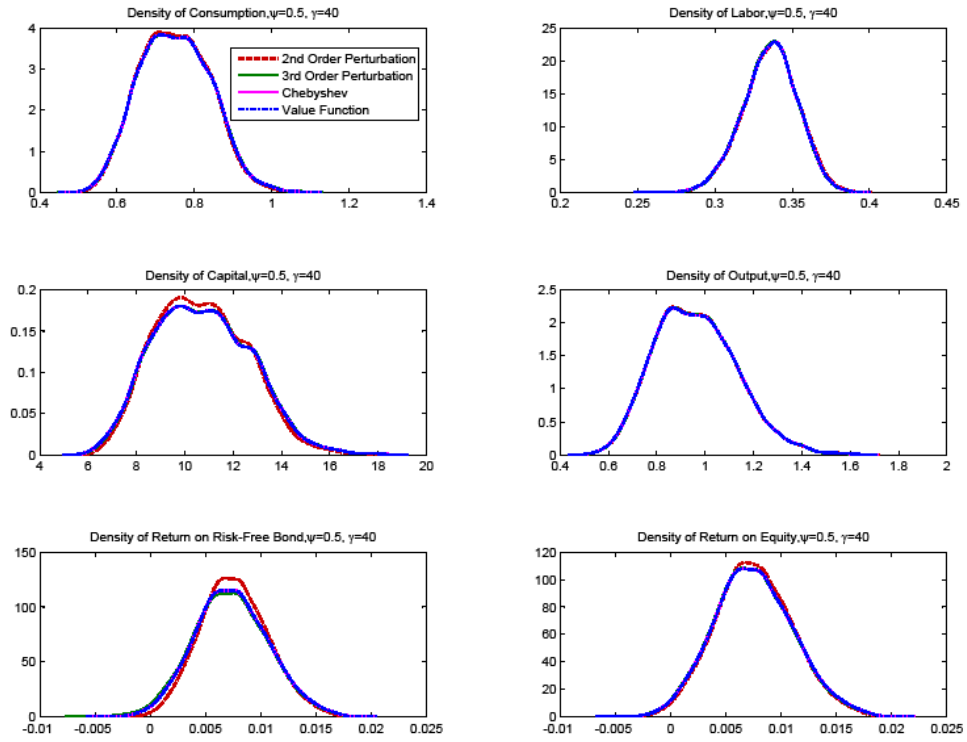

Figure 4: Densities, extreme calibration

Table 4 reports business cycle statistics. Differences across methods are minor. Looking at prices, the mean of the risk-free rate from the second-order perturbation is larger compared to the means produced by the other methods.[9] Importantly, the second-order perturbation produces a negative excess return, which is positive (although small) for all other methods. This suggests that a second-order perturbation may not be good enough if we care about the asset pricing properties of our model and that we face high variance of the shocks and/or high risk aversion. A third-order perturbation, in comparison, eliminates most of the differences and delivers nearly the same asset pricing implications as Chebyshev polynomials or VFI.

Table 4 Business Cycle Statistics - Extreme Calibration					
	Mean				
	c	y	i	$R^f(\%)$	$R^k(\%)$
Second-Order Perturbation	0.7497	0.9616	0.2099	0.7610	0.7484
Third-Order Perturbation	0.7510	0.9615	0.2109	0.7281	0.7519
Chebyshev Polynomial	0.7506	0.9612	0.2106	0.7404	0.7556
Value Function Iteration	0.7506	0.9612	0.2106	0.7404	0.7556
	Variance (%)				
Second-Order Perturbation	0.7985	2.8463	0.7803	0.0009	0.0012
Third-Order Perturbation	0.8387	2.8462	0.8447	0.0012	0.0013
Chebyshev Polynomial	0.8354	2.8643	0.8520	0.0011	0.0013
Value Function Iteration	0.8354	2.8643	0.8520	0.0011	0.0013

Finally, Table 5 presents the welfare cost of the business cycle. Now, in comparison with the benchmark calibration, the welfare cost of the business cycle is positive and significant, slightly higher than 3 percent. This is not a surprise, since we have both a large risk aversion and productivity shocks with a standard deviation five times as big as the observed one. All methods deliver numbers that are close.

Table 5: Welfare Costs of Business Cycle - Extreme Calibration			
2nd-Order Pert.	3rd-Order Pert.	Chebyshev	Value Function
3.1127e(-2)	3.1127e(-2)	3.3034e(-2)	3.3032e(-2)

5.3. Euler Equation Errors

While the plots of the decision rules and the computation of densities and business cycle statistics that we presented in the previous subsection are highly informative, it is also important

[9] The mean of the returns is now much smaller than in the benchmark calibration because precautionary savings raise average capital.

