FEDERAL
RESERVE
BANK
of ATLANTA

Comparing Solution Methods for Dynamic Equilibrium Economies

S. Boragan Aruoba, Jesus Fernandez-Villaverde,
and Juan Francisco Rubio-Ramirez

Working Paper 2003-27
October 2003

Working Paper Series

Federal Reserve Bank of Atlanta
Working Paper 2003-27
October 2003

Comparing Solution Methods for Dynamic Equilibrium Economies

S. Boragan Aruoba, University of Pennsylvania
Jesus Fernandez-Villaverde, University of Pennsylvania
Juan Francisco Rubio-Ramirez, Federal Reserve Bank of Atlanta

Abstract: This paper compares solution methods for dynamic equilibrium economies. The authors compute and simulate the stochastic neoclassical growth model with leisure choice using Undetermined Coefficients in levels and in logs, Finite Elements, Chebyshev Polynomials, Second and Fifth Order Perturbations and Value Function Iteration for several calibrations. The authors document the performance of the methods in terms of computing time, implementation complexity and accuracy and they present some conclusions about their preferred approaches based on the reported evidence.

JEL classification: C63, C68, E37

Key words: dynamic equilibrium economies, computational methods, linear and nonlinear solution methods

---

The authors gratefully acknowledge Jose Victor Rios-Rull and Stephanie Schmitt-Grohe and participants at several seminars for useful comments. They also thank Kenneth Judd for encouragement to study perturbation methods further and Mark Fisher for crucial help with Mathmatica idiosyncrasies. The views expressed here are the authors' and not necessarily those of the Federal Reserve Bank of Atlanta or the Federal Reserve System. Any remaining errors are the authors' responsibility.

Please address questions regarding content to S. Boragan Aruoba, University of Pennsylvania, 530 McNeil Building, 3718 Locust Walk, Philadelphia, Pennsylvania 19104, 215-573-2618, aruoba@econ.upenn.edu; Jesus Fernandez-Villaverde, University of Pennsylvania, 160 McNeil Bldg., 3718 Locust Walk, Philadelphia, Pennsylvania 19104 215-898-1504, jesusfv@econ.upenn.edu; or Juan Francisco Rubio-Ramirez, Federal Reserve Bank of Atlanta, 1000 Peachtree Street, N.E., Atlanta, Georgia 30309, 404-498-8057, juan.rubio@atl.frb.org.

Comparing Solution Methods for Dynamic Equilibrium Economies

1. Introduction

This paper addresses the following question: how different are the computational answers provided by alternative solution methods for dynamic equilibrium economies?

It is well known that most dynamic equilibrium models do not have an analytic, closed-form solution and that we need to use numerical methods to approximate their behavior. There are a number of procedures to undertake this task (see Judd, 1998 and Marimón and Scott, 1999 for a general exposition). However it is difficult to assess a priori how the quantitative characteristics of the computed equilibrium paths change when we move from one solution approach to another. Also the relative accuracies of the approximated equilibria are not well understood.

The properties of a solution method are not only of theoretical interest but crucial to assess how reliable the answers provided by quantitative exercises are. For example if we state, as in the classical measurement by Kydland and Prescott (1982), that the productivity shocks account for seventy percent of the fluctuations of the U.S. economy, we want to know that this number is not a by-product of numerical error. Similarly if we use the equilibrium model for estimation purposes we need an approximation that does not introduce bias in the estimates but yet is quick enough to make the exercise feasible.

Over 15 years ago a group of researchers compared solution methods for the stochastic growth model without leisure choice (see Taylor and Uhlig, 1990 and the companion papers). Since then, a number of nonlinear solution methods – several versions of projection (Judd, 1992 and McGrattan, 1999) and asymptotic procedures (Judd and Guu, 1997) – have been proposed as alternatives to more traditional (and relatively simpler) linear approaches and to Value Function Iteration. However, little is known about the relative performance of the new methods.[1] This is unfortunate since these new methods, built on the long experience of applied mathematics, promise superior performance. This paper tries to fill part of this gap

[1] For the stochastic growth model we are only aware of the comparison between Chebyshev Polynomials and different versions of the dynamic programming algorithm and policy iteration undertaken by Santos (1999) and Benítez-Silva *et al.* (2000). However those two paper (except one case in Santos, 1999) only deal with the model with full depreciation and never with the other nonlinear methods.

in the literature.

To do so, we use the canonical stochastic neoclassical growth model with leisure choice, the workhorse of modern macroeconomics. We solve and simulate the model using two linear approximations (based on the linearization of the model equilibrium conditions around the deterministic steady state in levels and in logs) and five nonlinear approximations (Finite Elements, Chebyshev Polynomials and three Perturbation Methods, 2^{nd} order in levels, 2^{nd} order in logs and 5^{th} order in levels). We also solve the model using Value Function Iteration with a multigrid scheme. The results of the Value Function Iteration method are a natural benchmark given our knowledge about the convergence and stability properties of the procedure (see Santos and Vigo, 1998 and references therein).

We report results for a benchmark calibration of the model and for alternative calibrations that change the variance of the productivity shock and the risk aversion. In that way we study the performance of the methods both for a nearly linear case (the benchmark calibration) and highly nonlinear cases (high variance/high risk aversion). In our simulations we keep a fixed set of stochastic shocks common for all methods. That allows us to observe the dynamic responses of the economy to the same driving process and how computed paths and their moments differ for each approximation. We also assess the accuracy of the solution methods by performing Euler Equation tests.

We document substantial differences in algorithmic complexity, time performance and accuracy of the solution. Four main results deserve to be highlighted. First, Perturbation Methods deliver an interesting compromise between accuracy, speed and programming burden. For example, we show how a 5^{th} order perturbation has an advantage in terms of accuracy over all other solution methods for the benchmark calibration around the deterministic steady state. However, perturbations suffer from the need of computing analytical derivatives, a difficult task in low level but efficient languages such as C++ or Fortran 95. Also, since their validity is local, they perform poorly away from the steady state and the simulations display an annoying tendency to explode in the highly nonlinear case.

Second, since higher order perturbations display a much superior performance over linear methods for a trivial marginal cost, we see a compelling reason to move most computations currently undertaken with linear methods to at least a 2^{nd} order perturbation approximation.

Third, even if the performance of linear methods is disappointing along a number of dimensions, linearization in levels is preferred to log-linearization for both the benchmark calibration and the highly nonlinear cases. This results is new and contradicts a common

practice based on a misleading test model.

Fourth, the Finite Elements method performs very well for all parametrizations. It is extremely stable and accurate over the range of the state space even for high values of the risk aversion and the variance of the shock. This property is crucial in estimation procedures where the accuracy is required to obtain unbiased estimates (see Fernández-Villaverde and Rubio-Ramírez, 2003a). However it suffers from being probably the most complicated method to implement in practice (although not the most intensive in computing time). Chebyshev polynomials share all the good properties of the Finite Elements Method and are easier to implement although they suffer from numerical instabilities.

Our results should serve as an encouragement of a wider use of Perturbation Methods, to suggest the reliance on projection for problems that demand high accuracy and stability and support the phasing out of pure linearizations.

The rest of the paper is organized as follows. Section 2 presents the canonical stochastic neoclassical growth model. Section 3 describes the different solution methods used to approximate the policy functions of the model. Section 4 presents the benchmark calibration and alternative robustness calibrations. Section 5 reports numerical results and section 6 concludes. A technical appendix provides further details about all the different methods.

2. The Stochastic Neoclassical Growth Model

As mentioned above we use the basic model in modern macroeconomics, the stochastic neoclassical growth model with leisure as our test model for comparing different approximation procedures. The popularity of the model and its applications (directly or with minor variants) to address a huge number of questions (see Cooley, 1995) makes it a natural laboratory to explore different approximation methods. An alternative (which is actually a special case) could have been the model with log utility function, no leisure choice and total depreciation, a case where a simple closed form solution exists (see Sargent, 1987). However since it is difficult to extrapolate the lessons from this particular example into statements for the more general case, we prefer to pay the cost of not having an explicit analytic solution.[2]

Since the model is well known we only go through the minimum exposition required to fix notation. There is a representative agent in the economy, whose preferences over stochastic

[2]Santos (2000) shows how changes in the curvature of the utility function influence the size of the Euler equation errors.

sequences of consumption and leisure are representable by the utility function

$$U = E_0 \sum_{t=1}^{\infty} \beta^{t-1} \frac{\left(c_t^{\theta} \left(1 - l_t\right)^{1-\theta}\right)^{1-\tau}}{1-\tau} \tag{1}$$

where $\beta \in (0,1)$ is the discount factor, τ is the elasticity of intertemporal substitution, θ controls labor supply and E_0 is the conditional expectation operator.

There is one good in the economy, produced according to the aggregate production function $y_t = e^{z_t} k_t^{\alpha} l_t^{1-\alpha}$ where k_t is the aggregate capital stock, l_t is aggregate labor and z_t is a stochastic process representing random technological progress. The technology follows the process $z_t = \rho z_{t-1} + \epsilon_t$ with $|\rho| < 1$ and $\epsilon_t \sim \mathcal{N}(0, \sigma^2)$. Capital evolves according to the law of motion $k_{t+1} = (1-\delta)k_t + i_t$ and the economy must satisfy the resource constraint $y_t = c_t + i_t$.

Since both welfare theorems hold in this economy, we can solve directly for the social planner's problem where we maximize the utility of the household subject to the production function, the evolution of the stochastic process, the law of motion for capital, the resource constraint and some initial conditions k_0 and z_0.

The solution to this problem is fully characterized by the equilibrium conditions:

$$\frac{\left(c_t^{\theta} \left(1 - l_t\right)^{1-\theta}\right)^{1-\tau}}{c_t} = \beta E_t \left\{ \frac{\left(c_{t+1}^{\theta} \left(1 - l_{t+1}\right)^{1-\theta}\right)^{1-\tau}}{c_{t+1}} \left(1 + \alpha e^{z_{t+1}} k_{t+1}^{\alpha-1} l_{t+1}^{1-\alpha} - \delta\right) \right\} \tag{2}$$

$$(1-\theta) \frac{\left(c_t^{\theta} \left(1 - l_t\right)^{1-\theta}\right)^{1-\tau}}{1 - l_t} = \theta \frac{\left(c_t^{\theta} \left(1 - l_t\right)^{1-\theta}\right)^{1-\tau}}{c_t} (1-\alpha) e^{z_t} k_t^{\alpha} l_t^{-\alpha} \tag{3}$$

$$c_t + k_{t+1} = e^{z_t} k_t^{\alpha} l_t^{1-\alpha} + (1-\delta) k_t \tag{4}$$

$$z_t = \rho z_{t-1} + \varepsilon_t \tag{5}$$

and the boundary condition $c(0, z_t) = 0$. The first equation is the standard Euler equation that relates current and future marginal utilities from consumption, the second one is the static first order condition between labor and consumption and the last two equations are the resource constraint of the economy and the law of motion of technology.

Solving for the equilibrium of this economy amounts to finding three policy functions for next period's capital $k(\cdot,\cdot)$, labor $l(\cdot,\cdot)$ and consumption $c(\cdot,\cdot)$ that deliver the optimal choice of the variables as functions of the two state variables, capital and the technology level.

All the computational methods used below except for the value function iteration exploit directly the equilibrium conditions. This characteristic makes the extension of the methods to non-pareto optimal economies – where we need to solve directly for the market allocation – straightforward. As a consequence we can export at least part of the intuition from the computational results in the paper to a large class of economies.

3. Solution Methods

The system of equations listed above does not have a known analytical solution and we need to use a numerical method to solve it.

