Computational Methods for Production-Based Asset Pricing Models with Recursive Utility

Eric M. Aldrich *
Department of Economics
Duke University

Howard Kung
Fuqua School of Business
Duke University

First draft: 15 March 2009
This draft: 12 November 2011

Abstract

We compare local and global polynomial solution methods for DSGE models with Epstein-Zin-Weil utility. We show that model implications for macroeconomic quantities are relatively invariant to choice of solution method but that a global method can yield substantial improvements for asset prices and welfare costs. The divergence in solution quality is highly dependent on parameters which effect value function sensitivity to TFP volatility, as well as the magnitude of TFP volatility itself. This problem is pronounced for calibrations at the extreme of those accepted in the asset pricing literature and disappears for more traditional macroeconomic parameterizations.

Keywords: DSGE Models, Nonlinear Solution Methods, Numerical Dynamic Programming, Recursive Utility, Asset Pricing.

JEL Classification: C63, C68, D53, E44, G12.

*Corresponding Author. Email: ealdrich@gmail.com; Phone: 919-638-6216. We are indebted to Juan Rubio-Ramirez for his guidance during this project and to both Juan Rubio-Ramirez and Jesus Fernandez-Villaverde for the use of their perturbation code. We would like to thank participants at the 2008 International Conference on Computing in Economics and Finance as well as the 2009 Joint Statistical Meetings for useful comments and criticisms. Finally, we are very grateful for insights provided by George Tauchen, Craig Burnside, Ron Gallant, Tim Bollerslev and other participants in the Financial Econometrics workshop at Duke as this work has progressed.

Introduction

This paper compares solution methods for production-based asset pricing models with recursive utility. In particular, we consider a standard neoclassical growth model with Epstein-Zin-Weil utility. Bansal and Yaron (2004) demonstrates the importance of recursive utility for explaining both asset prices and macroeconomic variables in a general equilibrium endowment setting. More recent extensions to production economies, e.g. Croce (2006), Campanale et al. (2010), Kaltenbrunner and Lochstoer (2010) and Rudebusch and Swanson (2008), show similar promise in reconciling financial and quantity dynamics and are becoming increasingly common in the literature. Accordingly, it is important to understand the behavior of solution methods for these models, and our work contributes to the literature along that dimension.

We focus on two standard methods of solution for dynamic stochastic general equilibrium (DSGE) models: a global projection method using Chebyshev basis functions and a local perturbation method of various orders. Aruoba et al. (2006) compares these solution methods for a neoclassical growth model with constant relative risk aversion utility, and demonstrates that both methods are broadly suitable. Our aim is to investigate the appropriateness of these solution methods when utility takes a more general, recursive form. Notably, we find that the two solution methods produce roughly equivalent implications for macroeconomic quantities, but can be very different with respect to asset prices and welfare costs. This discrepancy is exacerbated as the volatility of the total factor of productivity (TFP) is increased.

Several considerations drive our results. First, the value function corresponding to Epstein-Zin-Weil utility exhibits a high degree of curvature with respect to the TFP volatility; second, perturbation is a local Taylor approximation around the deterministic steady state, where the TFP volatility is zero; and third, in the presence of recursive utility, asset prices and welfare costs, not quantities, depend critically on the shape of the value function. The upshot is that for models with high output volatility, perturbation attempts to approximate a rapidly changing function at a locus (zero TFP volatility) that is far from the point of interest (high TFP volatility). Taylor's Theorem shows that the resulting approximation

error can be arbitrarily bad and can even increase with the order of approximation. Thus, when the TFP is calibrated to have high volatility, the resulting approximation of the value function obtained by perturbation can be very bad, yielding poor approximations for prices and welfare costs.

In addition to the direct effect of TFP volatility on approximation quality, we investigate the effect of other parameters. We find that the subjective discount factor, output growth and risk sensitivity are also important for solution accuracy, but in a more indirect manner: the solution methods diverge when these parameters increase the value function's sensitivity to TFP volatility. We describe the interplay between these parameters and solution accuracy.

Our problem dictates that, at a minimum, we must approximate the value function and policy function of the endogenous choice variable (consumption). In theory, if our solutions are sufficiently accurate, we can use them to accurately approximate any other variable in the model, even if it depends on the value and policy functions in a nonlinear way. The reverse is also true – a poor solution for *either* the value or policy functions will contaminate approximations of other model quantities. However, in the case of perturbation, we find that computing a local approximation of asset prices directly can ameliorate their dependency on the (poorly approximated) value function and improve their accuracy. We show how this modification can bring the model's asset pricing implications far closer to those produced by a global method. Unfortunately, the same result does not hold for welfare costs.

Caldara et al. (2009) compares these solution methods for a similar model with recursive utility. Their results suggest that while global methods are numerically superior, local methods are very accurate and strike an excellent balance between computation time and accuracy. Our results diverge from theirs because we consider a more broad set of parameter values - values which are often used in the asset pricing literature. That is, we are careful to stress that the aforementioned results depend critically on the calibrated level of TFP volatility. Hence, many asset pricing models which calibrate to annual, pre-war data will be badly approximated with a perturbation method. On the other hand, models which specify much smaller values of output volatility (e.g. macroeconomic models which calibrate to quarterly, post-war data, such as Caldara et al. (2009)) are likely to find perturbation to be an adequate, and even preferred solution method.

When perturbation is accurate, it is often the preferred solution method since it is extremely fast and can be efficiently nested within an estimation framework. For the particular problem we consider, the projection and perturbation methods are essentially equivalent in computation time, as the model only entertains one state variable. However, if the complexity of the model is increased (more state variables added), a global projection method would quickly deteriorate in computing time due to the curse of dimensionality. This curse is the relative disadvantage of a global method.

The results of this paper suggest that when using perturbation for production-based asset pricing models, it would be wise to compare the results with a global solution method. In particular, for models with a high calibration of TFP volatility, a global method, while computationally more burdensome, is more suitable since it yields a superior approximant to the value function.

This paper proceeds as follows. Section 1 describes our model in detail, Section 2 outlines the solution methods applied to the model, Section 3 reports the results of the two solution methods in addition to diagnostics that explore their accuracy and Section 4 concludes.

1 Model

We follow Kaltenbrunner and Lochstoer (2010) in specifying a basic neoclassical growth model with Epstein-Zin-Weil utility. As mentioned in the previous section, this model is very similar to that of Croce (2006), the primary difference being that Croce (2006) explicitly specifies a long-run risk component in the total factor of productivity process. While Croce (2006) does a better job of explaining both quantity and price dynamics, and while the model of Kaltenbrunner and Lochstoer (2010) exhibits several deficiencies with respect to asset prices, we choose the latter specification for two reasons. First, it is a more simple generalization of the neoclassical growth model, obtained by substituting Epstein-Zin-Weil utility for power utility, and nests the more standard model. Second, through appropriate normalization, the model has only one state variable. This simplicity allows us to emphasize the computational results with greater ease. We anticipate that the results will extend to more complicated, and perhaps appealing, situations.

1.1 Preferences

Our economy admits a representative agent whose utility function follows Epstein and Zin (1989) and Weil (1990):

$$U\left(\overline{C}_t\right) = \left((1-\beta)C_t^{\frac{1-\gamma}{\theta}} + \beta(\mathbb{E}_t[U\left(\overline{C}_{t+1}\right)^{1-\gamma}])^{\frac{1}{\theta}}\right)^{\frac{\theta}{1-\gamma}}, \tag{1}$$

where $0 < \beta < 1$ is the subjective discount factor, $\mathbb{E}_t$ is the conditional expectations operator, C_t denotes aggregate consumption, $\overline{C}_t = (C_t, C_{t+1}, \ldots)$, γ denotes the agent's coefficient of relative risk aversion, ψ denotes the agent's inter-temporal elasticity of substitution (IES), and $\theta = \frac{1-\gamma}{1-1/\psi}$. A particularly desirable feature of this utility function is that it separates the IES and risk aversion parameters, as opposed to the standard constant relative risk aversion (CRRA) utility, where IES and risk aversion are inversely related. In theory, it is not clear that there should be a tight link between these two parameters, as risk aversion is atemporal while IES is temporal.

As shown in Epstein and Zin (1989) and Weil (1989), the log of the stochastic discount factor, m_{t+1}, for these preferences is

$$m_{t+1} = \theta \ln \beta - \frac{\theta}{\psi}\Delta c_{t+1} - (1-\theta)r_{a,t+1}, \tag{2}$$

where Δc_{t+1} denotes log consumption growth and $r_{a,t+1}$ denotes the log gross return on the aggregate wealth portfolio. Of particular importance is the presence of $r_{a,t+1}$ in the specification of the discount factor, which makes innovations to expected consumption growth a priced risk factor; in the standard model with CRRA utility, $\theta = 1$ and the last term disappears. It is this feature of Epstein-Zin-Weil utility that makes agents concerned about shocks to expected future consumption growth and that allows us to amplify the equity premium while evading the risk-free rate puzzle (Mehra and Prescott (1985)).