to evaluate the accuracy of each of the procedures. Euler equation errors, introduced by Judd (1992), have become a common tool for determining the quality of the solution method. The idea is to observe that, in our model, the intertemporal condition:

$$u_{c,t} = \beta(\mathbb{E}_t V_{t+1}^{1-\gamma})^{\frac{1}{\theta}-1}\mathbb{E}_t\left(V_{t+1}^{\frac{(\gamma-1)(1-\theta)}{\theta}} u_{c,t+1} R\left(k_t, z_t; z_{t+1}\right)\right) \tag{4}$$

where $R\left(k_t, z_t; z_{t+1}\right) = 1 + \zeta e^{z_{t+1}} k_{t+1}^{\zeta-1} l_{t+1}^{1-\zeta} - \delta$ is the gross return of capital given states k_t, z_t and realization z_{t+1} should hold exactly for any given k_t, and z_t. However, since the solution methods we use are only approximations, there will be an error in (4) when we plug in the computed decision rules. This Euler equation error function $EE^i\left(k_t, z_t\right)$ is defined, in consumption terms:

$$EE^i\left(k_t, z_t\right) = 1 - \frac{\left[\frac{\beta(\mathbb{E}_t\left(V_{t+1}^i\right)^{1-\gamma})^{\frac{1}{\theta}-1}\mathbb{E}_t\left(\left(V_{t+1}^i\right)^{\frac{(\gamma-1)(1-\theta)}{\theta}} u_{c,t+1}^i R(k_t,z_t;z_{t+1})\right)}{\frac{1-\gamma}{\theta}\upsilon(1-l_t^i)^{(1-\upsilon)\frac{1-\gamma}{\theta}}}\right]^{\frac{1}{\upsilon\frac{1-\gamma}{\theta}-1}}}{c_t^i}$$

This function determines the (unit free) error in the Euler equation as a fraction of the consumption given the current states k_t, and z_t and solution method i. Following Judd and Guu (1997), we can interpret this error as the optimization error incurred by the use of the approximated decision rule and we report the absolute errors in base 10 logarithms to ease interpretation. Thus, a value of -3 means a \$1 mistake for each \$1000 spent, a value of -4 a \$1 mistake for each \$10,000 spent, and so on.

Figure 5 displays a transversal cut of the errors for the benchmark calibration when $z = 0$. Other transversal cuts at different technology levels reveal similar patterns. The first lesson from figure 5 is that all methods deliver high accuracy. We know from figure 3 that capital is nearly always between 8.5 and 10.5. In that range, the (log10) Euler equation errors are at most -5, and most of the time they are even smaller. For instance, the second- and third-order perturbations have an Euler equation error of around -7 in the neighborhood of the deterministic steady state, VFI of around -6, and Chebyshev an impressive -12/-13. The second lesson from figure 5 is that, as expected, global methods (Chebyshev and VFI) perform very well in the whole range of capital values, while perturbations deteriorate as we move away from the deterministic steady state. For second-order perturbation, the Euler error in the steady state is almost four orders of magnitudes smaller than on the boundaries. Third-order perturbation is around half an order of magnitude more accurate than second-order perturbation over the whole range of values (except in a very small region close to the deterministic steady state).

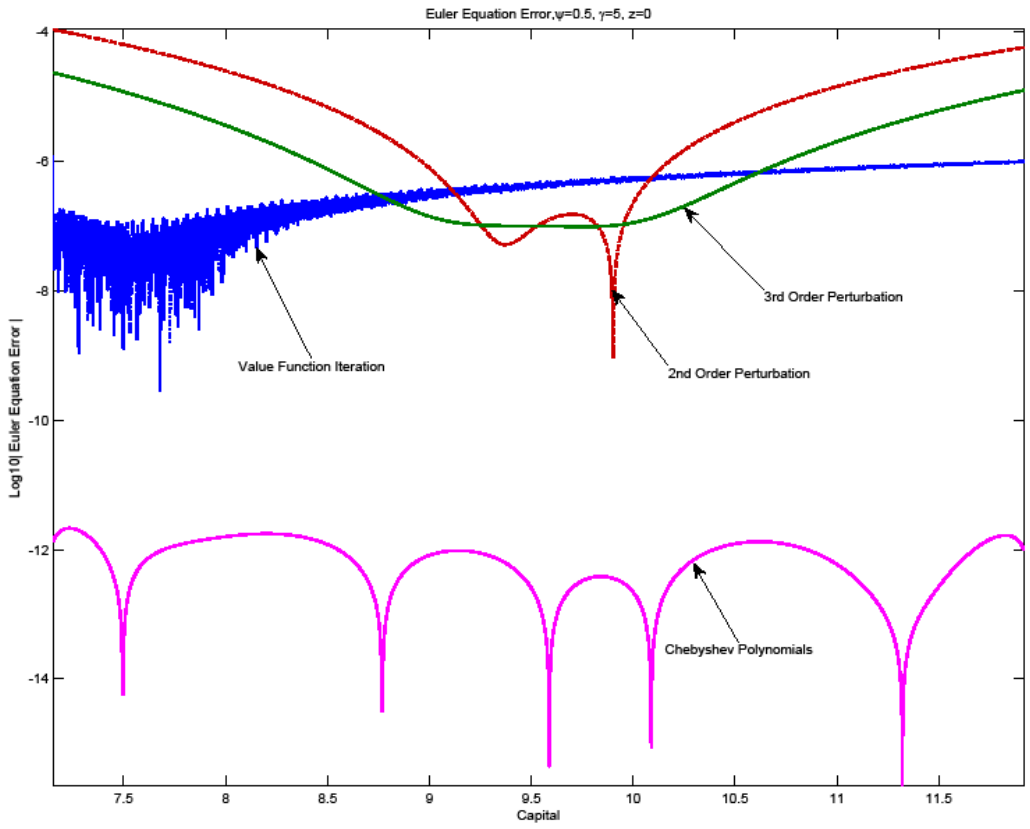

Figure 5: Euler Equation Error, benchmark calibration

Table 6 Euler errors - Benchmark Calibration		
	Max Euler Error	Integral of the Euler Errors
Second-Order Perturbation	-3.1421	-6.4360
Third-Order Perturbation	-3.2448	-6.9576
Chebyshev Polynomial	-11.2146	-12.4711
Value Function Iteration	-5.6908	-6.4246