The most direct approach is to attack the social planner's problem directly using Value Function Iteration. This procedure is safe and reliable and has useful convergence theorems (Santos and Vigo, 1998). However it is extremely slow (see Rust, 1996 and 1997 for accelerating algorithms) and suffers from a strong curse of the dimensionality. Also it is difficult to use in non-pareto optimal economies (see Kydland, 1989).

Because of these problems, the development of new solution methods for dynamic equilibrium models has been an important area of research in the last decades. These solution methods can be linear or nonlinear. The first ones exploit the fact that many dynamic equilibrium economies display behavior that is close to a linear law of motion.

The second group of methods correct the approximation for higher order terms. Two popular alternatives among these nonlinear approaches are perturbation (Judd and Guu, 1997) and projection methods (Judd, 1992 and McGrattan, 1999). These approaches are attractive because they are much faster than Value Function Iteration while sharing their convergence properties. This point is not only of theoretical importance but of key practical relevance. For instance in estimation problems, since an intermediate step in order to evaluate the likelihood function of the economy is to solve for the policy functions, we want to use a fast solution method since we may need to perform a huge number of these evaluations for different parameter values. Convergence properties assure us that, up to some fixed accuracy level, we are indeed getting the correct equilibrium path for economy.

In this paper we compare eight different methods. As our linear method, we use Undetermined Coefficients to solve for the unknown coefficients of the policy functions using linearized versions of the equilibrium equations of the model, both in levels and in logs.[3]

[3]Note that, subject to applicability, all different linear methods described in the literature -Linear

For the nonlinear methods we compute a Finite Elements method, a spectral procedure with Chebyshev Polynomials, three Perturbation Approaches (a 2^{nd} order expansion in levels, a 5^{th} order expansion in levels and a 2^{nd} order expansion in logs) and Value Function Iteration.[4]

We now briefly describe each of these methods. The technical appendix provides many more details about the procedures and the computational parameters choices.

3.1. Undetermined Coefficients in Levels

The basic idea of this approximation is to substitute the system of equilibrium conditions with a linearized version of it. Solving a system of linear difference equations is a well understood problem. A simple procedure is to guess a linear policy function with undetermined coefficients, plug it in the linear system and solve the resulting quadratic system for the unknown coefficients (see Uhlig, 1999 for details). Beyond simplicity and speed, the procedure also allows us to derive some analytical results about the model (see Campbell, 1994).

If we linearize the set of equilibrium conditions presented above around the steady state value x of the variables x_t we get the linear system[5]:

$$\frac{\theta(1-\tau)-1}{c}(c_t - c) - \frac{(1-\tau)(1-\theta)}{1-l}(l_t - l) = \tag{6}$$

$$E_t \left\{ \begin{array}{c} \frac{\theta(1-\tau)-1}{c}(c_{t+1} - c) + \left(\beta \frac{\alpha(1-\alpha)}{l}\frac{y}{k} - \frac{(1-\tau)(1-\theta)}{1-l}\right)(l_{t+1} - l) \\ +\alpha\beta\frac{y}{k} z_{t+1} + \beta\frac{\alpha(\alpha-1)y}{k^2}(k_{t+1} - k) \end{array} \right\} \tag{7}$$

$$\frac{1}{c}(c_t - c) + \frac{1}{(1-l)}(l_t - l) = z_t + \frac{\alpha}{k}(k_t - k) - \frac{\alpha}{l}(l_t - l) \tag{8}$$

$$(c_t - c) + (k_{t+1} - k) = y\left(z_t + \frac{\alpha}{k}(k_t - k) + \frac{1-\alpha}{l}(l_t - l)\right) + (1-\delta)(k_t - k) \quad \text{(intralineal)}$$

$$z_t = \rho z_{t-1} + \varepsilon_t \tag{9}$$

Quadratic approximation (Kydland and Prescott, 1982), the Eigenvalue Decomposition (Blanchard and Kahn, 1980 and King, Plosser and Rebelo, 2002), Generalized Schur Decomposition (Klein, 2000 or the QZ decomposition (Sims, 2002) among many others, should deliver the same results. The linear approximation of a differentiable function is unique and invariant to differentiable parameters transformations. Our particular choice of linear method is purely a matter of taste.

[4] We do not try to cover every single known method but rather to be selective and choose those methods that we find more promising based either on experience or on intuition from numerical analysis. Below we discuss how some apparently excluded methods are particular cases of some of our approaches.

[5] See the technical appendix for a discussion of alternatives points for the linearization.

Simplification and some algebra deliver:

$$A\widehat{k}_{t+1} + B\widehat{k}_t + C\widehat{l}_t + Dz_t = 0$$
$$E_t\left(G\widehat{k}_{t+1} + H\widehat{k}_t + J\widehat{l}_{t+1} + K\widehat{l}_t + Lz_{t+1} + Mz_t\right) = 0$$
$$E_t z_{t+1} = Nz_t$$

where the coefficients $A, B, C, ..., N$ are functions of the model parameters and $\widehat{x}_t = x_t - x$.

Now we guess policy functions of the form $\widehat{k}_{t+1} = P\widehat{k}_t + Qz_t$ and $\widehat{l}_t = R\widehat{k}_t + Sz_t$, plug them in the linear system and solve the resulting quadratic problem for the unknown coefficients P, Q, R and S that imply a stable solution. Note that the procedure delivers a linear law of motion for the choice variables that displays certainty equivalence (i.e. it does not depend on σ). This point will be important when we discuss our results. The other variables in the model are solved for using the linearized system and the computed policy functions.

3.2. Undetermined Coefficients in Logs

Since the exact solution of the stochastic neoclassical growth model in the case of log utility, total depreciation and no leisure choice is loglinear, a large share of practitioners have favored the loglinearization of the equilibrium conditions of the model over linearization. Some evidence in Christiano (1990) and Den Haan and Marcet (1994) suggest that this is the right practice but the question is not completely settled. To cast light on this question and perform a systematic comparison of both alternatives below, we repeat our undetermined coefficient procedure in logs: we loglinearize the equilibrium conditions instead of linearizing them but proceed otherwise as before.

In particular we take the equilibrium conditions of the model and we substitute each variable x_t by $xe^{\widehat{x}_t}$ where x is the steady state and $\widehat{x}_t = \log \frac{x_t}{x}$. Then we linearize with respect to $\widehat{x}_t$ around $\widehat{x}_t = 0$ (i.e. the steady state). After some algebra we get:

$$A\widehat{k}_{t+1} + B\widehat{k}_t + C\widehat{l}_t + Dz_t = 0$$
$$E_t\left(G\widehat{k}_{t+1} + H\widehat{k}_t + J\widehat{l}_{t+1} + K\widehat{l}_t + Lz_{t+1} + Mz_t\right) = 0$$
$$E_t z_{t+1} = Nz_t$$

where the coefficients $A, B, C, ..., N$ are functions of the parameters of the model.

We guess policy functions of the form $\widehat{k}_{t+1} = P\widehat{k}_t + Qz_t$ and $\widehat{l}_t = R\widehat{k}_t + Sz_t$, plug them in the linear system and solve for the unknown coefficients.[6]

3.3. Finite Elements Method

The Finite Elements Method (Hughes, 2000 and McGrattan, 1999) is the most widely used general-purpose technique for numerical analysis in engineering and applied mathematics. Beyond being conceptually simple and intuitive, the Finite Elements Method features several interesting properties. First, it provides us a lot of flexibility in the grid generation: we can create very small elements (and consequently very accurate approximations of the policy function) in the neighborhood of the mean of the stationary distribution of capital and larger ones in the areas of the state space less travelled. Second, large numbers of elements can be handled to exploit the sparsity of the problem. Third, the Finite Elements method is well suited for implementation in parallel machines with the consequent scalability of the problem.

The Finite Elements Method searches for a policy function for labor supply of the form $l_{fe}\left(k, z; \overline{\theta}\right) = \sum_{i,j} \overline{\theta}_{ij} \Psi_{ij}\left(k, z\right)$ where $\Psi_{ij}\left(k, z\right)$ is a set of basis functions and $\overline{\theta}$ is a vector of parameters to be determined. Note that given $l_{fe}\left(k, z; \overline{\theta}\right)$, the static first order condition and the resource constraint imply two policy function $c(k, z; l_{fe}\left(k, z; \overline{\theta}\right))$ and $k'\left(k, z; l_{fe}\left(k, z; \overline{\theta}\right)\right)$ for consumption and next period capital. The essence of the Finite Elements method is to chose basis functions that are zero for most of the state space except a small part of it, an interval in which they take a very simple form, typically linear.[7]

First we partition the state space Ω in a number of nonintersecting rectangles $[k_i, k_{i+1}] \times [z_j, z_{j+1}]$ where k_i is the ith grid point for capital and z_j is jth grid point for the technology

[6] Alternatively we could have taken the coefficients from the linearization in levels and transform them using a nonlinear change of variables and the Chain Rule. The final results would be the same as the ones in the paper. See Judd (2003) and Fernández-Villaverde and Rubio-Ramírez (2003b).

[7] We can have more elaborated basis functions as Chebyshev polynomials and solve the resulting Spectral-Finite Elements problem. These type of schemes, known as the *p-method* are much less used than the so-called *h-method* whereby the approximation error is reduced through successive mesh refinement.

shock. As basis functions we set $\Psi_{ij}(k,z) = \widehat{\Psi}_i(k)\,\widetilde{\Psi}_j(z)$ where

$$\widehat{\Psi}_i(k) = \begin{cases} \frac{k-k_i}{k_{i+1}-k_i} & \text{if } k \in [k_{i-1}, k_i] \\ \frac{k_{i+1}-k}{k_{i+1}-k_i} & \text{if } k \in [k_i, k_{i+1}] \\ 0 \text{ elsewhere} \end{cases}$$

$$\widetilde{\Psi}_j(z) = \begin{cases} \frac{z-z_j}{z_{j+1}-z_j} & \text{if } z \in [z_{j-1}, z_j] \\ \frac{z_{j+1}-z}{z_{j+1}-z_j} & \text{if } z \in [z_j, z_{j+1}] \\ 0 \text{ elsewhere} \end{cases}$$

Then we plug $l_{fe}\left(k,z;\overline{\theta}\right)$ and $c(k,z;l_{fe}\left(k,z;\overline{\theta}\right))$ and $k'\left(k,z;l_{fe}\left(k,z;\overline{\theta}\right)\right)$ in the Euler Equation to get a residual function $R(k_t,z_t;\overline{\theta})$.

A natural criterion for finding the $\overline{\theta}$ unknowns is to minimize this residual function over the state space given some weight function. To do so we employ a Galerkin scheme where the basis functions double as weights to get the nonlinear system of $\overline{\theta}$ equations

$$\int_{\Omega} \Psi_{i,j}(k,z)\, R(k,z;\overline{\theta}) dz dk = 0 \quad \forall i,j \tag{10}$$

on our $\overline{\theta}$ unknowns. Solving this system delivers our desired policy function $l_{fe}\left(k,z;\overline{\theta}\right)$ from which we can find all the other variables in the economy.[8]

3.4. Spectral Method (Chebyshev Polynomials)

Like Finite Elements, spectral methods (Judd, 1992) search for a policy function of the form $l_{sm}\left(k,z;\overline{\theta}\right) = \sum_{i,j} \overline{\theta}_{ij}\Psi_{ij}(k,z)$ where $\Psi_{ij}(k,z)$ is a set of basis functions and $\overline{\theta}$ is a vector of parameters to be determined. The difference with respect to the previous approach is that the basis functions are (almost surely) nonzero, i.e. we search for a global solution instead of pasting together local solutions as we did before.