1.2 Technology

A single firm owns the capital stock and produces a consumption good via Cobb-Douglass technology, using labor and capital as inputs:

$$Y_t = (Z_t H_t)^{(1-\alpha)} K_t^{\alpha},$$

where Z_t is the stochastic total factor of productivity, H_t denotes the number of hours worked, K_t represents capital and α is the share of capital in the production function. Under the assumption that $\overline{H}$ is the agent's total leisure endowment, it is clear that utility is going to be maximized when $H_t = \overline{H}$, $\forall t$, since H_t does not appear in the utility function. Normalizing $\overline{H} = 1$, the production function simplifies to

$$Y_t = Z_t^{(1-\alpha)} K_t^{\alpha}. \tag{3}$$

The log technology process, $z_t = \ln(Z_t)$, evolves exogenously according to

$$z_t = \mu t + \tilde{z}_t, \tag{4}$$

$$\tilde{z}_t = \varphi \tilde{z}_{t-1} + \sigma_z \epsilon_t, \tag{5}$$

$$\epsilon_t \sim N(0, 1). \tag{6}$$

We limit ourselves to the special case of $\varphi = 1$, since this allows us to retain only one state variable; our computational results are similar for the case of persistent, yet trend stationary z_t (when $|\varphi| < 1$). Hence, the log of TFP is a random walk with drift parameter μ, and in this special case shocks to technology are permanent.

1.3 Capital Accumulation

Following Jermann (1998), we allow capital adjustment costs in the accumulation equation

$$K_{t+1} = \phi\left(\frac{I_t}{K_t}\right) K_t + (1 - \delta) K_t, \tag{7}$$

where

$$\phi(x) = \frac{\alpha_1}{1 - 1/\xi} x^{1-1/\xi} + \alpha_2 \tag{8}$$

is an increasing, concave function which induces large changes in the capital stock to be more costly than successive small changes. The parameter ξ governs the degree of concavity and has the desirable feature that as $\xi \to \infty$, $\phi(x)$ becomes the identity function (with an appropriate specification of α_1 and α_2); that is, capital adjustment costs disappear. At the other extreme, as $\xi \to 0$, $I_t \to 0$, $\forall t$, and $\phi(x) \equiv \exp(\mu) - 1 + \delta$, allowing us to

obtain an endowment economy where all output is consumed each period and the capital stock grows deterministically at rate $\exp(\mu)$. Hence, for intermediate values, the adjustment cost parameter, ξ, allows us flexibility in matching the relative volatilities of consumption and output. As mentioned above, the remaining parameters are defined so as to eliminate adjustment costs in the deterministic steady state: $\alpha_1 = (\exp(\mu) - 1 + \delta)^{1/\xi}$ and $\alpha_2 = \frac{1}{1-\xi}(\exp(\mu) - 1 + \delta)$ (see Appendix A for a derivation).

1.4 Equilibrium

In equilibrium, the aggregate resource constraint is binding:

$$C_t + I_t = Y_t. \tag{9}$$

In this basic environment, the welfare theorems are satisfied and the solution to the social planner's problem yields the same allocations as a competitive equilibrium. The planner's problem is

$$V(K_t, Z_t) = \max_{C_t, K_{t+1}} W(C_t, K_{t+1})^{\frac{\theta}{1-\gamma}} \tag{10a}$$

subject to

$$K_{t+1} = \phi\left(\frac{Z_t^{1-\alpha}K_t^{\alpha} - C_t}{K_t}\right)K_t + (1-\delta)K_t, \tag{10b}$$

where

$$W(C_t, K_{t+1}) = (1-\beta)C_t^{\frac{1-\gamma}{\theta}} + \beta\left(\mathbb{E}_t\left[V(K_{t+1}, Z_{t+1})^{1-\gamma}\right]\right)^{\frac{1}{\theta}}. \tag{11}$$

Hence, the planner maximizes

$$\mathcal{L} = W(C_t, K_{t+1})^{\frac{\theta}{1-\gamma}} + \lambda_t\left(\phi\left(\frac{I_t}{K_t}\right)K_t + (1-\delta)K_t - K_{t+1}\right) \tag{12}$$

over C_t and K_{t+1}, where

$$Z_t = Z_{t-1}\exp(\mu + \sigma_z\varepsilon_t), \;\; \varepsilon_t \sim N(0,1). \tag{13}$$

The resulting first order conditions are

$$\frac{\partial \mathcal{L}}{\partial C_t} = (1-\beta)\check{W}_t^{\frac{1}{\psi-1}} \check{C}_t^{-\frac{1}{\psi}} - \check{\lambda}_t \phi'\left(\frac{\check{I}_t}{K_t}\right) = 0 \tag{14a}$$

$$\begin{aligned}\frac{\partial \mathcal{L}}{\partial K_{t+1}} &= \beta \check{W}_t^{\frac{1}{\psi-1}} \left(\mathbb{E}_t\left[\check{V}_{t+1}^{1-\gamma}\right]\right)^{\frac{1}{\theta}-1} \mathbb{E}_t\left[\check{V}_{t+1}^{-\gamma} \frac{\partial V_{t+1}}{\partial K_{t+1}}\bigg|_{\check{K}_{t+1}}\right] - \check{\lambda}_t \\ &= -\check{\lambda}_t + \beta \check{W}_t^{\frac{1}{\psi-1}} \left(\mathbb{E}_t\left[\check{V}_{t+1}^{1-\gamma}\right]\right)^{\frac{1}{\theta}-1} \\ &\quad \times \mathbb{E}_t\left[\check{V}_{t+1}^{-\gamma} \check{\check{\lambda}}_{t+1}\left(\phi'\left(\frac{\check{\check{I}}_{t+1}}{\check{K}_{t+1}}\right)\frac{(\alpha-1)\check{Y}_{t+1} + \check{\check{C}}_{t+1}}{\check{K}_{t+1}} \right.\right. \\ &\quad \left.\left. + \phi\left(\frac{\check{\check{I}}_{t+1}}{\check{K}_{t+1}}\right) + 1 - \delta\right)\right] = 0 \end{aligned} \tag{14b, 14c}$$

$$\frac{\partial \mathcal{L}}{\partial \lambda_t} = \phi\left(\frac{\check{I}_t}{K_t}\right) K_t + (1-\delta)K_t - \check{K}_{t+1} = 0, \tag{14d}$$

where $\check{C}_t$, $\check{K}_{t+1}$ and $\check{\lambda}_t$ are the optimal values of C_t, K_{t+1} and λ_t, $\check{\check{C}}_{t+1} = \check{C}_{t+1}|_{\check{K}_{t+1}}$, $\check{\check{\lambda}}_{t+1} = \check{\lambda}_{t+1}|_{\check{K}_{t+1}}$, $\check{Y}_{t+1} = Z_{t+1}^{1-\alpha}\check{K}_{t+1}^{\alpha}$, $\check{I}_t = Y_t - \check{C}_t$, $\check{\check{I}}_{t+1} = \check{Y}_{t+1} - \check{\check{C}}_{t+1}$, $V_t = V(K_t, Z_t)$, $\check{V}_{t+1} = V(\check{K}_{t+1}, Z_{t+1})$, $\check{W}_t = W(\check{C}_t, \check{K}_{t+1})$ and where Equation (14c) follows from (14b) by the envelope theorem:

$$\frac{\partial V_t}{\partial K_t} = \frac{\partial \mathcal{L}}{\partial K_t}\bigg|_{\check{C}_t, \check{K}_{t+1}, \check{\lambda}_t} = \check{\lambda}_t\left(\phi'\left(\frac{\check{I}_t}{K_t}\right)\frac{(\alpha-1)Y_t + \check{C}_t}{K_t} + \phi\left(\frac{\check{I}_t}{K_t}\right) + 1 - \delta\right). \tag{15}$$

Substituting Equation (14a) into (14c) and recognizing that $\check{W}_{t+1}|_{\check{K}_{t+1}} = \check{V}_{t+1}^{\frac{1-\gamma}{\theta}}$, we obtain the intertemporal Euler equation

$$\mathbb{E}_t\left[M_{t+1}\phi'\left(\frac{\check{I}_t}{K_t}\right)\left(\frac{(\alpha-1)\check{Y}_{t+1} + \check{\check{C}}_{t+1}}{\check{K}_{t+1}} + \frac{\phi\left(\frac{\check{\check{I}}_{t+1}}{\check{K}_{t+1}}\right) + 1 - \delta}{\phi'\left(\frac{\check{\check{I}}_{t+1}}{\check{K}_{t+1}}\right)}\right)\right] = 1 \tag{16}$$

where

$$M_{t+1} = \beta\left(\frac{\check{\check{C}}_{t+1}}{\check{C}_t}\right)^{-\frac{1}{\psi}} \frac{\check{V}_{t+1}^{\frac{1}{\psi}-\gamma}}{\left(\mathbb{E}_t\left[\check{V}_{t+1}^{1-\gamma}\right]\right)^{1-\frac{1}{\psi}}} \tag{17}$$

is an alternative expression for the Epstein-Zin-Weil stochastic discount factor, equivalent to $\exp(m_{t+1})$ in Equation (2). As shown in Appendix B, the return on equity is

$$R_{t+1}^E = \phi'\left(\frac{I_t}{K_t}\right)\left(\frac{(\alpha-1)Y_{t+1} + \check{C}_{t+1}}{K_{t+1}} + \frac{\phi\left(\frac{\check{I}_{t+1}}{K_{t+1}}\right) + 1 - \delta}{\phi'\left(\frac{\check{I}_{t+1}}{K_{t+1}}\right)}\right), \tag{18}$$

which is the term scaling the stochastic discount factor in Equation (16). Hence, the Euler equation can be written compactly as $\mathbb{E}_t[M_{t+1}\check{R}^E_{t+1}] = 1$, for an appropriate definition of $\check{R}^E_{t+1}$ in terms of optimal values. This expression for the return to equity will be useful when evaluating the quality of model solutions.