There are two complementary ways to summarize the information from Euler equation error functions. First, we report the maximum error in our interval (capital between 75 percent and 125 percent of the steady state and the 41 points of the productivity level) in the second column of table 6. The maximum Euler error is useful because it bounds the mistake owing to the approximation. Both perturbations have a maximum Euler error of around -3.2, VFI of -5.7, and Chebyshev, an impressive -11.2. We read this table as suggesting that all methods perform more than acceptably. The second procedure for summarizing Euler equation errors is to integrate the function with respect to the ergodic distribution of capital

and productivity to find the average error.[10] We can think of this exercise as a generalization of the Den Haan–Marcet test (Den Haan and Marcet, 1994). This integral is a welfare measure of the loss induced by the use of the approximating method. We report our results in the third column of table 6. Now, both perturbations and VFI have roughly the same performance (indeed, the third-order perturbation does better than VFI), while Chebyshev polynomials do fantastically well (the average loss of welfare is \$1 for each \$5 trillion, more than a third of U.S. output). But even an approximation with an average error of -6.96, such as the one implied by third-order perturbation must suffice for most relevant applications.

We repeat our exercise for the extreme calibration. Figure 6 displays the results for the extreme case. Again, we have changed the capital interval to make it representative of the behavior of the model in the ergodic distribution. Now, perturbations deteriorate more as we get further away from the deterministic steady state. However, in the relevant range of values of capital of [6, 17], we still have Euler equation errors smaller than -3. The performance of VFI and Chebyshev polynomials is roughly the same as in our benchmark calibration, although, since we extend the range of capital, the performance deteriorates around two orders of magnitude.

Table 7 reports maximum Euler equation errors and their integrals. The maximum Euler equation error is large for perturbation methods while it is remarkably small using Chebyshev polynomials. However, given the very large range of capital used in the computation, this maximum Euler error provides a too negative view of accuracy. We find the integral of the Euler equation error to be much more informative. With a second-order perturbation, we have -3.85, which is on the high side, but with a third-order perturbation we have -5, which is acceptable in most computations. Note that this integral is computed when we have extremely high risk aversion and large productivity shocks. Even in this challenging environment, a third-order perturbation delivers a high degree of accuracy. VFI and Chebyshev do not display a big loss of precision compared to the benchmark case, and in the case of Chebyshev polynomials, the performance is still outstanding.

[10] There is the technical consideration of which ergodic distribution to use for this task, since this is an object that can only be found by simulation. We use the ergodic simulation generated by VFI, which slightly favors this method over the other ones. However, we checked that the results are totally robust to using the ergodic distributions coming from the other methods.