Spectral methods have two main advantages over the Finite Elements method. First, they are generally much easier to implement. Second, since we can easily handle a large number of basis functions the accuracy of the procedure is potentially very high. The main drawback of

[8] Note that policy function iteration (see for example Coleman, 1990) is just a particular case of the Finite Elements when we pick a collocation scheme in the points of an exogenously given grid, linear basis functions and an iterative scheme to solve for the unknown coefficients. Experience from numerical analysis suggests that nonlinear solvers (as the simple Newton scheme that we used for our unknown coefficients) or multigrid schemes outperform pure iterative algorithms (see Briggs, Henson, and McCormick, 2000). Also Galerkin weigthings are superior to collocation for Finite Elements (Boyd, 2001).

the procedure is that it approximates the true policy function globally. If the policy function displays a rapidly changing local behavior the scheme may deliver a poor approximation.

A common choice for the basis functions is to set the tensor $\Psi_{ij}(k,z) = \widehat{\Psi}_i(k)\widetilde{\Psi}_j(z)$ where $\widehat{\Psi}_i(\cdot)$ and $\widetilde{\Psi}_j(\cdot)$ are Chebyshev polynomials (see Boyd, 2001 and Fornberg, 1998 for justifications of this choice of basis functions). These polynomials can be recursively defined by $T_0(x) = 1$, $T_1(x) = 1$ and for general n, $T_{n+1}(x) = 2xT_n(x) - T_{n-1}(x)$.[9]

As in the previous case we use the two Euler Equations with the budget constraint substituted in to get a residual function $R(k_t, z_t; \overline{\theta})$. Instead of a Galerkin weighting, computational experience (Fornberg, 1998) suggests that, for spectral methods, a collocation (also known as pseudospectral) criterion delivers the best trade-off between accuracy and ability to handle large number of basis functions. Collocation uses as weights the $n \times m$ dirac functions δ_j with unit mass in $n \times m$ points (n from the roots of the last polynomial used in the capital dimension and m from the points in Tauchen's, 1986 approximation to the stochastic process for technology). This scheme results in the nonlinear system of $n \times m$ equations

$$R(k_{ij}, z_{ij}; \overline{\theta}) = 0 \quad \text{for } \forall \ n \times m \text{ collocation points} \tag{11}$$

in $n \times m$ unknowns. This system is easier to solve than (10) since we will have in general less basis functions and we avoid the integral induced by the Galerkin weigthing.[10]

3.5. Perturbation

Perturbation methods (Judd and Guu, 1997) build a Taylor series expansion of the policy functions of the economy around some point of the state space k_0, z_0 and a perturbation parameter, in our case the standard deviation of the innovation to the productivity level σ.[11]

[9] The domain of the Chebyshev polynomials is $[-1, 1]$. Since our state space is different in general we use a linear mapping from $[a, b]$ into $[-1, 1]$.

[10] Parametrized expectations (see Marcet and Lorenzoni, 1999 for a description) is a spectral method that uses monomials (or exponents of) in the current states of the economy and montecarlo integration. Since monomials are highly collinear and determinist integration schemes are preferred for low dimensional problems over montecarlo approaches (Geweke, 1996), we stick with Chebyshev polynomials as our favorite spectral approximation. See Christiano and Fisher (2000) for a thorough explanation.

[11] The choice of perturbation parameter is model-dependent. Either the standard deviation (for discrete time models) or the variance (for continuous time models) are good candidates for stochastic equilibrium economies.

The policy function for consumption takes the form

$$c_p(k,z,\sigma) = \sum_{i,j,m} a_{ijm} \left(k - k_0\right)^i \left(z - z_0\right)^j \sigma^m$$

and all the other policy function will have analogous forms.

The key idea of the perturbation methods is to play with the parameter σ to find a case where the model can be solved analytically and to exploit implicit-function theorems to pin down the unknown coefficients a_{ijm} in a recursive fashion.[12] In particular we will set $a_{ijm} = \left.\frac{\partial^{i+j+m} c_p(k,z,\sigma)}{\partial k^i \partial z^j \partial \sigma^m}\right|_{k,z,0}$, i.e. equal to the derivative of the policy function evaluated at the steady state value of the variables and $\sigma = 0$, a point (the deterministic case) in which we can compute analytic expressions easily. Perturbations only deliver an asymptotically correct expression around the deterministic steady state for the policy function but given the positive experience of asymptotic approximations in other fields of applied mathematics, there is the potential for good nonlocal behavior (see Bender and Orszag, 1999).

Our perturbation scheme works as follows. We take the model equilibrium conditions and we substitute in the true (but unknown) policy functions $c(k,z,\sigma)$, $l(k,z,\sigma)$ and $k'(k,z,\sigma)$. Then we take successive derivatives with respect to the state variables and σ. Since the equilibrium conditions must be equal to zero for any value of k, z and σ, the derivatives of any order will also be equal to zero. Evaluating the derivatives at the steady state value of the variables and $\sigma = 0$ delivers a system of equations on the unknown coefficients a_{ijm}.

Solution of the systems is simplified because of the recursive structure of the problem. The constant term a_{000} must be equal to consumption in the deterministic steady state. Therefore, to pin it down we compute the deterministic steady state of the model and let $a_{000} = c$. Substituting this term (and other constant terms of the policy functions) in the system of first derivatives of the equilibrium conditions generates a set of quadratic equations in first order unknown coefficients. The solution to that system is the same that in our undetermined coefficient approximation before. This is the reason why we can interpret undetermined coefficients as a first order perturbation.

The next step is to take second order derivatives. In this system we substitute the known coefficients from the previous two steps and we find a linear system on the second order terms of the policy function. After solving for these terms we take the system of third

[12] In a complementary way we can also use σ to develop analytical asymptotic methods to characterize the average behavior of the economy. See Williams (2002).

order derivatives, substitute the known constant (linear and second order terms) and again we get a linear system that we can easily solve. Iterating the procedure we could see that all the coefficients a_{ijm} after the linear term are just the solution to linear systems. The intuition of why only the system for the linear term is quadratic is simple: the stochastic neoclassical growth model has two saddle paths and the quadratic system is just recording that information. Once we have picked the right path all the other terms are just refinements of the path including information on σ that does not appear in the linear approximation.

The burden of the method is taking all the required derivatives. Paper and pencil become virtually infeasible after the second derivatives. Higher order numerical derivatives accumulate enough errors to prevent their use. The best alternative is the use of symbolic manipulation software such as *Mathematica*. However that means that we lose the speed of low level languages as C++ or Fortran 95. In the absence of publicly available libraries for analytic derivation in this languages, the required use of less powerful software limits the applicability of perturbation.

When implementing the approximation we face two choices. First we need to decide the order of the perturbation. We choose 2^{nd} and 5^{th} order perturbations. Second order approximations have received attention because of the easiness of their computation (see Schmitt-Grohé and Uribe, 2002 and Sims, 2000) and we find of interest to assess how much gain is obtained by this simple correction of the linear policy functions. Then we pick a high order approximation. After the fifth order the coefficients are nearly equal to the machine zero (in a 32-bits architecture of the standards PCs) and further terms do not add much to the behavior of the approximation. The second choice is whether to undertake our perturbation in levels and logs. We performed both cases but because of space considerations we only present results in levels except for the 2^{nd} order approximation for the high variance/high risk aversion case where we report results both in levels and in logs. The omitted results were nearly indistinguishable from perturbations in levels since the additional quadratic term in both expansions corrected for the differences in the linear term between levels and logs.

3.6. Value Function Iteration

Finally we solve the model using value function iteration. Since the dynamic algorithm is well known we only present a sparse discussion.

We generate a grid for capital and we discretize the productivity level using the method proposed by Tauchen (1986). We use a multigrid scheme where the last step has a uniform

one million points grid, with 25000 points for capital and 40 for the productivity level. Then for each point in the grid we iteratively compute the Bellman operator:

$$TV(k,z) = \max_{c>0,0<l<1,k'>0} \frac{\left(c^{\theta}(1-l)^{1-\theta}\right)^{1-\tau}}{1-\tau} + \beta EV(k',z'|z) \quad (12)$$

$$c + k' = \exp(z) k^{\alpha} l^{1-\alpha} + (1-\delta) k \quad (13)$$

$$z' = \rho z + \varepsilon \quad (14)$$

We explore different interpolation schemes (linear, quadratic and Schumaker, 1993) for values of the function outside the grid until convergence and report the ones with better performance.

4. Calibration: Benchmarks Case and Robustness

To make our comparison results as useful as possible we pick a benchmark calibration and we explore how those results change as we move to different "unrealistic" calibrations.

We select the benchmark calibration values for the stochastic neoclassical growth model as follows. The discount factor $\beta = 0.9896$ matches an annual interest rate of 4% (see McGrattan and Prescott, 2000 for a justification of this number based on their measure of the return on capital and on the risk-free rate of inflation-protected U.S. Treasury bonds). The risk aversion $\tau = 2$ is a common choice in the literature. $\theta = 0.357$ matches the microeconomic evidence of labor supply to 31% of available time in the deterministic steady state. We set $\alpha = 0.4$ to match labor share of national income (after the adjustments to National Income and Product Accounts suggested by Cooley and Prescott, 1995). The depreciation rate $\delta = 0.0196$ fixes the investment/output ratio and $\rho = 0.95$ and $\sigma = 0.007$ match the stochastic properties of the Solow residual of the U.S. economy. The chosen values are summarized in table 4.1.

Table 4.1: Calibrated Parameters

Parameter	β	τ	θ	α	δ	ρ	σ
Value	0.9896	2.0	0.357	0.4	0.0196	0.95	0.007

To check robustness, we repeat our analysis for five other calibrations. As explained in the introduction this analysis allows us to study the relative performance of the methods both for a nearly linear case (the benchmark calibration) and for highly nonlinear cases (high variance/high risk aversion). We increase the risk aversion to 10 and 50 and the standard

deviation of the productivity shock to 0.035. Although below we concentrate on the results for the benchmark and the extreme case, the intermediate cases are important to check that our comparison across calibrations does not hide non-monotonicities. Table 4.2. summarizes our different cases.

Table 4.2: Sensitivity Analysis

case	$\sigma = 0.007$	$\sigma = 0.035$
$\tau = 2$	Benchmark	Intermediate Case 3
$\tau = 10$	Intermediate Case 1	Intermediate Case 4
$\tau = 50$	Intermediate Case 2	Extreme

Also we briefly discuss some results for the deterministic case $\sigma = 0$ since they well help us understand some characteristics of the proposed methods.

5. Numerical Results

In this section we report results from our different methods and calibrations. We concentrate on the benchmark and extreme calibrations, reporting the intermediate cases when they clarify the argument.[13] First we present and discuss the computed policy functions. Second we show some simulations. Third we perform accuracy tests, the χ^2 test proposed by Den Haan and Marcet (1990), the Euler Error function proposed by Judd (1992) and Judd and Guu (1997) and a weighting of the Euler Equation error using the simulated distributions. Finally we discuss some details about implementation and computing time.

5.1. Policy Functions

One of the first outputs of the computation is the policy functions. We plot the decision rules for labor supply when $z = 0$ over a capital interval centered around the deterministic steady state level of capital for the benchmark calibration in Figure 5.1.1 and for investment in Figure 5.1.2.[14] Labor supply is very similar in all methods, especially in the neighborhood of 23.14, the deterministic steady state level of capital. Only far away from that neighborhood we appreciate any relevant difference.[15] A similar description applies to the policy rule for

[13] All additional results are available upon request.

[14] Similar figures could be plotted for other values of z. We omit them because of space considerations.