To preserve stationarity in the economy, we normalize all variables by the level of the contemporaneous technology process:

$$\left\{\widehat{C}_t, \widehat{K}_t, \widehat{Z}_{t+1}, \widehat{I}_t, \widehat{Y}_t, \widehat{V}_t\right\} = \left\{\frac{C_t}{Z_t}, \frac{K_t}{Z_t}, \frac{Z_{t+1}}{Z_t}, \frac{I_t}{Z_t}, \frac{Y_t}{Z_t}, \frac{V_t}{Z_t}\right\}. \tag{19}$$

Alternatively, we could have normalized by Z_{t-1}, but the former specification results in only one state variable – a feature that allows us to place emphasis on the computational results. The normalized system of equilibrium conditions can then be expressed as (suppressing the $\check{x}$ notation for optimal values of x)

$$\widehat{V}(\widehat{K}_t) - \left((1-\beta)\widehat{C}_t^{\frac{1-\gamma}{\theta}} + \beta\left(\mathbb{E}_t\left[\widehat{Z}_{t+1}^{1-\gamma}\,\widehat{V}(\widehat{K}_{t+1})^{1-\gamma}\right]\right)^{\frac{1}{\theta}}\right)^{\frac{\theta}{1-\gamma}} = 0 \tag{20a}$$

$$\mathbb{E}_t\left[M_{t+1}\,\phi'\left(\frac{\widehat{I}_t}{\widehat{K}_t}\right)\left(\frac{(\alpha-1)\widehat{Y}_{t+1} + \widehat{C}_{t+1}}{\widehat{K}_{t+1}} + \frac{\phi\left(\frac{\widehat{I}_{t+1}}{\widehat{K}_{t+1}}\right) + 1 - \delta}{\phi'\left(\frac{\widehat{I}_{t+1}}{\widehat{K}_{t+1}}\right)}\right)\right] - 1 = 0 \tag{20b}$$

$$\widehat{K}_{t+1} - \frac{1}{\widehat{Z}_{t+1}}\left((1-\delta)\widehat{K}_t + \phi\left(\frac{\widehat{I}_t}{\widehat{K}_t}\right)\widehat{K}_t\right) = 0 \tag{20c}$$

where

$$\widehat{Z}_t = \exp(\mu + \sigma_z\varepsilon_t), \;\; \varepsilon_t \sim N(0,1). \tag{21}$$

2 Solution Methods

We now describe the two methods we use to solve the model of the previous section.

2.1 Perturbation

Perturbation methods, suggested for economic models by Judd and Guu (1997) and Judd (1998), and widely popularized by Schmitt-Grohé and Uribe (2004), build an asymptotically

valid polynomial approximation of a function around a point where the solution is known. In general notation, perturbation seeks a local approximation to a function, F, where

$$F(\boldsymbol{x}(\varepsilon), \varepsilon) = 0, \tag{22}$$

and where $F(\boldsymbol{x}(0), 0)$ is known. The typical specialization in economics is for F to represent a system of nonlinear stochastic difference equations,

$$F(\boldsymbol{x}(\varepsilon), \varepsilon) = \mathbb{E}_t[f(\boldsymbol{x}_{t+1}(\varepsilon), \boldsymbol{x}_t(\varepsilon), \varepsilon)] = 0, \tag{23}$$

where the deterministic steady state, $\mathbb{E}_t[f(\boldsymbol{x}_{t+1}(0), \boldsymbol{x}_t(0), 0)] = f(\boldsymbol{x}_{ss}, \boldsymbol{x}_{ss}, 0)$ is known. The canonical economic example is the neoclassical growth model, where f is a system of equations including the inter-temporal Euler equation and constraints, and where the polynomial approximation to f is a Taylor expansion. However, there is no *a priori* reason to restrict our attention to the inter-temporal Euler equation; since we are interested in computing financial moments and since bond prices in a recursive utility model depend on the value function, it is natural for us to approximate the value function directly. Judd and Guu (1997) and Judd (1998) are early examples of using the value function to generate perturbation conditions and van Binsbergen et al. (2008) provides an argument for this approach. Our particular solution method utilizes both the value function and the intertemporal Euler equation.

We use system (20) to build approximations of the value and policy functions:

$$\widetilde{V}_{pert}(\widehat{K}, \sigma_z) = \sum_{i,j} \widehat{V}_{ss}^{(i,j)} (\widehat{K} - \widehat{K}_{ss})^i \sigma_z^j \tag{24a}$$

$$\widetilde{C}_{pert}(\widehat{K}, \sigma_z) = \sum_{i,j} \widehat{C}_{ss}^{(i,j)} (\widehat{K} - \widehat{K}_{ss})^i \sigma_z^j \tag{24b}$$

where

$$\widehat{V}_{ss}^{(i,j)} = \left(\frac{1}{i!j!}\right) \left.\frac{\partial^{i+j} \widehat{V}_t}{\partial^i \widehat{K}_t \partial^j \sigma_z}\right|_{\widehat{K}_{ss},0} \tag{24c}$$

and

$$\widehat{C}_{ss}^{(i,j)} = \left(\frac{1}{i!j!}\right) \left.\frac{\partial^{i+j} \widehat{C}_t}{\partial^i \widehat{K}_t \partial^j \sigma_z}\right|_{\widehat{K}_{ss},0}. \tag{24d}$$

To obtain these approximations, we take successive derivatives of Equations (20a) and (20b) with respect to $\widehat{K}_t$ and σ_z, and evaluate the resulting systems of equations at the deterministic steady state to obtain closed form solutions for the coefficients in Equations (24c) and (24d).

For example, Equations (20a) - (21) evaluated at the deterministic steady state contain enough information to determine $\widehat{Z}_{ss}$, $\widehat{K}_{ss}$, $\widehat{V}_{ss}^{(0,0)}$ and $\widehat{C}_{ss}^{(0,0)}$. Taking first derivatives (with respect to $\widehat{K}_t$ and σ_z) of Equations (20a) and (20b) and again evaluating at the deterministic steady state allows us to solve for $\widehat{V}_{ss}^{(1,0)}$, $\widehat{V}_{ss}^{(0,1)}$, $\widehat{C}_{ss}^{(1,0)}$ and $\widehat{C}_{ss}^{(0,1)}$. Continuing in this fashion leads to the approximations in Equations (24a) and (24b), where the order of approximation of is equivalent to the number of times we have differentiated Equations (20a) and (20b). As mentioned in Aruoba et al. (2006), the first order solution involves a quadratic matrix equation, but each order of approximation thereafter only necessitates the solution of a linear system. Hence, higher order solutions only require a matrix inversion, albeit of rapidly increasing size.

With approximations $\widetilde{V}_{pert}(\widehat{K}, \sigma_z)$ and $\widetilde{C}_{pert}(\widehat{K}, \sigma_z)$ in hand, we can compute any other variable in the economy, where the accuracy of the approximation of those variables will depend on the underlying accuracy of our approximations for V_t and C_t. However, we can also approximate other variables of interest by augmenting system (20) with additional equilibrium conditions. In our case, we are interested in approximating both the risk-free rate and $\log(V_t/C_t)$ (which we will use in computing welfare costs). The requisite equilibrium conditions are

$$R_{t+1}^f - \mathbb{E}_t \left[M_{t+1}\right]^{-1} = 0 \tag{25a}$$

$$LVC_t - \log\left(\widehat{V}_t/\widehat{C}_t\right) = 0. \tag{25b}$$

Adding Equations (25a) and (25b) to system (20) allows us to obtain approximations

$$\widetilde{R^f}_{pert}(\widehat{K}, \sigma_z) = \sum_{i,j} \widehat{R^f}_{ss}^{(i,j)} (\widehat{K} - \widehat{K}_{ss})^i \sigma_z^j \tag{26a}$$

$$\widetilde{LVC}_{pert}(\widehat{K}, \sigma_z) = \sum_{i,j} \widehat{LVC}_{ss}^{(i,j)} (\widehat{K} - \widehat{K}_{ss})^i \sigma_z^j \tag{26b}$$

as outlined above.

2.2 Projection

Similar to perturbation, projection methods seek a polynomial approximation to a function, F, as in Equation (22), or more commonly to the special case, f, as in Equation (23). However, rather than using the known solution at $\varepsilon = 0$ to construct a local approximation, we specify a polynomial expansion, $\hat{\boldsymbol{x}}$, with coefficients chosen to minimize $f(\hat{\boldsymbol{x}})$ globally, over the domain of $\boldsymbol{x}$. As before, for the neoclassical growth model, f would be comprised of the inter-temporal Euler equation and constraints, and a projection solution would specify a polynomial expansion of the consumption policy that would minimize the Euler equation error.

The analogous approach to our problem would be to use Equations (20a) and (20b) to obtain approximations,

$$\widetilde{V}_{proj}(\widehat{K}) = \sum_{j=0}^{M} a_j \varphi_j(\widehat{K}) \tag{27a}$$

and

$$\widetilde{C}_{proj}(\widehat{K}) = \sum_{j=0}^{M} b_j \varphi_j(\widehat{K}), \tag{27b}$$

where M is the order of approximation and φ_j, $j = 1, 2, \ldots$, represent a set of linearly independent polynomial basis functions. That is, given an order of approximation M, we could specify a grid of $N \geq M$ points for $\widehat{K}$ and evaluate Equations (20a) and (20b) (coupled with the constraint (20c)) at those points to obtain a system of $2N$ equations in $2M$ unknowns. We could then use a nonlinear solution method to find the coefficients $\boldsymbol{a}$ and $\boldsymbol{b}$, in Equations (27a) and (27b), that best satisfy (20a) and (20b).