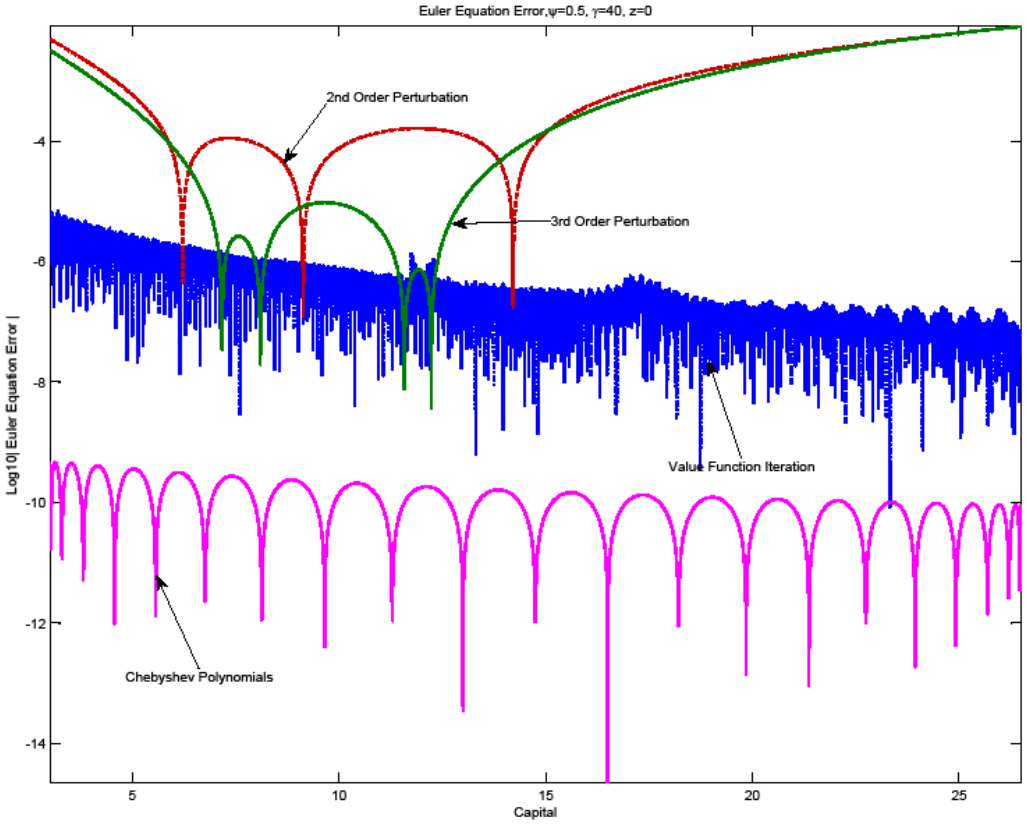

Figure 6: Euler Equation Errors, extreme calibration

Table 7 Euler errors - Extreme Calibration		
	Max Euler Error	Integral of the Euler Errors
Second-Order Perturbation	-1.6905	-3.8544
Third-Order Perturbation	-1.8598	-5.0616
Chebyshev Polynomial	-8.9540	-10.1028
Value Function Iteration	-4.7429	-6.5300

5.4. Robustness: Changing the EIS and Changing the Perturbation Point

In the results we reported above, we kept the EIS equal to 0.5, a conventional value in the literature, while we modified the risk aversion and the volatility of productivity shocks. However, since some recent papers prefer higher values of the EIS (see, for instance, Bansal and Yaron, 2004), we also computed our model with $\psi = 1.5$. Basically our results were unchanged. To save on space, we concentrate only on the Euler equation errors (decision rules and simulation paths are available upon request). In table 8, we report the maxima of the Euler equation errors and their integrals with respect to the ergodic distribution. The relative size and values of the entries of this table are quite similar to the entries in table 6 (except, partially, VFI that performs a bit better).

Table 8 Euler errors - Benchmark Calibration with $\psi = 1.5$		
	Max Euler Error	Integral of the Euler Errors
Second-Order Perturbation	-3.1536	-6.4058
Third-Order Perturbation	-3.2362	-6.8470
Chebyshev Polynomial	-11.9230	-12.4091
Value Function Iteration	-6.4694	-7.8523

Table 9 repeats the same exercise for the extreme calibration with high risk aversion and high volatility of productivity shocks. Again, the entries on the table are very close to the ones in table 7 (and now, VFI does not perform better than when $\psi = 0.5$).

Table 9 Euler errors - Extreme Calibration with $\psi = 1.5$		
	Max Euler Error	Integral of the Euler Errors
Second-Order Perturbation	-1.9532	-3.6788
Third-Order Perturbation	-2.1848	-4.9156
Chebyshev Polynomial	-8.4452	-9.6057
Value Function Iteration	-4.8484	-6.4759

As a final robustness test, we computed the perturbations not around the deterministic steady state but around a point close to the mode of the ergodic distribution of capital. This strategy could deliver better accuracy because we approximate the value function and decision rules in a region where the model spends more time. Disappointingly, we found only trivial improvements in terms of accuracy (for instance, the Euler equation errors improved by less than 1 percent). Moreover, expanding at a point different from the deterministic steady state has the disadvantage that the theorems that ensure the convergence of the Taylor approximation might fail (see theorem 6 in Jin and Judd, 2002).

5.5. Implementation and Computing Time

We briefly discuss implementation and computing time. For the benchmark calibration, second-order perturbation and third- order perturbation algorithms take only 0.1 second and 0.3 second, respectively, in a 2.2GHz Intel PC running Windows Vista (the reference computer for all times below), and it is simple to implement (732 lines of code in Fortran 95 for second order and 1492 lines of code for third order).[11] Although the number of lines doubles in the third order, the complexity in terms of coding does not increase much: the extra lines are mainly from declaring external functions and reading and assigning values to the perturbation coefficients. Fortran 95 borrows the analytical derivatives of the equilibrium conditions from a code written in `Mathematica 6`. This code has between 150 to 210 lines, although `Mathematica` is much less verbose. An interesting observation is that we only need to take the analytic derivatives once, since they are expressed in terms of parameters and not in terms of parameter values. This allows Fortran to evaluate the analytic derivatives extremely fast for new combinations of parameter values. This advantage of perturbation is particularly relevant when we need to solve the model repeatedly for many different parameter values, for example, when we are estimating the model. For completeness, the second-order perturbation was also run in `Dynare` (although we had to use version 4.0, which computes analytic derivatives, instead of previous versions, which use numerical derivatives that are not accurate enough for perturbation). This run was a double-check of the code and a test of the feasibility of using off-the-shelf software to solve DSGE models with recursive preferences.

The projection algorithm takes less than 30 seconds, but it requires a good initial guess for the solution of the system of equations. Finding the initial guess for some combination of parameter values proved to be challenging. The code is 695 lines of Fortran 95. Finally, the VFI code is 673 lines of Fortran 95, but it takes about 2 hours to run.

[11] We use lines of code as a proxy for the complexity of implementing the code. We do not count comment lines.