[15] We must be cautious mapping differences in choices into differences in utility (see Santos, 2000). The Euler Error function below provides a better view of the welfare consequences of different approximations.

investment except for the loglinear approximation where the rule is pushed away from the other ones for low and high capital. The difference is big enough that even the monotonicity of the policy function is lost. In this behavior rests already a hint of the problems with loglinearization that we will discuss below.

Dramatic differences appear as we begin to increase risk aversion and the variance of the shock. The biggest discrepancy are for the extreme calibration. The policy functions for this case are presented in Figures 5.1.3 and 5.1.4. In these figures we change the interval reported because, due to the risk aversion/high variance of the calibration, the equilibrium paths will fluctuate around much higher levels of capital (between 30 and 45) when the solution method accounts for that high variance (i.e. all except linearizations). Our new interval is consequently more representative.

We highlight several results. First, the linear and loglinear policy functions deviate from all the other ones: they imply much less labor (around 10%) and investment (up to 30%) than the group of nonlinear methods. This difference in level is due to the lack of correction for increased variance of the technology shock by these two approximations since they are certainty-equivalent. This is proof of how linearization and certainty equivalence produce biased results. Second just correcting for quadratic terms in the 2^{nd} order perturbation allows to get the right level of the policy functions. This is another key point in our argument in favor of phasing out linearizations and substitute them by at least 2^{nd} order perturbations. Third, the policy function for labor and investment approximated by the 5^{th} order perturbation changes from concavity into convexity for values of capital bigger than 45. (contrary to the theoretical results) This change of slope will cause problems below in our simulations.[16]

5.2. Simulations

Practitioners often rely in statistics from simulated paths of the economy. We computed 1000 simulations of 500 observations each for all different methods. To make comparisons meaningful we keep the productivity shock constant across methods for each particular simulation.

We plot in Figures 5.2.1-5.2.4 the histograms for output, capital, labor and consumption for the different methods for the benchmark calibration (where we have dropped the first 100 observations of each simulation as a burn-in period). As we could have suspected after

[16] One last result is that the policy functions have a positive slope. This is because households are so risk-adverse that they want to work hard when capital is high to accumulate even more capital and insure against future bad shocks. Numerically we found that the change in slope occurs for τ around 40.

looking at the policy functions, the histograms suggest that the different methods deliver basically the same behavior for the economy. That impression is reinforced by Figures 5.2.5 and 5.2.6 where we plot the paths of output and capital for one randomly chosen simulation. All paths are roughly equal. This similarity of the simulation paths causes that the business cycle statistics for the model under different solution methods (not reported here but available upon request) to be nearly identical.

We repeat the same exercise for the extreme calibration in Figures 5.2.7-10. We see three groups: first the two linear methods, second the perturbations and finally the three global methods (value function, Finite Elements and Chebyshev). The last two groups have the histograms shifted to the left: much more capital is accumulated and more labor supplied by all the methods that allow for corrections by variance. For example the empirical distributions of nonlinear methods accumulate a large percentage of their mass between 40 and 50 while the linear methods rarely visit that region. Also, it is clear from the figures that even different non-linear methods provide quite a diverse description of the behavior of economy. In particular the three global methods are in a group among themselves (nearly on top of each other) separated from perturbations that lack enough variance. Figures 5.2.11 and 5.2.12 plot a simulation of the economy randomly chosen where the differences in output and capital are easy to visualize.

Higher risk aversion/high variance also have an impact for business cycle statistics. For example investment is three times more volatile in the linear simulation than with Finite Elements despite the filtering of the data.

The simulations show an important drawback of using perturbations to characterize equilibrium economies. For example in 39 simulations out of the 1000 (not shown on the histograms) 5^{th} order perturbation generated a capital that exploded. The reason for that abnormal behavior is the change in the slope of the policy functions reported above. When the economy begins to travel that part of the policy functions the simulation falls in an unstable path and the results need to be disregarded. This instability property is an important problem of perturbations that may limit its use.[17]

[17] We also had problems in the high risk aversion/high variance with 1 2^{nd} order perturbations, 1 log 2^{nd} order perturbations and 65 linearizations in levels (those last ones because capital goes below zero). In the benchmark calibration we did not have any problems.

5.3. A χ^2 Accuracy Test

From our previous discussion it is clear that the consequences for simulated equilibrium paths of using different methods are important. A crucial step in our comparison is then the analysis of accuracy of the computed approximations to figure it out which one we should prefer.

We begin that investigation implementing the χ^2-test proposed by Den Haan and Marcet (1990). The authors noted that if the equilibrium of the economy is characterized by a system of equations $f(y_t) = E_t(\phi(y_{t+1}, y_{t+2}, ..))$ where the vector y_t contains all the n variables that describe the economy at time t, $f : \Re^n \to \Re^m$ and $\phi : \Re^n \times \Re^\infty \to \Re^m$ are known functions and $E_t(\cdot)$ represent the conditional expectation operator, then

$$E_t(u_{t+1} \otimes h(x_t)) = 0 \tag{15}$$

for any vector x_t measurable with respect to t with $u_{t+1} = \phi(y_{t+1}, y_{t+2}, ..) - f(y_t)$ and $h : \Re^k \to \Re^q$ being an arbitrary function.

Given one of our simulated series of length T from the method i in previous section, $\{y_t^i\}_{t=1}^T$, we can find $\left\{u_{t+1}^i, x_t^i\right\}_{t=1}^T$ and compute the sample analog of (15):

$$B_T^i = \frac{1}{T}\sum_{t=1}^{T} u_{t+1}^i \otimes h\left(x_t^i\right) \tag{16}$$

Clearly (16) would converge to zero as T increases almost surely if the solution method were exact. However, given the fact that we only have numerical methods to solve the problem, this may not be the case in general. However the statistic $T\left(B_T^i\right)'\left(A_T^i\right)^{-1}B_T^i$ where A_T^i is a consistent estimate of the matrix $\sum_{t=-\infty}^{\infty} E_t\left[(u_{t+1} \otimes h(x_t))(u_{t+1} \otimes h(x_t))'\right]$ given solution method i, converges in distribution to a χ^2 with qm degrees of freedom under the null that (15) holds. Values of the test above the critical value can be interpreted as evidence against the accuracy of the solution.

Since any solution method is an approximation, as T grows we will eventually always reject the null. To control for this problem, we can repeat the test for many simulations and report the percentage of statistics in the upper and lower critical 5% of the distribution. If the solution provides a good approximation, both percentages should be close to 5%.

We report results for the benchmark calibration in Table 5.3.1 and plot the Empirical CDF

in Figure 5.3.1.[18] All the methods perform similarly and reasonably close to the nominal coverages,with a small bias towards the right of the distribution. Also, and contrary to some previous findings for simpler models (as reported by Den Haan and Marcet, 1994 and Christiano, 1990) it is not clear that we should prefer loglinearization to linearization.

Table 5.3.1: χ^2 Accuracy Test, $\tau = 2/\sigma = 0.007$

	Less than 5%	More than 95%
Linear	3.10	5.40
Log-Linear	3.90	6.40
Finite Elements	3.00	5.30
Chebyshev	3.00	5.40
Perturbation 2	3.00	5.30
Perturbation 5	3.00	5.40
Value Function	2.80	5.70

We present the results for the extreme case in table 5.3.2 and Figure 5.3.2.[19] Now the performance of the linear methods deteriorates enormously, with quite unacceptable coverages (although again linearization in levels is no worse than loglinearization). On the other hand nonlinear methods deliver quite a good performance, with very reasonable coverages on the upper tail (except 2^{nd} order perturbations). The lower tail behavior is poor for all methods.

Table 5.3.2: χ^2 Accuracy Test, $\tau = 50/\sigma = 0.035$

	Less than 5%	More than 95%
Linear	0.43	23.42
Log-Linear	0.40	28.10
Finite Elements	1.10	5.70
Chebyshev	1.00	5.20
Perturbation 2	0.90	12.71
Perturbation 2-Log	0.80	22.22
Perturbation 5	1.56	4.79
Value Function	0.80	4.50

[18] We use a constant, k_t, k_{t-1}, k_{t-2} and z_t as our instruments, 3 lags and a Newey-West Estimator of the matrix of variances-covariances (Newey and West, 1987).

[19] The problematic simulations as described above are not included in these computations.

5.4. Non Local Accuracy test

The previous test is a simple procedure to evaluate the accuracy of the solution procedure. That approach may suffer, however, from three problems. First, since all methods are just approximations the test will display poor power. Second orthogonal residuals can be compatible with large deviations from the optimal policy. Third, by its design the model will spend most of the time in those regions where the density of the stationary distribution is higher. Often it is important to assess the accuracy of a model far away from the steady state as in estimation procedures where we want to explore the global shape of the likelihood function.

Judd (1992) proposes to determine the quality of the solution method defining normalized Euler Equation Errors. First note that in our model the intertemporal condition

$$u_c'\left(c\left(k_t\,,z_t\right),l\left(k_t\,,z_t\right)\right)=\beta E_t\left\{u_c'\left(c\left(k\left(k_t\,,z_t\right)\,,z_{t+1}\right),l\left(k\left(k_t\,,z_t\right)\,,z_{t+1}\right)\right)R\left(k_t\,,z_t\,,z_{t+1}\right)\right\} \tag{17}$$

where $R\left(k_t\,,z_t,z_{t+1}\right)=\left(1+\alpha e^{z_{t+1}}k\left(k_t\,,z_t\right)^{\alpha-1}l\left(k\left(k_t\,,z_t\right),z_{t+1}\right)^{1-\alpha}-\delta\right)$ is the gross return rate of capital, should hold exactly for given k_t and z_t.

Since the solution methods used are only approximations, (17) will not hold exactly when evaluated using the computed decision rules. Instead, for solution method i with associated policy rules $c^i\left(\cdot\,,\cdot\right)$, $l^i\left(\cdot\,,\cdot\right)$ and $k^i\left(\cdot\,,\cdot\right)$ and the implied gross return of capital $R^i\left(k_t\,,z_t,z_{t+1}\right)$, we can define the Euler Equation error function $EE^i\left(\cdot\,,\cdot\right)$ as:

$$EE^i\left(k_t\,,z_t\right)\equiv 1-\frac{\left(\frac{\beta E_t\left\{u_c'\left(c^i\left(k^i\left(k_t\,,z_t\right),z_{t+1}\right),l^i\left(k^i\left(k_t\,,z_t\right),z_{t+1}\right)\right)R^i\left(k_t\,,z_t,z_{t+1}\right)\right\}}{\theta\left(1-l^i\left(k^i\left(k_t\,,z_t\right),z_{t+1}\right)\right)^{(1-\theta)(1-\tau)}}\right)^{\frac{1}{\theta(1-\tau)-1}}}{c^i\left(k_t\,,z_t\right)} \tag{18}$$

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. Judd and Guu (1997) interpret this error as the relative optimization error incurred by the use of the approximated policy rule. For instance if $EE^i\left(k_t\,,z_t\right)=0.01$, then the agent is making a \$1.00 mistake for each \$100 spent. In comparison, $EE^i\left(k_t\,,z_t\right)=1e^{-8}$ implies that the agent is making a 1 cent mistake for each million of dollars spent.

The Euler Equation error is also important because we know that under certain conditions, the approximation error of the policy function is of the same order of magnitude as the size of the Euler equation residual and correspondingly the change in welfare is of the square order of the Euler equation residual (Santos, 2000).