As there is no theorem to guarantee convergence of the preceding approach, we follow an alternative methodology, suggested by Campanale et al. (2010), Croce (2006) and Kaltenbrunner and Lochstoer (2010), which is to couple polynomial approximations of the value and policy functions with value function iteration. Specifically, we seek a polynomial approximation to the value function as in Equation (27a). Letting $N_k \geq M$, we specify a (not necessarily equally spaced) grid for $\widehat{K}$, spanning the values $(0.1\widehat{K}_{ss}, 1.9\widehat{K}_{ss})$. Additionally, we set $N_\varepsilon = \lceil \frac{M+1}{2} \rceil$ and confine ε to the order N_ε Gauss-Hermite abscissae. To ease notation,

we suppress time subscripts, collect the N_k values of $\widehat{K}$ in the vector $\widehat{\boldsymbol{K}}$ and group the basis functions evaluated at each value of $\widehat{K}$ in the matrix $\Phi(\widehat{\boldsymbol{K}})$, where $\Phi(\widehat{\boldsymbol{K}})_{ij} = \varphi_j(\widehat{K}_i)$. Using this notation, $\widetilde{V}_{proj}(\widehat{K}) = \Phi(\widehat{K})\boldsymbol{a}$. Our algorithm proceeds in the following manner:

Set $\tau = 0.00000001(1-\beta)$, $\Delta = 1$, $l = 0$ and $\boldsymbol{a}^0 = \mathbf{1}$.

while $\Delta > \tau$ **do**

  for $i = 1$ to N_k **do**

    Solve $\widehat{V}_i^* = \max_{\widehat{C}} \left\{ (1-\beta)\widehat{C}^{\frac{1-\gamma}{\theta}} + \beta \left(\mathcal{E}(\widehat{K}_i, \widehat{C}) \right)^{\frac{1}{\theta}} \right\}^{\frac{\theta}{1-\gamma}}$ where

    for $j = 1$ to N_ε **do**

$$\widehat{Z}'_j = \exp(\mu + \sigma_z \varepsilon'_j)$$

$$\widehat{K}'_{i,j}(\widehat{C}) = \frac{1}{\widehat{Z}'_j} \left((1-\delta)\widehat{K}_i + \phi\left(\frac{\widehat{K}_i^\alpha - \widehat{C}}{\widehat{K}_i} \right) \widehat{K}_i \right)$$

and

$$\Psi(\widehat{K}_i, \widehat{C}, \varepsilon'_j) = \Phi(\widehat{K}'_{i,j}(\widehat{C}))\boldsymbol{a}^l.$$

    end for

$$\mathcal{E}(\widehat{K}_i, \widehat{C}) = \sum_{j=1}^{N_\varepsilon} \omega_j \widehat{Z}'^{1-\gamma}_j \Psi(\widehat{K}_i, \widehat{C}, \varepsilon'_j)^{1-\gamma},$$

and where ω_j, $j = 1, 2, \ldots, N_\varepsilon$, are the Gauss-Hermite quadrature weights. Denote the argmax by $\widehat{C}_i^*$. Clearly, $\Psi(\widehat{K}_i, \widehat{C}, \varepsilon'_j)$ is an approximation of $\widehat{V}(\widehat{K}')$, given $\widehat{K}_i$, $\widehat{C}$ and ε'_j, and $\mathcal{E}(\widehat{K}_i, \widehat{C})$ is an approximation of $\mathbb{E}_t \left[\widehat{Z}'^{1-\gamma} \widehat{V}(\widehat{K}')^{1-\gamma} \right]$, given $\widehat{K}_i$ and $\widehat{C}$.

  end for

  Update the coefficients by solving the linear system

$$\boldsymbol{a}^{l+1} = (\Phi(\widehat{\boldsymbol{K}})^T \Phi(\widehat{\boldsymbol{K}}))^{-1} \Phi(\widehat{\boldsymbol{K}})^T \widehat{\boldsymbol{V}}^*,$$

where $\widehat{\boldsymbol{V}}^*$ is the vector comprised of $\widehat{V}_i^*$, $i = 1, 2, \ldots, N_k$.

  Set $\Delta = \max \left\{ \Phi(\widehat{\boldsymbol{K}})\boldsymbol{a}^{l+1} - \Phi(\widehat{\boldsymbol{K}})\boldsymbol{a}^l \right\}$ and $l = l + 1$.

end while

13: Solve for the coefficients of the consumption policy approximant

$$\boldsymbol{b}^{l+1} = (\Phi(\widehat{\boldsymbol{K}})^T \Phi(\widehat{\boldsymbol{K}}))^{-1} \Phi(\widehat{\boldsymbol{K}})^T \widehat{\boldsymbol{C}}^*,$$

where $\widehat{\boldsymbol{C}}^*$ is the vector comprised of $\widehat{C}_i^*$, $i = 1, 2, \ldots, N_k$.

The maximization step in line 2.2 of the algorithm can be performed in a variety of ways; we use a binary search method that exploits the monotonicity of the value function with respect to $\widehat{C}$ (see Appendix C). We also speed the algorithm with a Howard improvement step, performing the maximization in line 2.2 only when $l - 100\lfloor l/100 \rfloor = 0$ (that is, when l modulo 100 is zero) and otherwise computing $\widehat{V}_i^*$ by substituting the contemporaneous value of $\widehat{C}_i^*$. The resulting polynomial approximations are

$$\widetilde{V}_{proj}(\widehat{K}) = \Phi(\widehat{K})\boldsymbol{a}^l \tag{28}$$

and

$$\widetilde{C}_{proj}(\widehat{K}) = \Phi(\widehat{K})\boldsymbol{b}^l. \tag{29}$$

For our particular implementation of the projection algorithm, we use Chebyshev basis functions and their collocation points; this method allows us to choose the N_k values of $\widehat{K}$ so that the interpolation errors are uniformly minimized and so that $N_k = M$. We find that a value as low as $M = 6$ (an order 5 polynomial) provides accurate solutions to the problem.

3 Results

We now apply the solution methods outlined in the previous section to the model of Section 1 and state the main result of our paper: while the quantity dynamics of the two methods are essentially equivalent for a variety of parameter values, the same is not true of variables that are tightly linked to the value function, such as asset prices and welfare costs. We discuss the reasons for this result and outline a very simple motivating example that provides intuition for the particular problem we consider. We conclude the section by reporting diagnostics which compare the accuracy of the solution methods.

3.1 Calibration

We fix several parameters of our model and report them in Table 1. These values are widely accepted in the literature; in particular, the depreciation rate and share of capital, δ and α, respectively, are identical to those of Jermann (1998). The quarterly growth rate, μ, implies annual growth of 1.6 percent and the intertemporal elasticity of substitution (IES) parameter, ψ, is set in the middle of the range ((1,2]) advocated by Bansal et al. (2007). In reality, we considered alternative values of μ and ψ but do not report the corresponding solutions and simulations as they do not alter the qualitative nature of our results. Finally, the adjustment cost parameter, ξ, was chosen so that the ratio of volatilities of log consumption growth to log output growth matches empirical estimates (in the vicinity of 0.5), which depend on the time period and frequency of the data (see discussion below).

α	δ	ψ	μ	ξ
0.36	0.025	1.5	0.004	13

Table 1: Quarterly model calibration

To understand our parameterization of the TFP volatility, it is instructive to consider the data moments reported in Table 2. The table contains means and volatilities for GDP, aggregate consumption and the 90 T-bill, both at annual and quarterly frequencies, for several sample periods. We highlight two important features of the data. First, the volatility of log output growth is markedly different between pre-war and post-war samples, the former being roughly 2.5 to 3 times as great as the latter. Second, the mean of the risk-free rate increases and its volatility decreases as the time horizon is curtailed to include fewer years. In the case of the mean, values in later samples are up to twice as large as that of the pre-war sample.

As a result of the variance in sample moments across sub-periods, we observe a wide range of calibrated values for the TFP volatility, σ_z, and the discount factor, β, in the literature. Since σ_z determines output volatility, models that calibrate to quarterly (post-war) data often specify much smaller values of σ_z than models which calibrate to annual data.

	1929-2008 (A)	1950-2008 (Q)	1960-2008 (Q)	1970-2008(Q)
Std (Δc)	0.0108	0.00490	0.00451	0.00429
Std (Δy)	0.0246	0.00980	0.00857	0.00840
Std (Δc)/Std (Δy)	0.439	0.500	0.526	0.510
Mean(r^f)	0.00847	0.0146	0.0173	0.0168
Std (r^f)	0.0119	0.00747	0.00723	0.00801

Table 2: Data moments for different periods and frequencies. 'A' denotes annual frequency and 'Q' denotes quarterly frequency. Quarterly samples begin with the first quarter of the stated year and end with the final quarter of 2008. 'c' and 'y' denote the log of real consumption (nondurables plus services) and GDP, respectively, and are obtained from NIPA Tables 1.1.4 - 1.1.6, with annual values scaled to quarterly for comparison. 'r_f' denotes the net return on the 90 T-bill, obtained from CRSP (monthly frequency for all horizons), converted to real by subtracting the 12 month lagged moving average of CPI return (as a forecast of expected inflation). The risk-free is annualized by a simple scale factor.

Hence, we allow $\sigma_z \in \{0.01, 0.02, 0.03, 0.04\}$. Since Std $(\Delta y) \approx (1-\alpha)\sigma_z$ (the volatility of log capital growth is very small) in our model, our choices of σ_z correspond to Std $(\Delta y) \in \{0.0064, 0.0128, 0.0192, 0.0256\}$, a range that encompasses the moments reported in Table 2.