6. Conclusions

In this paper, we have compared different solution methods for DSGE models with recursive preferences. We evaluated the different algorithms based on accuracy, speed, and programming burden. We learned that all of the most promising methods (perturbation, projection, and VFI) do a fair job in terms of accuracy. We were surprised by how well simple second-order and third-order perturbations perform. For an extreme calibration, a second-order perturbation suffered somewhat in terms of accuracy, in particular regarding asset prices, but a third order perturbation still held its place. We were impressed by how accurate Chebyshev polynomials can be, even in highly challenging calibrations. However, their computational cost was higher and we are concerned about the curse of dimensionality. In any case, it seems clear to us that, when accuracy is the key consideration, Chebyshev polynomials are the way to go. Finally, we were disappointed by VFI since even with 123,000 points in the grid, it still could not beat perturbation and it performed much worse than Chebyshev polynomials. This suggests that unless there are compelling reasons such as non-differentiabilities or non-convexities in the model, we better avoid VFI.

A theme we have not developed in this paper is the possibility of interplay among different solution methods. For instance, we can compute extremely easily a second-order approximation to the value function and use it as an initial guess for VFI. This second-order approximation is such a good guess that VFI will converge in few iterations. We verified this idea in non-reported experiments, where VFI took one-tenth of the time to converge once we used the second-order approximation to the value function as the initial guess. This approach may even work when the true value function is not differentiable at some points or has jumps, since the only goal of perturbation is to provide a good starting point, not a theoretically sound approximation. This algorithm may be particularly useful in problems with many state variables. More research in this type of hybrid method is a natural extension of our work.

We close the paper by pointing out that recursive preferences are only one example of a large class of non-standard preferences that have received much attention by theorists and applied researchers over the last years (see the review of Backus, Routledge, and Zin, 2004). Having fast and reliable solution methods for this class of new preferences will help researchers to sort out which of these preferences deserve further attention and to derive empirical implications. Thus, this paper is a first step in the task of learning how to compute DSGE models with non-standard preferences.

7. Appendix

In this appendix, we present the steady state of the model and the alternative perturbation approach, the value function perturbation (VFP).

7.1. Steady State of the Model

To solve the system:

$$
\begin{aligned}
V_{ss} &= c_{ss}^{\upsilon}\left(1-l_{ss}\right)^{1-\upsilon} \\
\left(\zeta k_{ss}^{\zeta-1} l_{ss}^{1-\zeta}+1-\delta\right) &= 1/\beta \\
\frac{1-\nu}{\nu}\frac{c_{ss}}{(1-l_{ss})} &= (1-\zeta)k_{ss}^{\zeta} l_{ss}^{-\zeta} \\
m_{ss}R_{ss}^{f} &= 1/\beta \\
c_{ss}+i_{ss} &= k_{ss}^{\zeta} l_{ss}^{1-\zeta} \\
i_{ss} &= \delta k_{ss}
\end{aligned}
$$

note first that:

$$
\frac{k_{ss}}{l_{ss}} = \left(\frac{1}{\zeta}\left(\frac{1}{\beta}-1+\delta\right)\right)^{\frac{1}{\zeta-1}} = \Omega
$$

Now, from the leisure-consumption condition:

$$
\frac{c_{ss}}{1-l_{ss}} = \frac{\upsilon}{1-\upsilon}\left(1-\zeta\right)\Omega^{\zeta} = \Phi \Rightarrow c_{ss} = \Phi\left(1-l_{ss}\right)
$$

Then:

$$
c_{ss}+\delta k_{ss} = k_{ss}^{\zeta} l_{ss}^{1-\zeta} = \Omega^{\zeta} l_{ss} \Rightarrow c_{ss} = \left(\Omega^{\zeta}-\delta\Omega\right) l_{ss}
$$

and:

$$
\begin{aligned}
\Phi\left(1-l_{ss}\right) &= \left(\Omega^{\zeta}-\delta\Omega\right) l_{ss} \Rightarrow \\
l_{ss} &= \frac{\Phi}{\Omega^{\zeta}-\delta\Omega+\Phi} \\
k_{ss} &= \frac{\Phi\Omega}{\Omega^{\zeta}-\delta\Omega+\Phi}
\end{aligned}
$$

from which we can find V_{ss} and i_{ss}.

7.2. Value Function Perturbation (VFP)

We mentioned in the main text that instead of perturbing the equilibrium conditions of the model, we could directly perturb the value function in what we called value function perturbation (VFP). To undertake the VFP, we write the value function as:

$$V\left(k_t, z_t; \chi\right) = \max_{c_t, l_t} \left[\begin{array}{c} \left(1-\beta\right)\left(c_t^{\upsilon}\left(1-l_t\right)^{1-\upsilon}\right)^{\frac{1-\gamma}{\theta}} \\ +\beta\left(\mathbb{E}_t V^{1-\gamma}\left(e^{z_t}k_t^{\zeta}l_t^{1-\zeta} + \left(1-\delta\right)k_t - c_t, \lambda z_{t-1} + \chi\sigma\varepsilon_t; \chi\right)\right)^{\frac{1}{\theta}} \end{array} \right]^{\frac{\theta}{1-\gamma}}$$

To find a second-order approximation to the value function, we take derivatives of the value function with respect to controls (c_t, l_t), states (k_t, z_t), and the perturbation parameter χ.