Figures 5.4.1-5.4.10 present the Euler Equation Error functions for our benchmark calibration. Figure 5.4.1 shows the results for the linear approximation to the equilibrium conditions for capital between 70% and 130% of the deterministic steady state level (23.14) and for a range of technology shocks from -0.065 to 0.065 (with zero being the level of technology in the deterministic case).[20] We plot the absolute errors in base 10 logarithms to ease interpretation. A value of -3 means \$1 mistake for each \$1000, a value of -4 a \$1 mistake for each \$10000 and so on. As intuition would suggest the error is much lower around the central regions, closer to the point around which we make our linear expansion. The quality of the approximation deteriorates as we move away from the central regions and quickly reaches -3. Figure 5.4.2 follows the same convention and plots the errors of the loglinear approximation. We can see a pattern with two narrow valleys of high accuracy surrendered by regions with worse errors. The origin of these valleys will be explained below. As in the case of the χ^2-test, from the comparison of figures 5.4.1 and 5.4.2 it is not obvious that we should prefer loglinear approximations to straight linearizations.

The next two figures display the results for Finite Elements (figure 5.4.3) and Chebyshev polynomials (figure 5.4.4). Finite Elements delivers a very robust performance along the state space, specially for technology shocks between -0.02 and 0.02 (our mesh is finer in this region) where the errors fluctuate around -7 (with some much better points around the nodes of the elements). Only for large shocks (where our mesh is coarser) the performance of finite elements deteriorates. Chebyshev polynomials emerge from figure 5.4.4 as a very competitive solution method: the error is consistently below -8. Given the much lower computational burden of Chebyshev polynomials versus Finite Elements, the result is encouraging for spectral methods.

Figures 5.4.5 and 5.4.6 present the Euler Error functions for the 2^{nd} and 5^{th} order perturbations. Figure 5.4.5 proves how we can strongly improve the accuracy of the solution over a linear approximation paying only a trivial additional cost that delivers a result nearly as good as Finite Elements. Correcting for variance and quadratic terms reduces Euler errors by an order of magnitude over the results from linear methods. The 5^{th} order approximation performance is superb. Over the whole range, its error is less than -7 and in the central regions up to –8.

Finally Figure 5.4.7 graphs the Euler Errors for the Value Function iteration that fluctuate around -5 with the ups and downs induced by the grid and the expected uniform performance over the state space.

[20] 0.065 corresponds to roughly 99.5^{th} percentile of the normal distribution given our parameterization.

To get a better view of the relative performance of each approximation and since plotting all the error functions in a same plot is cumbersome, Figure 5.4.8 displays a transversal cut of the errors when the technology is equal to zero. Here we can see many of the same results we just discussed. The loglinear approximation is worse than the linearization except at the two valleys. Finite Elements and Chebyshev polynomials perform much better than the linear methods (three orders of magnitude even at the steady state) and perturbations' accuracy is impressive. Other transversal cuts at different technology levels reveal similar patterns.

To explore further the origin of the errors we plot in figure 5.4.9 the level of the euler errors at $z = 0$. With this graph we can explain the two valleys of the loglinearization: at the deterministic steady state level of capital loglinearization induces a negative bias in the euler equation while the errors tend to grow quickly away from it. The two valleys are just the two neighborhoods where the parabola crosses the zero. The parabola of the linearization is always positive (something by itself neutral) but much flatter. Our reading of these shapes is that linearization may be better than loglinearization after all. Figure 5.4.10 plots the same figure eliminating the two linear approximations to zoom the behavior of the error of all the other methods, that are of a much smaller magnitude.

We can combine the information from the simulations and from the Euler Errors integrating the (absolute) Euler errors using the computed distribution. This exercise is a generalization of the Den Haan-Marcet test where instead of using the conditional expectation operator we estimate an unconditional expectation using the population distribution. This integral is a welfare measure of the loss induced by the use of the approximating method over the exact solution.

Table 5.4.1: Integral of the Euler Errors ($x10^{-4}$)

Linear	0.2291
Log-Linear	0.6306
Finite Elements	0.0537
Chebyshev	0.0369
Perturbation 2	0.0481
Perturbation 5	0.0369
Value Function	0.0224

Results are presented in Table 5.4.1.[21] Our interpretation of the numbers reinforces our

[21] We use the distribution from Value Function Iteration. Since the distributions are nearly identical for all

belief that linearization in levels must be preferred over loglinearization for the benchmark calibration of the stochastic neoclassical growth model with leisure and that the performance of perturbation methods is excellent.

Another view of the same information is provided by Figure 5.4.11 where we plot a nonparametric estimate of the marginal distribution of capital around $z = 0$ and the Euler Errors around $z = 0$. This figure allow us to get a feeling of where the stationary distribution is spending time and how big are the Euler errors there.

The problems of linearization are not as much due to the presence of uncertainty but to the curvature of the exact policy functions and second that the loglinear approximation is clearly inferior to a linearization in levels. Even with no uncertainty the Euler Errors of the linear methods (not reported here) are very poor in comparison with the nonlinear procedures.

Figures 5.4.12-5.4.20 display results for the extreme calibration $\tau = 50$ and $\sigma = 0.035$ (again we have changed the capital interval to make it representative). Figure 5.4.12 shows the huge errors of the linear approximation, of order -3 in the relevant parts of the state space. Figure 5.4.13 plots even worse error for the log-linear approximation, of around -2. Figure 5.4.14 shows how Finite Elements still displays a robust and stable behavior over the state space. This result is not a big surprise since the global character of the method allows it to pick the strong nonlinearities induced by high risk aversion and high variance. Chebyshev's performance is also very good and delivers similar accuracies. The perturbations of order 2 and 5 keep their ground and perform relatively well for a while but then, around 40 strongly deteriorate. Value Function Iteration gets a relatively uniform -5. We plot a transversal cut in Figure 5.4.20. This graph summarizes much of the discussion above including the fact that the errors of the perturbations (especially of second order) are not completely competitive against projection methods.

This intuition is reinforced by Table 5.4.2 with the integral of the Euler Errors computed as in the benchmark calibration. From the table we can see two clear winner (Finite Elements and Chebyshev) and a clear loser (log-linear) with the other results somehow in the middle. The poor performance of the 5^{th} order approximation is due to the very quick deterioration of the approximation outside the range of capital between 20 and 45. It is interesting to note

methods, the table is also nearly the same if we use any other distributions. The only caveat is that using that distribution slightly favors the integral from Value Function Iterations.

that the 2^{nd} order perturbation in logs does better than in levels.[22]

Table 5.4.2: Integral of the Euler Errors ($x10^{-4}$)

Linear	7.12
Log-Linear	24.37
Finite Elements	0.34
Chebyshev	0.22
Perturbation 2	7.76
Perturbation 5	8.91
Perturbation 2 (log)	6.47
Value Function	0.32

We finish remarking that the results the four intermediate parametrizations, not included here, did not uncover any non-monoticity of the Euler Errors as they moved in the directions expected when we changed the parameters.

5.5. Implementation and Computing Time

We conclude this section by briefly discussing implementation and computing time. Traditionally (for example Taylor and Uhlig (1990)) computational papers have concentrated in the discussion of the running times of the approximation. Being an important variable, sometimes it is of minor relevance in comparison with the implementation time of an algorithm (i.e. the programming and debugging time). A method that may run in a fraction of a second in a regular PC but requires thousands a line of code may be less interesting than a method that takes a minute but only has a few dozens of lines of code unless we need to repeat the computation once and again (as in an estimation problem). Of course implementing time is a much more subjective measure than running time but we feel that some comments about it are important. In particular we use lines of code as a proxy for the implementation complexity.[23]

[22] We again use the stationary distribution of capital from Value Function Iteration. The results with any other two global nonlinear method are nearly the same (see again figure 5.2.8 where the three distribution are on top of each other).

[23] Being the same programmers across all different methods somehow controls for the fixed effects of stylistic idiosyncrasies. We must repeat however that comparison of lines is only an imperfect measure of complexity and that some correction is required for the relative verbosities of *Matlab*, *Mathematica* and Fortran 95.

The Undetermined Coefficients method (both in level and in logs), computed in *Matlab* 6.5, takes only a fraction of a second in a 1.7 Mhz Xeon PC running Windows XP (the reference computer for all times below), and it is very simple to implement (the code for both methods takes less than 140 lines of code with generous comments). Similar in complexity is the code for the perturbations, only around 64 lines of code in *Mathematica* 4.1 (code that can also compute linearization as a special case) although *Mathematica* is much less verbose than *Matlab*. The code runs in between 2 and 10 seconds depending on the order of the expansion. This observation is the basis of our comment the marginal cost of perturbations over linearizations is close to zero. The Finite Elements method is perhaps the most complicated method to implement: our code in Fortran 95 has above 2000 lines and requires some ingenuity. Running time is moderate, around 20 minutes. Chebyshev polynomials are an intermediate case. The code is much shorter, around 750 lines of Fortran 95. Computation time varies between 20 seconds and 3 minutes but the solution of the system of equations requires some effort searching for an appropriate initial guess which is included in the computation time. Finally, Value Function Iteration code is around 600 lines of Fortran 95 but it takes between 20 and 250 hours to run.

6. Conclusions

In this paper we have compared a set of different solution methods for dynamic equilibrium economies. We have found that perturbation methods are an attractive compromise between accuracy, speed and programming burden but they suffer from the need of computing analytical derivatives and from some instabilities problems for highly nonlinear problems. In any case they must clearly be preferred to linear methods as the simple, default procedure for computing dynamic equilibrium economies. In the case that a linear method is required (for instance if we want to apply the Kalman filter for estimation purposes), the paper suggest that is better to linearize in levels than in logs. The Finite Elements method is a robust, solid method that conserves its accuracy over a long range of the state space even for high values of the risk aversion and the variance of the shock and that is perfectly suited for parallelization and estimation purposes (see also Fernández-Villaverde and Rubio, 2002). However it is costly to implement and moderately intensive in running time. We also found that Chebyshev Polynomials share most of the good properties of Finite Elements and it maybe easier to implement.

We finish by pointing out to several lines of future research. First the accuracy of 2^{nd} order approximations indicates that powerful additional analytical results regarding the stochastic growth model can be obtained extending Campbell's (1994) exercise to the quadratic terms of the policy function. Similarly the results in Williams (2002) suggest that further work integrating perturbation method with small noise asymptotics are promising. Finally we are exploring in a companion paper (Fernández-Villaverde and Rubio-Ramírez, 2003c) the application of newer nonlinear methods as the Adaptive Finite Element method (Verfürth, 1996), the Weighted extended B-splines Finite Element approach (Höllig, 2003) and Element-Free Galerkin Methods (Belytschko *et al.*, 1996) that improve on the basic Finite Elements approach exploiting local information and error estimator values for the elements.

7. Technical Appendix

In this technical appendix we offer some additional details on the implementation of our approximations.

7.1. Undetermined Coefficients in Levels

First we find the deterministic steady state of the model: $k = \frac{\Psi}{\Omega+\varphi\Psi}$, $l = \varphi k, c = \Omega k$ and $y = k^{\alpha}l^{1-\alpha}$ where $\varphi = \left(\frac{1}{\alpha}\left(\frac{1}{\beta} - 1 + \delta\right)\right)^{\frac{1}{1-\alpha}}$, $\Omega = \varphi^{\frac{1}{\alpha}} - \delta$ and $\Psi = \frac{\theta}{1-\theta}(1-\alpha)\varphi^{-\alpha}$.