For the remaining parameters, the discount rate, β, and coefficient of relative risk aversion, γ, we entertain $\beta \in [0.980, 0.998]$ and $\gamma \in \{2, 5, 10\}$. We choose these values because they not only encompass accepted values in the literature, but they allow a broad enough range of parameterizations to investigate their effect on the sensitivity of the value function to σ_z. In general, we are primarily concerned with $\beta > 0.99$, as these higher values are requisite for matching moments of the risk-free asset.

3.2 Model Implications

3.2.1 Low Volatility

Table 3 reports simulation results for both projection and perturbation methods when $\sigma_z = \{0.01, 0.02\}$ and when $\gamma = 5$. We set $\beta = 0.998$, which simultaneously yields a risk-free

rate in the neighborhood of those observed in post-war data and that allows us to closely approximate the historical pre-war risk-free rate of 0.00847 (see Table 4) when $\sigma_z > 0.02$. We use fifth order Cheybshev polynomials for projection and third order Taylor expansions in the case of perturbation – for the former method, higher order approximations make little material difference to the stated results, and for the latter, numerical instabilities lead to potentially greater discrepancies than those reported. Finally, moments are computed by simulating 100,000 quarterly observations and then aggregating financial variables to an annual frequency by a simple scale factor.

	$\sigma_z = 0.01$			$\sigma_z = 0.02$		
	Proj	Pert	NPert	Proj	Pert	NPert
$Std(\Delta c)$	0.00353	0.00352	0.00352	0.00704	0.00702	0.00702
$Std(\Delta y)$	0.00643	0.00643	0.00643	0.0129	0.0129	0.0129
$Std(\Delta c)/Std(\Delta y)$	0.549	0.549	0.549	0.548	0.546	0.546
$Std(\Delta i)/Std(\Delta y)$	1.85	1.85	1.85	1.84	1.84	1.84
$\mathbb{E}[R_f]$	0.0182	0.0181	0.0190	0.0163	0.0161	NaN
$\text{Std}(R_f)$	0.00116	0.00115	0.00114	0.00232	0.00229	NaN
$\mathbb{E}[R^E - R_f]$	0.0000821	0.000213	-0.000658	0.000653	0.000845	NaN
$\text{Std}(R^E - R_f)$	0.00221	0.00221	0.00221	0.00440	0.00440	NaN
$\text{SR}(R^E)$	0.0371	0.0964	-0.297	0.148	0.192	NaN
$\mathbb{E}(\log V/C)$	3.01	3.00	2.94	2.31	2.12	NaN

Table 3: Simulation moments for both projection (5th order) and perturbation (3rd order) methods, for $\sigma_z = \{0.01, 0.02\}$, $\gamma = 5$ and $\beta = 0.998$. Simulations are quarterly and financial moments are annualized.

We begin by considering the projection results in Table 3. As previously mentioned, the model was calibrated to closely approximate the volatilities of log consumption and output growth in annual pre-war data; $\sigma_z = 0.01$ was chosen to yield an output volatility slightly lower than observed in quarterly data and ξ allows us to fix the ratio $Std(\Delta c)/Std(\Delta y)$. Hence, it is not surprising that the standard deviations of consumption and output are not

drastically different than their counterparts in the data. The remaining moments are freely determined, and in some cases are quite different from observed values. In particular, the equity premium and its volatility are extremely low and the volatility of the risk-free is about six to seven times smaller than what would be expected in the data. In fact, Kaltenbrunner and Lochstoer (2010) find that while holding the Sharpe ratio of equity fixed, there is a trade-off in matching the mean and variance of the equity asset and the volatility of the risk-free. We emphasize that a simple modification to the model à la Croce (2006) (explicitly parameterizing a time varying growth rate in the TFP process) can rectify some of these issues. However, in order to highlight our computational results, we favor parsimony and forsake the additional state variable.

The remaining columns of Table 3 report simulation results for perturbation, for both the case where R_t^f and $\log(V_t/C_t)$ are computed with a direct local approximation (the column denoted 'Pert') and where they are computed nonlinearly with the local solutions of the value function and consumption policy (the column denoted 'NPert'). Regardless of the solution method and the value of σ_z, we see that a third order perturbation yields quantity dynamics that are almost identical to those of projection. The same is not true of asset pricing moments and $\log(V_t/C_t)$. When $\sigma_z = 0.01$, both variants of the perturbation method generate simulated moments that are in close agreement with projection, the one exception being the equity premium, which is extremely close to zero in all cases. However, increasing the TFP volatility to $\sigma_z = 0.02$ renders the nonlinear perturbation unable to compute asset prices and $\log(V_t/C_t)$. The reason is that in the presence of higher volatility, instability of the value function solution results in negative values under a radical or log function, precluding our ability to compute the corresponding moments. These values are reported as 'NaN'. Alternatively, with the direct perturbation, obtained by augmenting the perturbation conditions with Equations (26a) and (26b), we are able to drastically improve the simulated moments of the local method; in this case, column 'Pert' shows that the local method only exhibits slight deviations from the global method for asset prices, again with the exception of the equity premium which is very close to zero in both cases. The deviation for $\log(V_t/C_t)$ is slightly larger, but not horrendous. As we mention in Section 2, there is no theoretical reason to resort to direct approximations for ancillary model variables; in fact,

if the solutions for V_t and C_t were good enough, any nonlinear function of them would also yield a highly accurate approximation. However, as we will see below, perturbation delivers a poor solution of V_t when σ_z is high. The result is that a direct local approximation of R_t^f and $\log(V_t/C_t)$ reduces the dependency of these variables on the value function and improves their accuracy.

3.2.2 High Volatility

Asset pricing papers that match moments of annual, pre-war data sets are too numerous to cite. Table 4 reports simulation results for both projection and perturbation methods when $\sigma_z = \{0.03, 0.04\}$, the latter being a value that conforms to pre-war, annual parameterizations. As before, $\gamma = 5$ and $\beta = 0.998$.

	$\sigma_z = 0.03$			$\sigma_z = 0.04$		
	Proj	Pert	Npert	Proj	Pert	NPert
$Std(\Delta c)$	0.0105	0.0105	0.0105	0.0140	0.0138	0.0138
$Std(\Delta y)$	0.0193	0.0193	0.0193	0.0257	0.0257	0.0257
$Std(\Delta c)/Std(\Delta y)$	0.547	0.543	0.543	0.543	0.537	0.537
$Std(\Delta i)/Std(\Delta y)$	1.82	1.83	1.83	1.80	1.81	1.81
$\mathbb{E}[R_f]$	0.0130	0.0127	NaN	0.00847	0.00779	NaN
$\text{Std}(R_f)$	0.00345	0.00343	NaN	0.00455	0.00468	NaN
$\mathbb{E}[R^E - R_f]$	0.00166	0.00195	NaN	0.00299	0.00370	NaN
$\text{Std}(R^E - R_f)$	0.00651	0.00654	NaN	0.00855	0.00860	NaN
$\text{SR}(R^E)$	0.254	0.299	NaN	0.350	0.430	NaN
$\mathbb{E}(\log V/C)$	1.44	0.663	NaN	0.561	-1.38	NaN

Table 4: Simulation moments for both projection (5th order) and perturbation (3rd order) methods, for $\sigma_z = \{0.03, 0.04\}$, $\gamma = 5$ and $\beta = 0.998$. Simulations are quarterly and financial moments are annualized.

The previous discrepancies now become exaggerated: while the global method and both variations of the local method show high agreement for quantity dynamics, solutions for asset

moments and $\log(V_t/C_t)$ diverge as σ_z increases. As before, nonlinear perturbation is unable to compute asset prices and $\log(V_t/C_t)$ for high σ_z, but the direct local approximations ameliorate the problem. However, in the most extreme case of $\sigma_z = 0.04$, even the direct perturbation and projection show moderate discrepancies for asset prices, and for both values of σ_z the means of $\log(V_t/C_t)$ are quite different.

The inability of the local method to approximate $\log(V_t/C_t)$ is crucial for welfare analysis. For example, to compute the welfare costs of TFP volatility, one would simply evaluate the difference $\log(V_t^l/C_t^l) - \log(V_t^h/C_t^h)$, where V_t^l and C_t^l are computed under low volatility and V_t^h and C_t^h are computed under high volatility. The resulting value is interpreted as the percentage change in the agent's utility (as a fraction of consumption) as volatility changes. These differences are easily computed from the values reported in Tables 3 and 4: according to Chebyshev projection, a one percent increase in TFP volatility from $\sigma_z = 0.03$ to $\sigma_z = 0.04$ results in a welfare loss of 1.44 - 0.561 = 0.879, while the analogous computation due to perturbation is 0.663 + 1.38 = 2.04 – more than twice the value of the global method.

We will see that the findings in this section are a result of the fact that perturbation is a local approximation around the deterministic steady state ($\sigma_z = 0$), and that the value function exhibits a high degree of curvature in the direction of σ_z. For this reason, perturbation has difficulty achieving an accurate approximation as the calibrated value of σ_z moves away from zero. We examine this result in more detail below, but first consider a very simple example that provides intuition for the source of the discrepancy.