$*$Derivative with respect to c_t:

$$\left(1-\beta\right)\upsilon\frac{\left(c_t^{\upsilon}\left(1-l_t\right)^{1-\upsilon}\right)^{\frac{1-\gamma}{\theta}}}{c_t} - \beta\left(\mathbb{E}_t V_{t+1}^{1-\gamma}\right)^{\frac{1}{\theta}-1}\mathbb{E}_t\left[V_{t+1}^{-\gamma}V_{1,t+1}\right] = 0$$

where we have used the notation $V_t = V\left(k_t, z_t; \chi\right)$ (and the analogous notation for partial derivatives).

$*$Derivative with respect to l_t:

$$-\left(1-\beta\right)\left(1-\upsilon\right)\frac{\left(c_t^{\upsilon}\left(1-l_t\right)^{1-\upsilon}\right)^{\frac{1-\gamma}{\theta}}}{1-l_t} + \beta\left(\mathbb{E}_t V_{t+1}^{1-\gamma}\right)^{\frac{1}{\theta}-1}\mathbb{E}_t\left[V_{t+1}^{-\gamma}V_{1,t+1}\left(1-\zeta\right)e^{z_t}k_t^{\zeta}l_t^{-\zeta}\right] = 0,$$

$*$Derivative with respect to k_t:

$$V_{1,t} = V_t^{1-\frac{1-\gamma}{\theta}}\left[\beta\left(\mathbb{E}_t\left[V_{t+1}^{1-\gamma}\right]\right)^{\frac{1}{\theta}-1}\mathbb{E}_t\left[V_{t+1}^{-\gamma}V_{1,t+1}\right]\left(\zeta e^{z_t}k_t^{\zeta-1}l_t^{1-\zeta} + 1 - \delta\right)\right].$$

$*$Derivative with respect to z_t:

$$V_{2,t} = V_t^{1-\frac{1-\gamma}{\theta}}\left[\beta\left(\mathbb{E}_t\left[V_{t+1}^{1-\gamma}\right]\right)^{\frac{1}{\theta}-1}\mathbb{E}_t\left[V_{t+1}^{-\gamma}\left(V_{1,t+1}e^{z_t}k_t^{\zeta}l_t^{1-\zeta} + V_{2,t+1}\lambda\right)\right]\right].$$

$*$Derivative with respect to χ:

$$V_{3,t} = V_t^{1-\frac{1-\gamma}{\theta}}\left[\beta\left(\mathbb{E}_t\left[V_{t+1}^{1-\gamma}\right]\right)^{\frac{1}{\theta}-1}\mathbb{E}_t\left[V_{t+1}^{-\gamma}\left(V_{2,t+1}\chi\sigma\varepsilon_{t+1} + V_{3,t+1}\right)\right]\right].$$

In the last three equations, we apply the envelope theorem to eliminate the derivatives of the value function with respect to c_t, and hence the derivatives of consumption with respect to

k, z, and χ.

We collect all the equations:

$$c_t + k_{t+1} = e^{z_t} k_t^{\zeta} l_t^{1-\zeta} + (1-\delta) k_t$$

$$V_t = \left[(1-\beta) \left(c_t^{\upsilon} (1-l_t)^{1-\upsilon} \right)^{\frac{1-\gamma}{\theta}} + \beta \left(\mathbb{E}_t \left[V_{t+1}^{1-\gamma} \right] \right)^{\frac{1}{\theta}} \right]^{\frac{\theta}{1-\gamma}}$$

$$(1-\beta)\, \upsilon \frac{\left(c_t^{\upsilon} (1-l_t)^{1-\upsilon} \right)^{\frac{1-\gamma}{\theta}}}{c_t} - \beta \left(\mathbb{E}_t V_{t+1}^{1-\gamma} \right)^{\frac{1}{\theta}-1} \mathbb{E}_t \left[V_{t+1}^{-\gamma} V_{1,t+1} \right] = 0$$

$$-(1-\beta)(1-\upsilon) \frac{\left(c_t^{\upsilon} (1-l_t)^{1-\upsilon} \right)^{\frac{1-\gamma}{\theta}}}{1-l_t} + \beta \left(\mathbb{E}_t V_{t+1}^{1-\gamma} \right)^{\frac{1}{\theta}-1} \mathbb{E}_t \left[V_{t+1}^{-\gamma} V_{1,t+1} (1-\zeta) e^{z_t} k_t^{\zeta} l_t^{-\zeta} \right] = 0$$

$$V_1(k_t, z_t; \chi) = V_t^{1-\frac{1-\gamma}{\theta}} \left[\beta \left(\mathbb{E}_t \left[V_{t+1}^{1-\gamma} \right] \right)^{\frac{1}{\theta}-1} \mathbb{E}_t \left[V_{t+1}^{-\gamma} V_{1,t+1} \right] \left(\zeta e^{z_t} k_t^{\zeta-1} l_t^{1-\zeta} + 1 - \delta \right) \right]$$

$$V_2(k_t, z_t; \chi) = V_t^{1-\frac{1-\gamma}{\theta}} \left[\beta \left(\mathbb{E}_t \left[V_{t+1}^{1-\gamma} \right] \right)^{\frac{1}{\theta}-1} \mathbb{E}_t \left[V_{t+1}^{-\gamma} \left(V_{1,t+1} e^{z_t} k_t^{\zeta} l_t^{1-\zeta} + V_{2,t+1} \lambda \right) \right] \right]$$

$$V_3(k_t, z_t; \chi) = V_t^{1-\frac{1-\gamma}{\theta}} \left[\beta \left(\mathbb{E}_t \left[V_{t+1}^{1-\gamma} \right] \right)^{\frac{1}{\theta}-1} \mathbb{E}_t \left[V_{t+1}^{-\gamma} \left(V_{2,t+1} \chi \sigma \varepsilon_{t+1} + V_{3,t+1} \right) \right] \right]$$

$$z_t = \lambda z_{t-1} + \chi \sigma \varepsilon_t$$

or, in the more compact notation of the main text

$$\widetilde{F}(k_t, z_t, \chi) = \mathbf{0}$$

where the hat over F emphasizes that now we are dealing with a slightly different set of equations than the F in the main text.