If we linearize the set of equilibrium conditions around those variables values we get:

$$\frac{\theta(1-\tau)-1}{c}(c_t - c) - \frac{(1-\tau)(1-\theta)}{1-l}(l_t - l) =$$
$$E_t\left\{\begin{array}{c} \frac{\theta(1-\tau)-1}{c}(c_{t+1} - c) + \left(\beta\frac{\alpha(1-\alpha)}{l}k^{\alpha-1}l^{1-\alpha} - \frac{(1-\tau)(1-\theta)}{1-l}\right)(l_{t+1} - l) + \\ \alpha\beta k^{\alpha-1}l^{1-\alpha}z_{t+1} + \beta\frac{\alpha(\alpha-1)}{k}k^{\alpha-1}l^{1-\alpha}(k_{t+1} - k) \end{array}\right\}$$
$$\frac{1}{c}(c_t - c) + \frac{1}{(1-l)}(l_t - l) = z_t + \frac{\alpha}{k}(k_t - k) - \frac{\alpha}{l}(l_t - l)$$
$$(c_t - c) + (k_{t+1} - k) = k^{\alpha}l^{1-\alpha}\left(z_t + \frac{\alpha}{k}(k_t - k) + \frac{1-\alpha}{l}(l_t - l)\right) + (1-\delta)(k_t - k)$$
$$z_t = \rho z_{t-1} + \varepsilon_t$$

or

$$\alpha_1(c_t - c) + \alpha_2(l_t - l) = E_t\{\alpha_1(c_{t+1} - c) + \alpha_3(l_{t+1} - l) + \alpha_4 z_{t+1} + \alpha_5(k_{t+1} - k)\}$$
$$(c_t - c) = cz_t + \frac{\alpha}{k}c(k_t - k) + \alpha_6(l_t - l)$$
$$(c_t - c) + (k_{t+1} - k) = yz_t + y\frac{\alpha}{k}(k_t - k) + \alpha_7(l_t - l) + (1-\delta)(k_t - k)$$
$$z_t = \rho z_{t-1} + \varepsilon_t$$

where

$$\begin{array}{ll} \alpha_1 = \frac{\theta(1-\tau)-1}{c} & \alpha_2 = -\frac{(1-\tau)(1-\theta)}{1-l} \\ \alpha_3 = \beta\frac{\alpha(1-\alpha)}{l}k^{\alpha-1}l^{1-\alpha} - \frac{(1-\tau)(1-\theta)}{1-l} & \alpha_4 = \alpha\beta k^{\alpha-1}l^{1-\alpha} \\ \alpha_5 = \beta\frac{\alpha(\alpha-1)}{k}k^{\alpha-1}l^{1-\alpha} & \alpha_6 = -\left(\frac{\alpha}{l} + \frac{1}{(1-l)}\right)c \\ \alpha_7 = y\frac{1-\alpha}{l} & y = k^{\alpha}l^{1-\alpha} \end{array}$$

We group terms to eliminate one of the equations of the system and obtain the system:

$$A\widehat{k}_{t+1} + B\widehat{k}_t + C\widehat{l}_t + Dz_t = 0$$
$$E_t\left(G\widehat{k}_{t+1} + H\widehat{k}_t + J\widehat{l}_{t+1} + K\widehat{l}_t + Lz_{t+1} + Mz_t\right) = 0$$
$$E_t z_{t+1} = Nz_t$$

where $A=1$, $B=\frac{\alpha}{k}c-y\frac{\alpha}{k}-(1-\delta)$, $C=\alpha_6-\alpha_7$, $D=c-y$, $G=\alpha_1\frac{\alpha}{k}c+\alpha_5$, $H=-\alpha_1\frac{\alpha}{k}c$, $J=\alpha_1\alpha_6+\alpha_3$, $K=-(\alpha_1\alpha_6+\alpha_2)$, $L=(\alpha_1 c+\alpha_4)$, $M=-\alpha_1 c$, $N=\rho$ and $\widehat{x}_t=x_t-x$.

Now we can guess policy functions of the form $\widehat{k}_{t+1}=P\widehat{k}_t+Qz_t$ and $\widehat{l}_t=R\widehat{k}_t+Sz_t$, plug them in and get:

$$A\left(P\widehat{k}_t+Qz_t\right)+B\widehat{k}_t+C\left(R\widehat{k}_t+Sz_t\right)+Dz_t=0$$
$$G\left(P\widehat{k}_t+Qz_t\right)+H\widehat{k}_t+J\left(R\left(P\widehat{k}_t+Qz_t\right)+SNz_t\right)+K\left(R\widehat{k}_t+Sz_t\right)+(LN+M)z_t=0$$

Since these equations need to hold for any value $\widehat{k}_t$ or z_t we need to equate each coefficient to zero, on $\widehat{k}_t$:

$$\begin{aligned} AP+B+CR &= 0 \quad (19)\\ GP+H+JRP+KR &= 0 \quad (20)\end{aligned}$$

and on z_t:

$$\begin{aligned} AQ+CS+D &= 0 \quad (21)\\ (G+JR)Q+JSN+KS+LN+M &= 0 \quad (22)\end{aligned}$$

To solve these system of four equations on four unknowns, we solve for R on (19):

$$R=-\frac{1}{C}(AP+B)=-\frac{1}{C}AP-\frac{1}{C}B$$

and plug in (20) and grouping terms:

$$P^2+\left(\frac{B}{A}+\frac{K}{J}-\frac{GC}{JA}\right)P+\frac{KB-HC}{JA}=0$$

a quadratic equation with solutions:

$$P=-\frac{1}{2}\left(-\left(\frac{B}{A}+\frac{K}{J}-\frac{GC}{JA}\right)\pm\sqrt{\left(\frac{B}{A}+\frac{K}{J}-\frac{GC}{JA}\right)^2-4\left(\frac{KB-HC}{JA}\right)}\right)$$

one associated with the stable saddle path and another with the unstable.

If we pick the stable root and find $R=-\frac{1}{C}(AP+B)$ we reduce (21) and (22) to a system of two linear equations on two unknowns with solution:

$$\begin{aligned} Q &= \frac{-D(JN+K)+CLN+CM}{AJN+AK-CG-CJR}\\ S &= \frac{-ALN-AM+DG+DJR}{AJN+AK-CG-CJR}\end{aligned}$$

completing the solution of the model.

A modification of the procedure would expand the model around some other point to

correct for the difference between the mean of the variables in the stochastic steady state and the deterministic steady state values. A simple algorithm would compute an approximation around the deterministic steady state, simulate the model, find the mean of the variables in the simulation, expand around that mean and iterate until convergence (see Collard and Juillard (2001)). This bias correction procedure is however intensive in time and prone to problems induced by the fact that the linear policy is independent of the variance of the driving stochastic process for the economy no matter where the linearization is performed. For example in our simulations the mean of the simulated capital series was not always higher than the deterministic steady state level of capital and consequently the bias correction procedure might not have any chance of success. Also it is not obvious that the leading term of an asymptotically valid approximation should be taken around that mean point of the stationary distribution of the state variable. As we argued in the main text a simple correction for the first few next terms of the asymptotic expansion performs extremely well for a trivial marginal cost and is to be preferred to bias correction.

7.2. Undetermined Coefficients in Logs

First we substitute each variable x_t by $xe^{\widehat{x}_t}$ where x is the steady state and $\widehat{x}_t = \log \frac{x_t}{x}$ in the model equilibrium equations. After some simplification

$$\frac{\left(\left(ce^{\widehat{c}_t}\right)^{\theta}\left(1-le^{\widehat{l}_t}\right)^{1-\theta}\right)^{1-\tau}}{ce^{\widehat{c}_t}} =$$

$$\beta E_t \frac{\left(\left(ce^{\widehat{c}_{t+1}}\right)^{\theta}\left(1-le^{\widehat{l}_{t+1}}\right)^{1-\theta}\right)^{1-\tau}}{ce^{\widehat{c}_{t+1}}}\left(1+\alpha e^{z_{t+1}}\left(ke^{\widehat{k}_{t+1}}\right)^{\alpha-1}\left(le^{\widehat{l}_{t+1}}\right)^{1-\alpha}-\delta\right)$$

$$\frac{ce^{\widehat{c}_t}}{1-le^{\widehat{l}_t}} = \frac{\theta}{1-\theta}(1-\alpha)e^{z_t}\left(ke^{\widehat{k}_t}\right)^{\alpha}\left(le^{\widehat{l}_t}\right)^{-\alpha}$$

$$ce^{\widehat{c}_t} + ke^{\widehat{k}_{t+1}} = e^{z_t}\left(ke^{\widehat{k}_t}\right)^{\alpha}\left(le^{\widehat{l}_t}\right)^{1-\alpha} + (1-\delta)ke^{\widehat{k}_t}$$

$$z_t = \rho z_{t-1} + \varepsilon_t$$

Loglinearing the conditions delivers:

$$(\theta(1-\tau)-1)\widehat{c}_t - (1-\tau)(1-\theta)\frac{l}{1-l}\widehat{l}_t =$$

$$E_t\left\{\begin{array}{c}(\theta(1-\tau)-1)\widehat{c}_{t+1} - (1-\tau)(1-\theta)\frac{l}{1-l}\widehat{l}_{t+1} + \\ \alpha\beta k^{\alpha-1}l^{1-\alpha}\left(z_{t+1}+(\alpha-1)\widehat{k}_{t+1}+(1-\alpha)\widehat{l}_{t+1}\right)\end{array}\right\}$$

$$\widehat{c}_t + \frac{l}{1-l}\widehat{l}_t = z_t + \alpha\widehat{k}_t - \alpha\widehat{l}_t$$

$$c\widehat{c}_t + k\widehat{k}_{t+1} = k^{\alpha}l^{1-\alpha}\left(z_t + \alpha\widehat{k}_t + (1-\alpha)\widehat{l}_t\right) + (1-\delta)k\widehat{k}_t$$

$$z_t = \rho z_{t-1} + \varepsilon_t$$

or:

$$E_t\left(\alpha_1\widehat{c_{t+1}}-\alpha_2\widehat{l_{t+1}}+\alpha_3 z_{t+1}-\alpha_4\widehat{k_{t+1}}+\alpha_4\widehat{l_{t+1}}-\alpha_1\widehat{c_t}+\alpha_2\widehat{l_t}\right)=0$$

$$\widehat{c_t}+\alpha_5\widehat{l_t}-z_t-\alpha\widehat{k_t}=0$$

$$c\widehat{c_t}+k\widehat{k_{t+1}}-yz_t-\alpha y\widehat{k_t}-y\left(1-\alpha\right)\widehat{l_t}-\left(1-\delta\right)k\widehat{k_t}=0$$

$$z_t=\rho z_{t-1}+\varepsilon_t$$

where

$$\begin{array}{lll}\alpha_1=\left(\theta\left(1-\tau\right)-1\right) & \alpha_2=\left(1-\tau\right)\left(1-\theta\right)\frac{l}{1-l} & \alpha_3=\alpha\beta k^{\alpha-1}l^{1-\alpha}\\ \alpha_4=\alpha_3\left(1-\alpha\right) & \alpha_5=\left(\frac{l}{1-l}+\alpha\right) & y=k^{\alpha}l^{1-\alpha}\end{array}$$

After some algebra the system is reduced to:

$$A\widehat{k}_{t+1}+B\widehat{k}_t+C\widehat{l}_t+Dz_t=0$$

$$E_t\left(G\widehat{k}_{t+1}+H\widehat{k}_t+J\widehat{l}_{t+1}+K\widehat{l}_t+Lz_{t+1}+Mz_t\right)=0$$

$$E_t z_{t+1}=Nz_t$$

where $A=k$, $B=\alpha\left(c-y\right)-\left(1-\delta\right)k$, $C=y\left(\alpha-1\right)-\alpha_5 c$, $D=c-y$, $G=(\alpha_1\alpha-\alpha_4)$, $H=-\alpha_1\alpha$, $J=\alpha_4-\alpha_1\alpha_5-\alpha_2$, $K=\alpha_2+\alpha_1\alpha_5$, $L=\alpha_3+\alpha_1$, $M=-\alpha_1$ and $N=\rho$.