3.3 Motivating Example

We now consider the very simple example of approximating the square root function, $f(x) = \sqrt{x}$, with a Taylor polynomial of various orders[1]. Judd (1998) provides similar numerical results for $x^{1/4}$. Generally speaking, any (analytic) continuously differentiable function can be written as,

$$f(x) = \sum_{i=0}^{\infty} \frac{f^{(i)}(x_0)}{i!}(x - x_0)^i, \tag{30}$$

[1]We thank George Tauchen for suggesting this basic example to illustrate the intuition of our main result.

for x in a neighborhood of x_0, where $f^{(i)}$ denotes the ith derivative of f and where $f^{(0)} = f$. For the square root function, it is not difficult to show that,

$$f^{(i)}(x) = (-1)^{i+1}\frac{(2i-3)!!}{2^i}x^{\frac{1}{2}-i}, \tag{31}$$

where $n!!$ denotes the double factorial of n [2]. Thus, the square root function can be written as

$$\begin{aligned}\sqrt{x} &= \sqrt{x_0} + \sum_{i=1}^{\infty}(-1)^{i+1}\frac{(2i-3)!!}{2^i i!}x_0^{\frac{1}{2}-i}(x-x_0)^i \\ &= \sqrt{x_0} + \sum_{i=1}^{\infty} a_i(x-x_0)^i,\end{aligned} \tag{32}$$

where $a_i = \frac{f^{(i)}(x_o)}{i!}$. For a general power series of the form in Equation (32), the radius of convergence is defined as the value $r \in \overline{\mathbb{R}}^+$ such that (32) converges for $|x - x_0| \leq r$; that is, the radius of convergence identifies the neighborhood for which the function converges. A simple way to determine the radius of convergence is

$$r = \lim_{i\to\infty}\left|\frac{a_i}{a_{i+1}}\right|. \tag{33}$$

Hence, for the square root function (assuming $x_0 > 0$),

$$\begin{aligned}r_{sq}(x_0) &= \lim_{i\to\infty}\left|\frac{(-1)^{i+1}\frac{(2i-3)!!}{2^i i!}x_0^{\frac{1}{2}-i}}{(-1)^{i+2}\frac{(2i-1)!!}{2^{i+1}(i+1)!}x_0^{\frac{1}{2}-i-1}}\right| \\ &= \lim_{i\to\infty}\frac{\frac{(2i-3)!!}{2^i i!}x_0^{\frac{1}{2}-i}}{\frac{(2i-1)!!}{2^{i+1}(i+1)!}x_0^{\frac{1}{2}-i-1}} \\ &= \lim_{i\to\infty}\frac{2(i+1)}{2i-1}x_0 \\ &= x_0.\end{aligned} \tag{34}$$

From Equation (34) we understand that the Taylor series expansion of the square root function around the point x_0 is only guaranteed to converge for $x \in (0, 2x_0)$. Outside of this range, the series expansion will diverge.

[2] For n odd, $n!!$ is the product of all odd numbers less than or equal to n. Similarly for n even. The double factorial of 0 and -1 are defined to be 1.

Intuitively, the radius of convergence for a Taylor series depends on the rate at which the derivatives of the target function diminish at the point of approximation. For a function with little shape, the high order derivatives drop quickly to zero, forcing r to be quite high – the extreme case being a polynomial of finite order, with an infinite radius of convergence (the Taylor series converges for all $x \in \mathbb{R}$). Conversely, in cases where the high order derivatives do not exhibit quick decay, the radius of convergence is low, resulting in a Taylor series expansion which is only applicable over a small portion of the domain[3].

To illustrate this concept, we approximate the square root function at two values: $x_0 = 1$ and $x_0 = 2$. In the first case the radius of convergence is 1 and we anticipate that a Taylor series approximation will only be appropriate in the range (0,2). Figure 1 depicts the first nine Taylor polynomial approximations of $f(x) = \sqrt{x}$ around $x_0 = 1$. Clearly, the polynomial approximations are adequate for $x \in (0, 2)$, but diverge outside of that range; while increasing the order of approximation to arbitrary levels allows us to fit the function at any desired level of precision over the interval (0,2), the approximations become erratic outside of that interval for high orders. In the second case, the radius of convergence is 2, indicating that the Taylor series will converge on the interval (0,4). Figure 2 depicts the first nine Taylor polynomial approximations around $x_0 = 2$, confirming our prior intuition.

This example illustrates that there is an inverse relationship between the degree of curvature of a function at a point of interest and the size of the interval (around that point) over which a Taylor polynomial approximation is adequate. It is precisely this concept which drives the main result of our paper, as we will see below.

3.4 Graphical Evidence

We re-consider the model solutions of Section 3.2. Figure 3 shows policy function approximations for 5th order Chebyshev projection and perturbation of orders 1,2 and 3, all for the case of $\sigma_z = 0.04$ and $\beta = 0.998$ (for C_t and V_t, the 'Pert' and 'NPert' solutions are identical). Figure 4 depicts similar approximations for the value function.

These plots clarify the results reported in Table 4: both projection and perturbation

[3]den Haan and de Wind (2009) provides a more general and thorough discussion of the relationship between polynomial approximations and the radius of convergence.

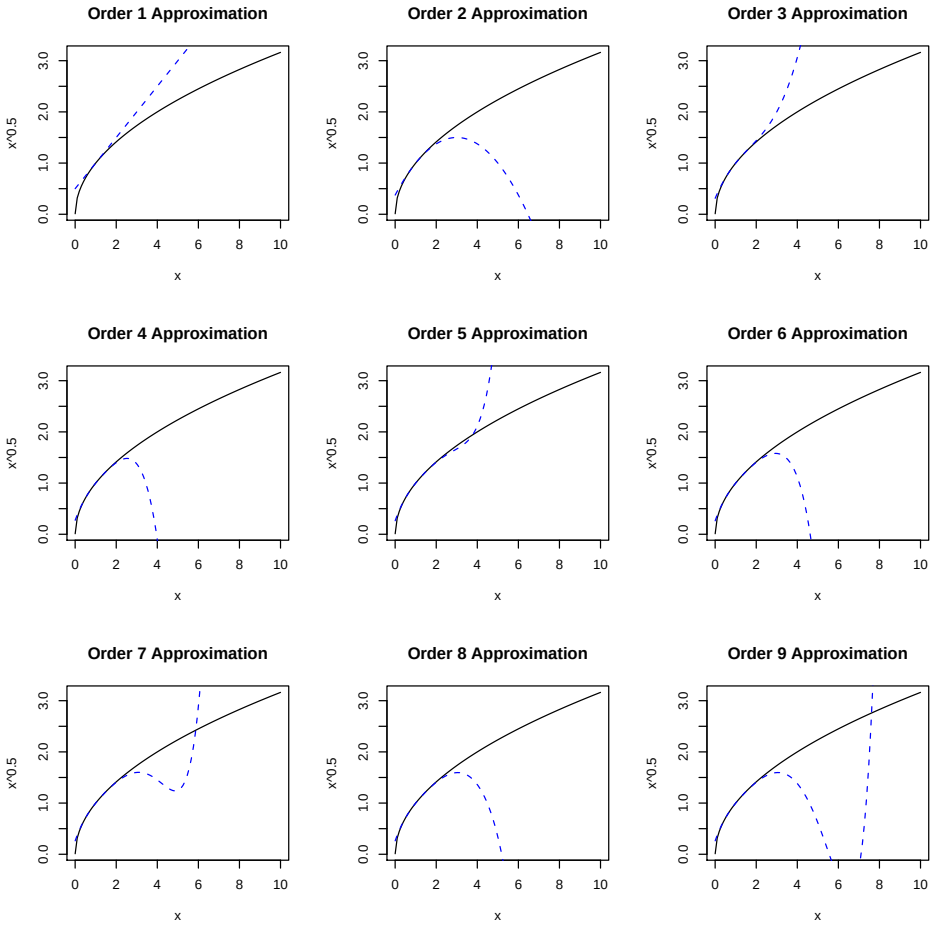

Figure 1: First nine Taylor polynomial approximations of $\sqrt{x}$ around $x_0 = 1$.

produce policy functions that are in close agreement, while there is a wide discrepancy in their solutions for the value function, even among among the different order perturbations. Since the quantity dynamics of the solutions are not sensitive to the value function, it is not surprising that the simulated macroeconomic moments of the two methods do not differ by a great amount. However, the risk-free rate and $\log(V_t/C_t)$ both depend directly on the level and shape of the value function, and hence are quite different across methods. Figures 5 and 6 depict the same approximations for the case of $\sigma_z = 0.01$, and demonstrate that when the TFP volatility is low, the solution methods are far more similar, as expected from the simulation output in Table 3. In this case, the consumption policy approximations overlap to an even greater extent and the value function approximations are separated by only a