The steady state of this system is the same as the one in the previous subsection of this appendix, except that now we have to compute three more objects from the equations:

$$(1-\beta)\, \upsilon \frac{V_{ss}^{\frac{1-\gamma}{\theta}}}{c_{ss}} - \beta V_{ss}^{\frac{1-\gamma}{\theta}-1} V_{1,ss} = 0$$

$$V_{2,ss} = \beta \left[V_{1,ss} k_{ss}^{\zeta} l_{ss}^{1-\zeta} + V_{2,ss} \lambda \right]$$

$$V_{3,ss} = \beta V_{3,ss}$$

With some algebra

$$V_{1,ss} = \frac{1-\beta}{\beta} \upsilon \frac{V_{ss}}{c_{ss}}$$

$$V_{2,ss} = \frac{\beta}{1-\beta\lambda} V_{1,ss} k_{ss}^{\zeta} l_{ss}^{1-\zeta}$$

$$V_{3,ss} = 0$$

that are the terms we need to build the first-order approximation of the value function:

$$V\left(k_t, z_t; \chi\right) \simeq V_{ss} + V_{1,ss}\left(k_t - k_{ss}\right) + V_{2,ss} z_t + V_{3,ss}\chi$$

Now, as we did with ECP, we take derivatives of the function $\widetilde{F}$ with respect to k_t, z_t, and χ

$$\widetilde{F}_i\left(k_{ss}, 0; 0\right) = \mathbf{0} \text{ for } i = \{1, 2, 3\}$$

and we solve for the unknown coefficients. This solution will give us a second-order approximation of the value function but only a first-order approximation of the decision rules. By repeating these steps n times, we can obtain the $n+1$-order approximation of the value function and the n-order approximation of the decision rules. It is straightforward to check that the coefficients obtained by ECP and VFP are the same. Thus, the choice for one approach or the other should be dictated by expediency.

References

[1] Anderson, E.W., L.P. Hansen, E.R. McGrattan, and T.J. Sargent (1996). "On the Mechanics of Forming and Estimating Dynamic Linear Economies," in H.M. Amman *et al.* (eds.) *Handbook of Computational Economics*, Elsevier.

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

[3] Backus, D.K., B.R. Routledge, and S.E. Zin (2004). "Exotic Preferences for Macroeconomists." *NBER Macroeconomics Annual* 2004, 319-390.

[4] Backus, D.K., B.R. Routledge, and S.E. Zin (2007). "Asset Pricing in Business Cycle Analysis." *Mimeo*, New York University.

[5] Bansal, R. and I. Shaliastovich (2007). "Risk and Return in Bond, Currency and Equity Markets." *Mimeo*, Duke University.

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

[7] Bansal, R., R. Dittman, and D. Kiku (2009). "Cointegration and Consumption Risks in Asset Returns." *Review of Financial Studies* 22, 1343-1375

[8] Bansal, R., R. Gallant, and G. Tauchen (2008). "Rational Pessimism, Rational Exuberance, and Asset Pricing Models." *Review of Economic Studies* 74, 1005-1033.

[9] Benigno, P. and M. Woodford (2006). "Linear-Quadratic Approximation of Optimal Policy Problems." *Mimeo*, Columbia University.

[10] Binsbergen, J.H. van, J. Fernández-Villaverde, R.S.J. Koijen, and J. Rubio-Ramírez (2009). "Likelihood Estimation of DSGE Models with Epstein-Zin Preferences." *Mimeo*, University of Pennsylvania.

[11] Campanale, C., R. Castro, and G. Clementi (2007). "Asset Pricing in a Production Economy with Chew-Dekel Preferences." *Working Papers* 07-13, New York University, Leonard N. Stern School of Business, Department of Economics.

[12] Campbell, J.Y. (1993). "Intertemporal Asset Pricing without Consumption Data." *American Economic Review* 83, 487-512.

[13] Campbell, J.Y. (1996). "Understanding Risk and Return." *Journal of Political Economy* 104, 298-345.

[14] Campbell, J.Y. and L. Viceira (2001). "Who Should Buy Long-Term Bonds?" *American Economic Review* 91, 99-127.

[15] Chen, X., J. Favilukis, and S. Ludvigson (2007). "An Estimation of Economic Models with Recursive Preferences." *Mimeo*, New York University.

[16] Cho, J-O. and T.C. Cooley (2000). "Business Cycle Uncertainty and Economic Welfare." *Mimeo*, New York University.

[17] Chow, C.S., and J.N. Tsitsiklis (1991). "An Optimal One-Way Multigrid Algorithm for Discrete-Time Stochastic Control." *IEEE Transaction on Automatic Control* 36, 898-914.