Since the resulting system is equivalent to the previous one in the linearization case, we proceed analogously to solve for the four unknown coefficients.

7.3. Finite Elements Method

The first step in the Finite Elements method is to note that we can rewrite the Euler equation for consumption as

$$U_c(k_t,z_t)=\frac{\beta}{\sqrt{2\pi\sigma}}\int_{-\infty}^{\infty}\left[U_c(k_{t+1},z_{t+1})(1+\alpha e^{z_{t+1}}k_{t+1}^{\alpha-1}l(k_{t+1},z_{t+1})^{1-\alpha}-\delta)\right]\exp(-\frac{\epsilon_{t+1}^2}{2\sigma^2})d\epsilon_{t+1} \tag{23}$$

where $U_c(t)=U_c(k_t,z_t)$, $k_{t+1}=e^{z_{t+1}}k_t^{\alpha}l_t^{1-\alpha}+(1-\delta)k_t-c(k_t,z_t)$ and $z_{t+1}=\rho z_t+\epsilon_{t+1}$.

The problem is to find two policy functions $c(k,z):R^+\times[0,\infty]\rightarrow R^+$ and $l(k,z):R^+\times[0,\infty]\rightarrow[0,1]$ that satisfy the model equilibrium conditions. Since the static first order condition gives a relation between the two policy functions, we only need to solve for one of them. For the rest of the exposition we will assume that we actually solve for $l(k,z)$ and then we find $c\left(l(k,z)\right)$.

First we bound the domain of the state variables to partition it in nonintersecting elements. To bound the productivity level of the economy define $\lambda_t=\tanh(z_t)$. Since $\lambda_t\in[-1,1]$ we can write the stochastic process as $\lambda_t=\tanh(\rho\tanh^{-1}(z_{t-1})+\sqrt{2}\sigma v_t)$ where $v_t=\frac{\epsilon_t}{\sqrt{2}\sigma}$. Now,

since $\exp(\tanh^{-1}(z_{t-1})) = \frac{\sqrt{1+\lambda_{t+1}}}{\sqrt{1-\lambda_{t+1}}} = \widehat{\lambda}_{t+1}$, we rewrite (23) as

$$U_c(t) = \frac{\beta}{\sqrt{\pi}} \int_{-1}^{1} \left[U_c(k_{t+1}, z_{t+1}) \left(1 + \alpha \widehat{\lambda}_{t+1} k_{t+1}^{\alpha-1} l(k_{t+1}, z_{t+1})^{1-\alpha} + \delta \right) \right] \exp(-v_{t+1}^2) dv_{t+1} \quad (24)$$

where $k_{t+1} = \widehat{\lambda}_{t+1} k_t^{\alpha} l\left(k_t, z_t\right)^{1-\alpha} + (1-\delta) k_t - c\left(l(k_t, z_t)\right)$ and $z_{t+1} = \tanh(\rho \tanh^{-1}(z_t) + \sqrt{2}\sigma v_{t+1})$. For convenience we use the same notation for $l\left(\cdot\right)$ in both (23) and (24) although they are not the same function since their domain is different. To bound the capital we fix an ex-ante upper bound $\overline{k}$, picked sufficiently high that it will only bind with an extremely low probability.

Then define $\Omega = \left[0, \overline{k}\right] \times [-1, 1]$ as the domain of $l_{fe}(k, z; \overline{\theta})$ and divide Ω into nonoverlapping rectangles $[k_i, k_{i+1}] \times [z_j, z_{j+1}]$, where k_i is the ith grid point for capital and z_j is jth grid point for the technology shock. Clearly $\Omega = \cup_{i,j} [k_i, k_{i+1}] \times [z_j, z_{j+1}]$. These elements may be of unequal size. In our computations we have small elements in the areas of Ω where the economy will spend most of the time while just a few, big size elements will cover wide areas of the state space infrequently visited.[24]

Next we set $l_{fe}\left(k, z; \overline{\theta}\right) = \sum_{i,j} \overline{\theta}_{ij} \Psi_{ij}\left(k, z\right) = \sum_{i,j} \overline{\theta}_{ij} \widehat{\Psi}_i\left(k\right) \widetilde{\Psi}_j\left(z\right)$ where

$$\widehat{\Psi}_i\left(k\right) = \begin{cases} \frac{k-k_i}{k_{i+1}-k_i} & \text{if } k \in [k_{i-1}, k_i] \\ \frac{k_{i+1}-k}{k_{i+1}-k_i} & \text{if } k \in [k_i, k_{i+1}] \\ 0 \text{ elsewhere} \end{cases} \qquad \widetilde{\Psi}_j\left(z\right) = \begin{cases} \frac{z-z_j}{z_{j+1}-z_j} & \text{if } z \in [z_{j-1}, z_j] \\ \frac{z_{j+1}-z}{z_{j+1}-z_j} & \text{if } z \in [z_j, z_{j+1}] \\ 0 \text{ elsewhere} \end{cases}$$

First, note that $\Psi_{ij}\left(k, z\right) = 0$ if $(k, z) \notin [k_{i-1}, k_i] \times [z_{j-1}, z_j] \cup [k_i, k_{i+1}] \times [z_j, z_{j+1}]$ $\forall i, j$, i.e. the function is 0 everywhere except inside two elements. Second $l_{fe}(k_i, z_j; \overline{\theta}) = \overline{\theta}_{ij}$ $\forall i, j$, i.e. the values of $\overline{\theta}$ specify the values of c_{fe} at the corners of each subinterval $[k_i, k_{i+1}] \times [z_j, z_{j+1}]$.

Let us define $U_c(k_{t+1}, z_{t+1})_{fe}$ be the marginal utility of consumption evaluated at the finite element approximation values of consumption and leisure. In this case, from the Euler equation we have a residual equation:

$$R(k_t, z_t; \theta) = \frac{\beta}{\sqrt{\pi}} \int_{-1}^{1} \left[\frac{U_c(k_{t+1}, z_{t+1})_{fe}}{U_c(k_{t+1}, z_{t+1})_{fe}} \left(1 + \alpha \widehat{\lambda}_{t+1} k_{t+1}^{\alpha-1} l_{fe}^{1-\alpha} - \delta \right) \right] \exp(-v_{t+1}^2) dv_{t+1} - 1 \quad (25)$$

A Galerkin scheme implies that we weight the residual function by the basis functions and solve the system of $\overline{\theta}$ equations

$$\int_{\left[0, \overline{k}\right] \times [-1,1]} \Psi_{i,j}\left(k, z\right) R(k, z; \overline{\theta}) dz dk = 0 \quad \forall i, j \quad (26)$$

on the $\overline{\theta}$ unknowns.

Since $\Psi_{ij}\left(k, z\right) = 0$ if $(k, z) \notin [k_{i-1}, k_i] \times [z_{j-1}, z_j] \cup [k_i, k_{i+1}] \times [z_j, z_{j+1}]$ $\forall i, j$ we can rewrite 26 as

$$\int_{[k_{i-1}, k_i] \times [z_{j-1}, z_j] \cup [k_i, k_{i+1}] \times [z_j, z_{j+1}]} \Psi_{i,j}\left(k, z\right) R(k, z; \overline{\theta}) dz dk = 0 \quad \forall i, j \quad (27)$$

[24] There is a whole area of research concentrated on the optimal generation of an element grid. See Thomson, Warsi and Mastin (1985).

Finally, in order to solve the system we use Gauss-Hermite for the integral in the residual equation and Gauss-Legendre for the integrals in (27) (Press *et al.*, 1992).

We use 71 unequal elements in the capital dimension and 31 on the λ axis. To solve the associated system of 2201 nonlinear equations we use a Quasi-Newton algorithm.

7.4. Spectral Methods

We approximate the decision rules for labor as $l_t = \sum_{i=1}^{n} \theta_i \psi_i (k_t, z_t)$ where $\{\psi_i (k, z)\}_{i=1}^{n}$ are basis functions and $\theta = [\{\theta_i\}_{i=1}^{n}]$ unknown coefficients and use that policy function to solve for consumption using the static first order condition.

We pick Chebyshev Polynomials as our basis functions and build a residual function $R(k, z, \theta)$ using the Euler equation and the static first order condition. Then we choose θ setting a weighted average of the residual function over all possible levels of the state variables equal to zero:

$$\int_{[k_{\min}, k_{\max}]} \int_{[z_{\min}, z_{\max}]} \phi_i (k, z) R(k, z, \theta) = 0 \quad \text{for } i = 1, ..., n \tag{28}$$

where $\{\phi_i (k, z)\}_{i=1}^{n}$ are some weight functions.

We use a collocation method that sets $\phi_i (k, z) = \delta (k - k_j, z - z_v)$ where $\delta (.)$ is the dirac delta function, $j = 1, ..., n_1$, $v = 1, ..., n_2$ and $n = n_1 \times n_2$. The points $\{k_j\}_{j=1}^{n_1}$ and $\{z_v\}_{v=1}^{n_2}$ are called the collocation points. The roots of the n_1^{th} order Chebyshev polynomial[25] as the collocation points for capital. This choice is called orthogonal collocation since the basis functions constitute an orthogonal set. These points are attractive because by the Chebyshev Interpolation Theorem if an approximating function is exact at the roots of the n_1^{th} order Chebyshev polynomial then as $n_1 \to \infty$ the approximation error becomes arbitrarily small. For the technology shock we use Tauchen (1986)'s finite approximation to an AR(1) process and obtain n_2 points. We also use the transition probabilities implied by this approximation to compute the relevant integrals.

Then we have a system of n equations $R(k_i, z_i, \theta) = 0$ in n unknowns θ that we solve using a Quasi-Newton method. Since we had problems to get the system to converge we use an iteration based on the increment of the number of basis functions and a nonlinear transform of the objective function (see Judd (1992)). First we solve a system with only three collocation points for capital (and n_2 points for the technology shock), then we use that solution as a guess for a system with one more collocation point for capital (with the new coefficients being guessed equal to zero), get a new solution and continue in the procedure until we use up to 11 polynomials in the capital dimension and 9 in the productivity axis.

7.5. Perturbation Methods

Beyond the description in the main text we only need to add that we compute the derivatives of the equilibrium conditions of the model using *Mathematica* 4.1, that we feed the deterministic steady state computed above and that we use the `NSolve` function to simultaneously

[25] The roots are given by $x_j = \cos\left\{\frac{\pi[2(n-j+1)-1]}{2n}\right\}$, $j = 1, .., n$.

find for both the coefficients associated with capital and the productivity shock in each order of the perturbation.