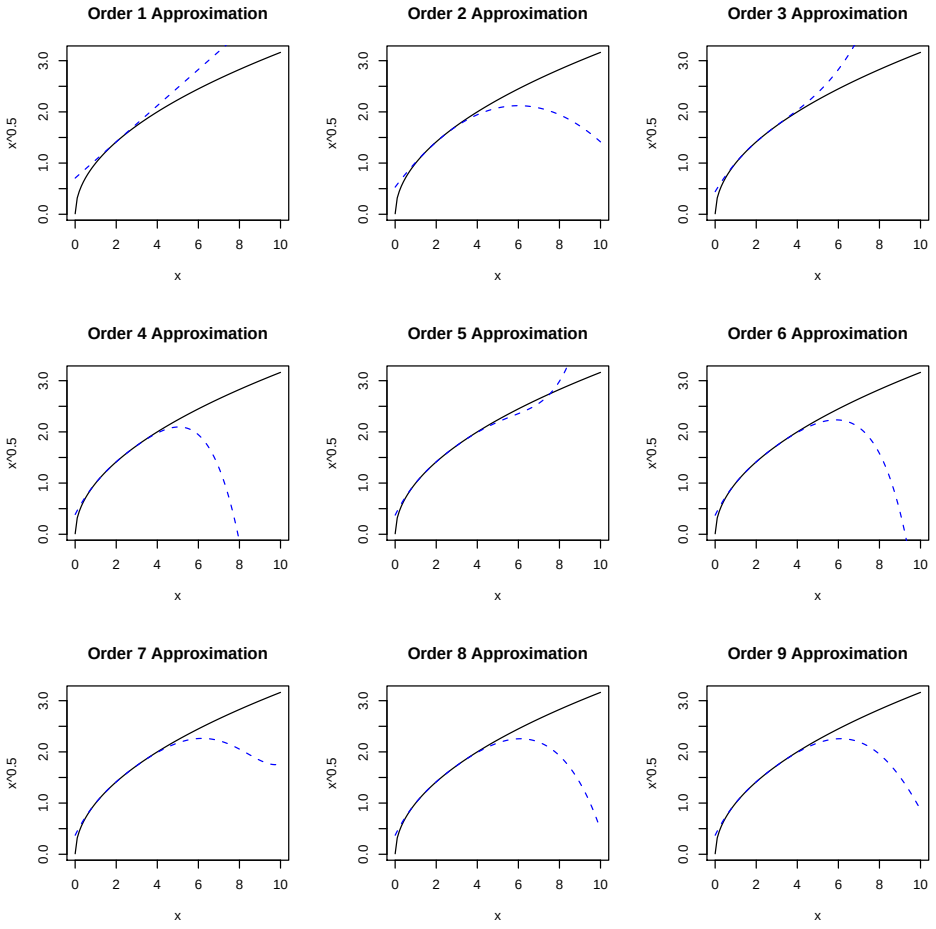

Figure 2: First nine Taylor polynomial approximations of $\sqrt{x}$ around $x_0 = 2$.

(relatively) small level shift. This latter shift diminishes as we shrink σ_z toward zero.

We note that in our model it is possible to analytically solve for the return on equity in terms of aggregate variables (see Appendix B), and hence it is unaffected by poor local approximations of the value function. It follows that local approximations of the equity premium are only affected by the value function via the risk-free rate. This result may not extrapolate to more general models where analytical expressions of the return on equity are not available.

To understand why the value functions for the two solution methods diverge for large σ_z, it is useful to think of the value function and consumption policy as functions of both the state variable, $\widehat{K}$, *and* the TFP volatility parameter, σ_z. Figures 7 and 8 show policy and value

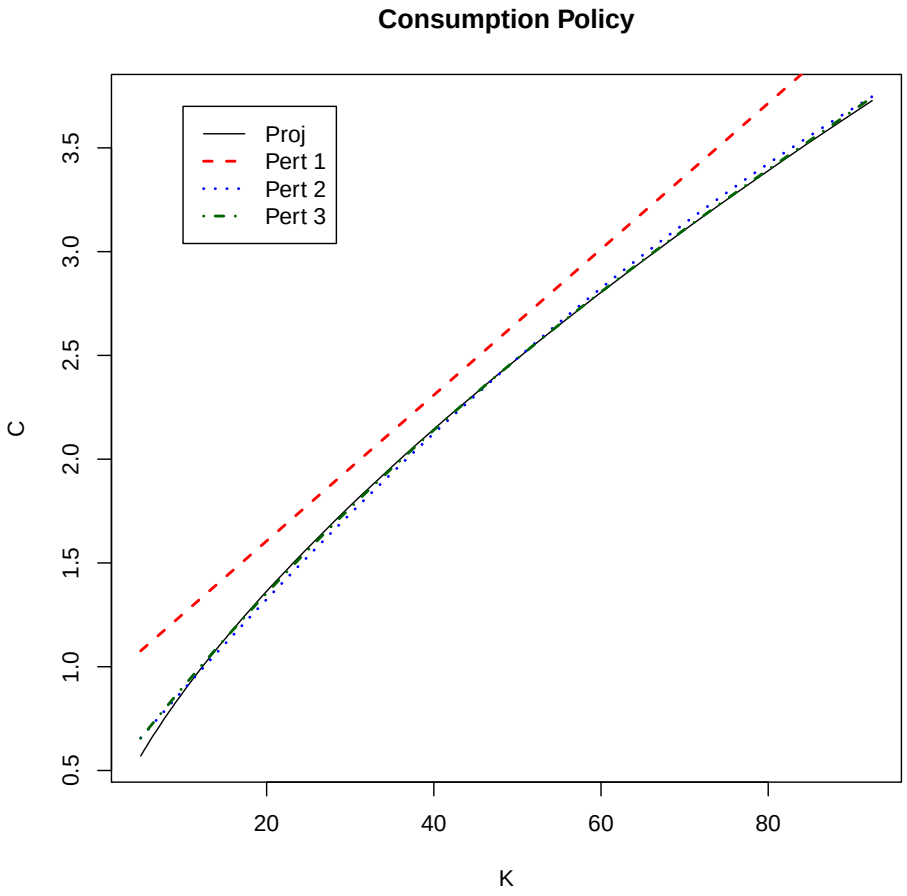

Figure 3: Consumption policy approximations for Chebyshev projection (5th order) and perturbation of orders 1,2 and 3, where $\sigma_z = 0.04$ and $\beta = 0.998$.

function approximations for 5th order Chebyshev projection and 3rd order perturbation, when $\sigma_z \in [0, 0.04]$. Thus, the approximations in Figures 3 – 6 are simply cross sections (fixing σ_z) of the functions depicted in Figures 7 and 8. It becomes apparent from inspecting these surfaces that the value function exhibits a high degree of curvature in the direction of σ_z, with the amount of curvature increasing as σ_z approaches zero, whereas the consumption policy is quite flat. As a result, similar to the square root function we previously considered, we anticipate that the radius of convergence of a Taylor polynomial approximation of the value function will diminish for approximations centered at points very close to $\sigma_z = 0$. This is a exactly what a perturbation solution is: a local Taylor approximation *at* $\sigma_z = 0$. It follows that, for the particular case of the value function, we have no guarantee of

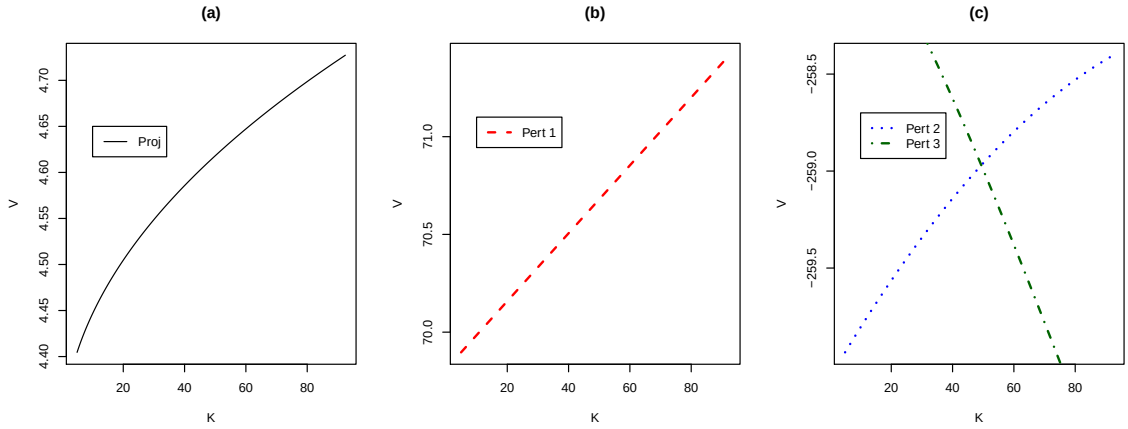

Figure 4: Value function approximations for Chebyshev projection (5th order) and perturbation of orders 1, 2 and 3, where $\sigma_z = 0.04$ and $\beta = 0.998$. Note the widely different scales on the vertical axes.

convergence for values of σ_z far away from zero, and in fact we should not be surprised to see divergent behavior, as suggested by the previous plots. This result is congruent with den Haan and de Wind (2009) who find that nonlinearities in DSGE models can render high order perturbation solutions that are explosive. On the other hand, the radius of convergence for the consumption policy is likely to be quite large, and we expect local Taylor approximations to converge for a wide range of σ_z – this is corroborated by Figure 7, where we are unable to distinguish the two surfaces.

The upshot of the foregoing results is that a global projection method is more robust to value function curvature, since it seeks to minimize an error equation expressed in terms of the true (unknown) value function, rather than approximating the truth at a distant focal point.