[18] Christiano, L., M. Eichenbaum, and C.L. Evans (2005). "Nominal Rigidities and the Dynamic Effects of a Shock to Monetary Policy." *Journal of Political Economy* 113, 1-45.

[19] Croce, M. (2006). "Welfare Costs and Long-Run Consumption Risk in a Production Economy." *Mimeo*, University of North Carolina.

[20] Den Haan, W. J., and A. Marcet (1994). "Accuracy in Simulations." *Review of Economic Studies* 61, 3-17.

[21] Dolmas, J. (1998). "Balanced-growth-consistent Recursive Utility." *Journal of Economic Dynamics and Control* 204, 657-680.

[22] Dolmas, J. (2006). "Real Business Cycle Dynamics under First Order Risk Aversion." *Mimeo*, Federal Reserve Bank of Dallas.

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

[24] Epstein, L., and S.E. Zin (1991). "Substitution, Risk Aversion, and the Temporal Behavior of Consumption and Asset Returns: An Empirical Analysis." *Journal of Political Economy* 99, 263-286.

[25] Fernandez-Villaverde, J. and J. Rubio-Ramirez (2006). "Solving DSGE Models with Perturbation Methods and a Change of Variables." *Journal of Economic Dynamics and Control* 30, 2509-2531.

[26] Fornberg, B., 1998. *A Practical Guide to Pseudospectral Methods*. Cambridge University Press, Cambridge.

[27] Gomes, F. and A. Michealides (2005). "Optimal Life-Cycle Asset Allocation: Understanding the Empirical Evidence." *Journal of Finance* 60, 869-904.

[28] Hansen, L.P., and T.J. Sargent (1995). "Discounted Linear Exponential Quadratic Gaussian Control." *IEEE Transactions on Automatic Control* 40, 968-971.

[29] Hansen, L.P., J. Heaton, and N. Li (2008). "Consumption Strikes Back?: Measuring Long Run Risk." *Journal of Political Economy* 16, 260-302.

[30] Jin, H. and K.L. Judd (2002). "Perturbation Methods for General Dynamic Stochastic Models." *Mimeo,* Hoover Institution.

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

[32] Judd, K.L. (1998). *Numerical Methods in Economics*. MIT Press.

[33] Judd, K.L. and S.M. Guu (1992). "Perturbation Solution Methods for Economic Growth Models." In H. Varian (ed.), *Economic and Financial Modelling with Mathematica*. Springer Verlag.

[34] Judd, K.L. and S.M. Guu (1997). "Asymptotic Methods for Aggregate Growth Models." *Journal of Economic Dynamics and Control* 21, 1025-1042.

[35] Kaltenbrunner, G. and L. Lochstoer (2007). "Long-Run Risk through Consumption Smoothing." *Mimeo*, London Business School.

[36] Kreps, D.M. and E.L. Porteus (1978). "Temporal Resolution of Uncertainty and Dynamic Choice Theory." *Econometrica* 46, 185-200.

[37] Kruger, D. and F. Kubler (2005). "Pareto Improving Social Security When Financial Markets Are Incomplete." *American Economic Review* 96, 737-755.

[38] Lettau, M. and H. Uhlig (2002). "Sharpe Ratios and Preferences: An Analytical Approach." *Macroeconomic Dynamics* 6, 242-265.

[39] Levine, P., J. Pearlman, and R. Pierse (2007). "Linear-Quadratic Approximation, External Habit, and Targeting Rules." *Mimeo*, University of Surrey.

[40] Piazzesi, M. and M. Schneider (2006). "Equilibrium Yield Curves." *NBER Macroeconomics Annual* 2006, 389-442.

[41] Restoy, F. and P. Weil (1998). "Approximate Equilibrium Asset Prices." *NBER Working Papers* 6611.

[42] Rudebusch, G. and E. Swamson (2008). The Bond Premium in a DSGE Model with Long-Run Real and Nominal Risks." *Mimeo*, Federal Reserve Bank of San Francisco.

[43] Schmitt-Grohé, S. and M. Uribe (2004). "Solving Dynamic General Equilibrium Models Using a Second-Order Approximation to the Policy Function.". *Journal of Economic Dynamics and Control* 28, 755-775.

[44] Schmitt-Grohé, S. and M. Uribe (2005). "Optimal Fiscal and Monetary Policy in a Medium Scale Macroeconomic Model." *NBER Macroeconomic Annual 2005*, 382-425.

[45] Smets, F. and R. Wouters (2007). "Shocks and Frictions in US Business Cycles: A Bayesian DSGE Approach." *American Economic Review* 97, 586-606.

[46] Tallarini., T.D. (2000). "Risk-Sensitive Real Business Cycles." *Journal of Monetary Economics* 45, 507-532.

[47] Tauchen, G. (1986). "Finite State Markov-Chain Approximations to Univariate and Vector Autoregressions." *Economics Letters* 20, 177-181.

[48] Uhlig, H. (2007). "Leisure, Growth, and Long Run Risk." *Mimeo*, University of Chicago.

[49] Weil, P. (1990). "Nonexpected Utility in Macroeconomics." *Quarterly Journal of Economics* 105, 29-42.

[50] Woodford, M.D. (2003). *Interest and Prices*. Princeton University Press.