7.6. Value Function Iteration

To solve the Bellman operator defined in the main body of the paper we define a grid $G_k \equiv \{k_1, k_2, \ldots, k_M\}$ on k and use the Tauchen's (1986) method to discretize the stochastic process z with $G_z \equiv \{z_1, z_2, \ldots, z_N\}$ and Π_N being the resulting transition matrix with generic element π_N^r

The algorithm to compute the Value function for a given grid is given by:

I. Set $n = 0$ and $V_0(k, z) = \frac{\left(c^\theta (1-l)^{1-\theta}\right)^{1-\tau}}{1-\tau}$.

II. Set $i = 1$.

a. Set $j = 1$ and $r = 1$.

b. For k_i and z_j use the static first order condition to set $c = (1-\alpha)\exp(z_j) k_i^\alpha l^{-\alpha}(1-l)$ and rewrite the resource constraint equation as

$$(1-\alpha)\exp(z_j) k_i^\alpha l^{-\alpha}(1-l) + k' = \exp(z_j) k_i^\alpha l^{1-\alpha} + (1-\delta) k_i \tag{29}$$

1. Set $s = r$ and $U_{i,j}^s = -Inf$.
2. Use Newton method to find l_s that solves

$$(1-\alpha)\exp(z_j) k_i^\alpha l^{-\alpha}(1-l) + k_s = \exp(z_j) k_i^\alpha l^{1-\alpha} + (1-\delta) k_i$$

3. Compute

$$U_{i,j}^s = \frac{\left(\left((1-\alpha)\exp(z_j) k_i^\alpha l_s^{-\alpha}(1-l_s)\right)^\theta (1-l_s)^{1-\theta}\right)^{1-\tau}}{1-\tau} + \beta \sum_{r=1}^{N} \pi_N^r V_n(k_s, z_r)$$

4. If $U_{i,j}^{s-1} \leq U_{i,j}^s$, then $s \rightsquigarrow s+1$ and go to ii.
5. Define

$$U(k; k_i, z_j) = \frac{\left(\left((1-\alpha)\exp(z_j) k_i^\alpha l^{-\alpha}(1-l)\right)^\theta (1-l)^{1-\theta}\right)^{1-\tau}}{1-\tau} + \beta \sum_{r=1}^{N} \pi_N^r \widehat{V}_n(k, z_r)$$

for $k \in [k_{s-2}, k_s]$, where l solves

$$(1-\alpha)\exp(z_j) k_i^\alpha l^{-\alpha}(1-l) + k = \exp(z_j) k_i^\alpha l^{1-\alpha} + (1-\delta) k_i$$

and $\widehat{V}_n(k, z_r)$ is computed using Lagrangian interpolation.
6. Calculate $\arg\max U(k; k_i, z_j)$.
7. Set r such that $k_{i,j}^* \in [k_s, k_{s+1}]$ and $V_{n+1}(k_i, z_j) = TV_n(k_i, z_j) = U\left(k_{i,j}^*; k_i, z_j\right)$.

c. If $j < N$, then $j \rightsquigarrow j+1$ and go to b.

III. If $i < N$, $i \rightsquigarrow i+1$ and go to a.

IV. If $\sup_{i,j} |V_{n+1}(k_i, z_j) - V_n(k_i, z_j)| / V_n(k_i, z_j) \geq 1.0e^{-8}$, then $n \rightsquigarrow n+1$ and go to II.[26]

To accelerate convergence we use a multigrid scheme (see Chow and Tsitsiklis (1991) and Rüde (1993)). That scheme begins computing the value function in a small grid, refines the grid with more points (with linear interpolation to fill the unknown values) and recomputes the value function. Iterating with this procedure we move from an initial small grid (8000 points) into a final one with one million points.

We interpolate using linear, quadratic and Schumaker (1983) schemes. Results were very similar with all three methods. Our intuition is that the final grid was so fine that how interpolation was done did not really matter. If anything Schumaker performed slightly worse than linear because the need to find the numerical derivative of the value function overtook the advantage of additional curvature. Consequently the results in the paper are those with linear interpolation.

[26] We also monitored convergence in the policy function that was much quicker.

References

[1] Belytschko T., Y. Krongauz, D. Organ, M. Fleming, and P. Krysl (1996), "Meshless Methods: An Overview and Recent Developments". *Computer Methods in Applied Mechanics and Engineering* 139, 3-47.

[2] Bender, C.M. and S.A. Orszag (1999), *Advanced Mathematical Methods for Scientists and Engineers: Asymptotic Methods and Perturbation Theory*. Springer Verlag.

[3] Benítez-Silva, H., G. Hall, G.J. Hitsch, G. Pauletto and J. Rust (2000), "A Comparison of Discrete and Parametric Approximation Methods for Continuous-State Dynamic Programming Problems". *Mimeo*, SUNY at Stony Brook.

[4] Briggs, W.L., V.E. Henson and S.F. McCormick (2000), *A Multigrid Tutorial*, Second Edition. Society for Industrial and Applied Mathematics.

[5] Blanchard, O.J. and C.M. Kahn (1980), "The Solution of Linear Difference Models under Linear Expectations". *Econometrica* 48, 1305-1311.

[6] Boyd, J.P. (2001). *Chebyshev and Fourier Spectral Methods*. Second Edition. Dover Publications.

[7] Briggs, W.L., V.E. Henson and S.F. McCormick (2000), *A Multigrid Tutorial*, Second Edition. Society for Industrial and Applied Mathematics.

[8] Campbell, J.Y. (1994), "Inspecting the Mechanism: An Analytical Approach to the Stochastic Growth Model". *Journal of Monetary Economics*, 33, 463-506.

[9] Christiano, L.J. (1990), "Linear-Quadratic Approximation and Value-Function Iteration: A Comparison". *Journal of Business Economics and Statistics* 8, 99-113.

[10] Christiano, L.J. and J.D.M. Fisher (2000), "Algorithms for Solving Dynamic Models with Occasionally Binding Constraints", *Journal of Economic Dynamics and Control* 24, 1179-1232

[11] 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.

[12] Coleman, W.J. (1990), "Solving the Stochastic Growth Model by Policy-Function Iteration". *Journal of Business and Economic Statistics* 8, 27-30.

[13] Collard, F. and M. Juillard (2001), "Perturbation Methods for Rational Expectations Models". *Mimeo*, CEPREMAP.

[14] Cooley, T.F. (1995), *Frontiers of Business Cycle Research*. Princeton University Press.

[15] Cooley, T.F. and E.C. Prescott (1995), "Economic Growth and Business Cycles" in T. F. Cooley. (ed), *Frontiers of Business Cycle Research*. Princeton University Press.

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

[17] Fernández-Villaverde, J. and J. Rubio-Ramírez (2002), "Estimating Nonlinear Dynamic Equilibrium Economies: A Likelihood Approach". *Mimeo,* University of Pennsylvania. Available at `www.econ.upenn.edu/~jesusfv`

[18] Fernández-Villaverde, J. and J. Rubio-Ramírez (2003a), "Comparing Dynamic Equilibrium Models to Data". *Journal of Econometrics,* forthcoming.

[19] Fernández-Villaverde, J. and J. Rubio-Ramírez (2003b), "Some Results on the Solution of the Neoclassical Growth Model". *Mimeo,* University of Pennsylvania.

[20] Fernández-Villaverde, J. and J. Rubio-Ramírez (2003c), "Refining the Finite Elements Method for Dynamic Equilibrium Economies". *Mimeo,* University of Pennsylvania.

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

[22] Geweke, J. (1996), "Monte Carlo Simulation and Numerical Integration" in H. Amman, D. Kendrick and J. Rust (eds), *Handbook of Computational Economics*. Elsevier-North Holland.

[23] Höllig, K. (2003). *Finite Element Methods with B-Splines*. Society for Industrial and Applied Mathematics.

[24] Hughes, T.R.J. (2000). *The Finite Element Method, Linear Static and Dynamic Finite Element Analysis*. Dover Publications.

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

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

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

[28] Judd, K.L. (2003), "Perturbation Methods with Nonlinear Changes of Variables". *Mimeo*, Hoover Institution.

[29] King, R.G., C.I. Plosser and S.T. Rebelo (2002), "Production, Growth and Business Cycles: Technical Appendix ". *Computational Economics* 20, 87-116

[30] Klein, P. (2000), "Using the Generalized Schur Form to Solve a Multivariate Linear Rational Expectations Model", *Journal of Economic Dynamics and Control* 24(10), 1405-1423.

[31] Kydland, F.E. (1989), "Monetary Policy in Models with Capital" in F. van der Ploeg and A.J. de Zeuw (eds), *Dynamic Policy Games in Economies*. North-Holland

[32] Kydland, F.E. and E.C. Prescott (1982), "Time to Build and Aggregate Fluctuations". *Econometrica* 50, 1345-1370.

[33] Marcet, A. and G. Lorenzoni (1999), "The Parametrized Expectations Approach: Some Practical Issues" in R. Marimon and A. Scott (eds), *Computational Methods for the Study of Dynamic Economies*. Oxford University Press.

[34] Marimón, R. and A. Scott (1999), *Computational Methods for the Study of Dynamic Economies*. Oxford University Press.

[35] McGrattan, E.R. (1999), "Application of Weigthed Residuals Methods to Dynamic Economic Models" in R. Marimon and A. Scott (eds), *Computational Methods for the Study of Dynamic Economies*. Oxford University Press.

[36] McGrattan, E. and E.C. Prescott (2000), "Is the Stock Market Overvalued?". *Mimeo*, Federal Reserve Bank of Minneapolis.

[37] Newey, W. and K.D. West (1987), "A Simple, Positive, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix". *Econometrica*, 55, 703-705.

[38] Press, W.H., S.A. Teukolsky, W.T. Vetterling and B.P. Flannery (1996). *Numerical Recipes in Fortran 77: The Art of Scientific Computing*. Cambridge University Press.

[39] Rüde, U. (1993). *Mathematical and Computational Techniques for Multilevel Adaptive Methods*. Society for Industrial and Applied Mathematics.

[40] Rust, J. (1996), "Numerical Dynamic Programming in Economics" in H. Amman, D. Kendrick and J. Rust (eds), *Handbook of Computational Economics*. Elsevier-North Holland.

[41] Rust, J. (1997), "Using Randomization to Break the Curse of Dimensionality". *Econometrica* 65, 487-516.

[42] Santos, M.S. (1999), "Numerical Solution of Dynamic Economic Models" in J.B. Taylor and M. Woodford (*eds*), *Handbook of Macroeconomics*, volume 1a, North Holland.

[43] Santos, M.S. (2000), "Accuracy of Numerical Solutions Using the Euler Equation Resituduals". *Econometrica* 68, 1377-1402.

[44] Santos, M.S. and J. Vigo (1998). "Analysis of Error for a Dynamic Programming Algorithm". *Econometrica* 66, 409-426.

[45] Sargent, T.J. (1987), *Dynamic Macroeconomic Theory*. Harvard University Press.

[46] Schumaker, L.L. (1983), "On Shape-Preserving Quadratic Spline Interpolation". *SIAM Journal of Numerical Analysis* 20, 854-864.

[47] Sims, C.A. (2000), "Second Order Accurate Solution of Discrete Time Dynamic Equilibrium Models". *Mimeo*, Princeton University.

[48] Sims, C.A. (2002), "Solving Linear Rational Expectations Models ". *Computational Economics* 20, 1-20.

[49] Schmitt-Grohé, S. and M. Uribe (2002), "Solving Dynamic General Equilibrium Models Using a Second-Order Approximation to the Policy Function". *NBER Technical Working Paper* 282.

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

[51] Taylor, J.B and H. Uhlig (1990), "Solving Nonlinear Stochastic Growth Models: A Comparison of Alternative Solution Methods". *Journal of Business Economics and Statistics* 8, 1-17.

[52] Thomson, J.F., Z.U.A. Warsi and C.W. Mastin (1985). *Numerical Grid Generation: Foundations and Applications*. Elsevier.

[53] Uhlig, H. (1999), "A Toolkit for Analyzing Nonlinear Dynamic Stochastic Models Easily" in R. Marimón and A. Scott (eds) *Computational Methods for the Study of Dynamic Economies*. Oxford University Press.

[54] Verfürth, R. (1996). *Posteriori Error Estimation and Adaptive Mesh-Refinement Techniques*. Wiley-Teubner.

[55] Williams, N. (2002), "Small Noise Asymptotics for a Stocastic Growth Model". *Mimeo*, Princeton University.