3.5 Solution Evaluation and Sensitivity Analysis

In the preceding analysis we have merely shown some conditions under which the two solution methods we consider are different; we have not formally investigated their relative accuracy. We now undertake the important task of determining which of the solutions is a closer approximation to the unknown truth and do so for a variety of model parameter values. While our primary evaluation criterion will be Euler equation errors, we will conclude the

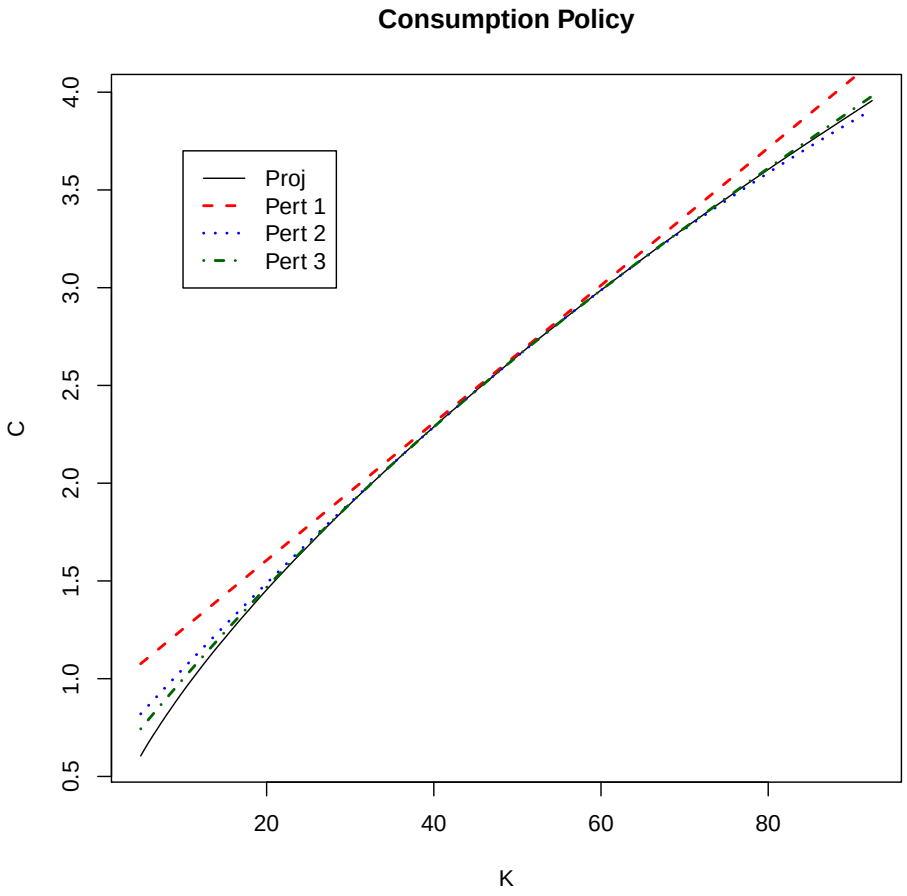

Figure 5: Consumption policy approximations for Chebyshev projection and perturbation of orders 1,2 and 3, where $\sigma_z = 0.01$ and $\beta = 0.998$.

section with a discussion of the den Haan-Marcet statistic (den Haan and Marcet (1994)).

3.5.1 Pricing Errors

The fundamental asset pricing equation is $1 = \mathbb{E}_t[M_{t+1}R_{t+1}]$, where M_{t+1} is the time t stochastic discount factor, and R_{t+1} is the return for any asset between t and $t+1$. Thus, from Equation (20b) we have

$$\begin{aligned} 1 &= \mathbb{E}_t[M_{t+1}R^e_{t+1}] \\ &= \mathbb{E}_t\left[M_{t+1}\,\phi'\left(\frac{\widehat{I}_t}{\widehat{K}_t}\right)\left(\frac{(\alpha-1)\widehat{Y}_{t+1}+\widehat{C}_{t+1}}{\widehat{K}_{t+1}}+\frac{\phi\left(\frac{\widehat{I}_{t+1}}{\widehat{K}_{t+1}}\right)+1-\delta}{\phi'\left(\frac{\widehat{I}_{t+1}}{\widehat{K}_{t+1}}\right)}\right)\right]. \end{aligned} \tag{35}$$

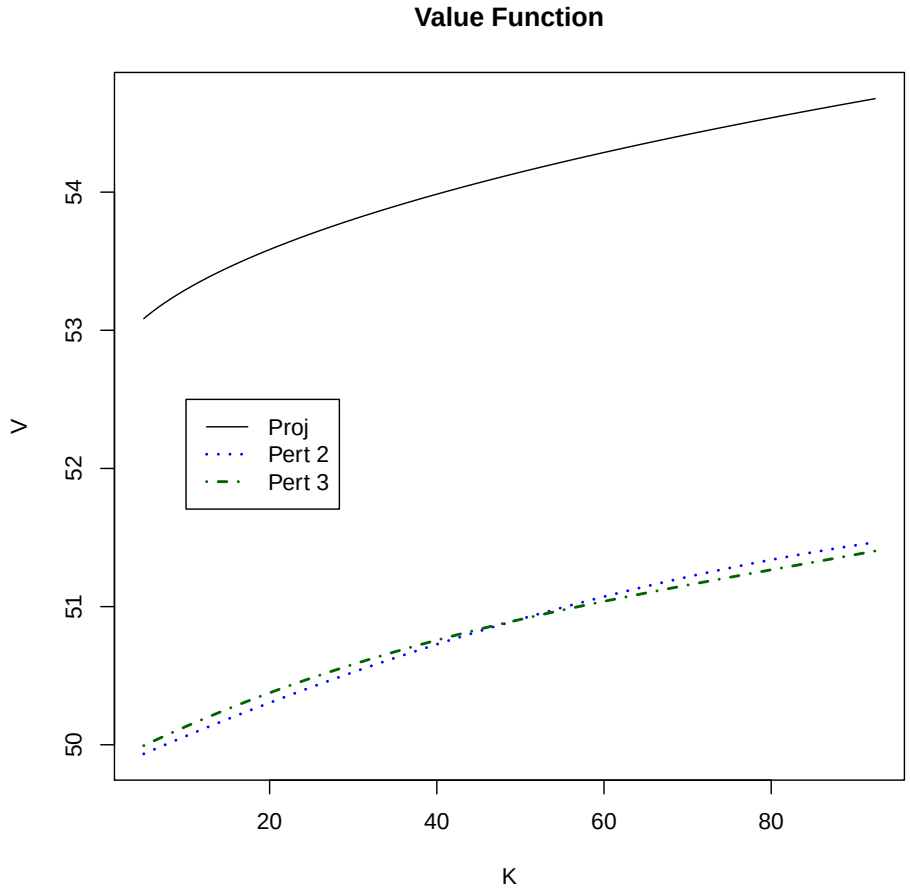

Figure 6: Value function approximations for Chebyshev projection and perturbation of orders 2 and 3, where $\sigma_z = 0.01$ and $\beta = 0.998$.

Since the stochastic discount factor M_{t+1} incorporates current consumption, C_t, in its denominator (see Equation (17)), we can interpret pricing errors as a fraction of contemporaneous consumption. As suggested by Judd and Guu (1997), Aruoba et al. (2006) and Caldara et al. (2009), base 10 logarithms of pricing errors in Equation (35) can be interpreted in the following manner: a value of -1 corresponds to a 10% consumption error, a value of -2 corresponds to a 1% consumption error, a value of -3 corresponds to a 0.1% consumption error, etc.. Combining Equation (35) with the long simulations of ε_t (see Equation (21)) used in Section 3.2, we can compute the mean of the pricing errors implied by the model, for each solution method. The expectation is approximated by a Gauss-Hermite quadrature rule, with the order chosen so as to exactly compute the integral for the finite polynomial so-

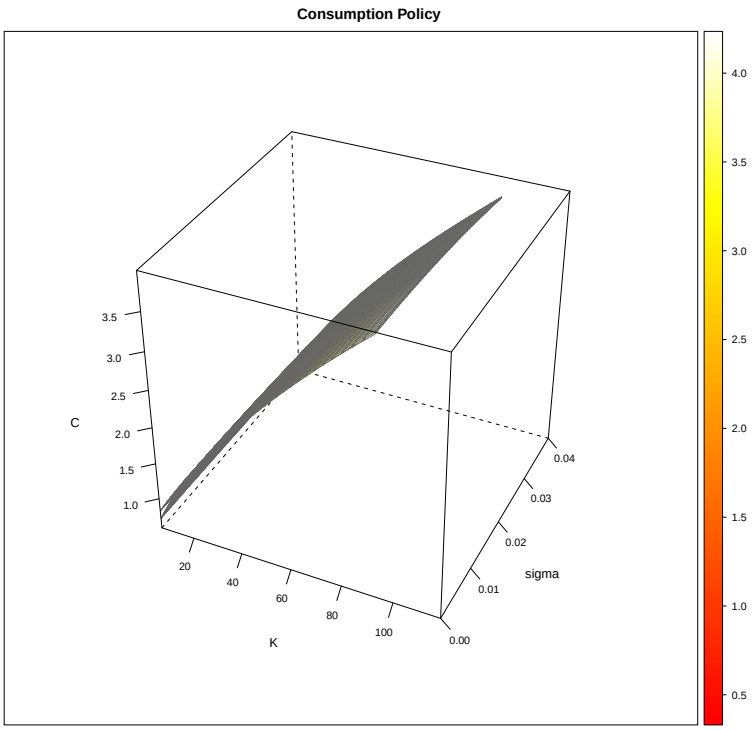

Figure 7: Consumption policy approximations for 5th order Chebyshev projection and 3rd order perturbation, for $\sigma_z \in [0, 0.04]$.

lutions. The pricing errors are reported graphically in the upper rows of Figures 9 - 12. The individual plots depict how mean Euler equation errors vary with β, where $\beta \in [0.980, 0.998]$ – in general, perturbation solution quality degrades as β rises. Moving across the upper rows, from left to right, we are then able to observe the effect of increasing risk aversion, γ, and moving between the four figures we observe the effect of increasing TFP volatility, σ_z – as with β, the quality of the perturbation solution degrades as each of these parameters increases. At the lower extreme, in Figure 9, when $\sigma_z = 0.01$ and $\gamma = 2$, a 3rd order perturbation dominates a 5th order projection for virtually all values of β that we consider. However, both solutions produce errors that most would consider economically insignificant (less than 0.01% of consumption). Holding σ_z fixed and increasing γ, the perturbation errors rise to levels as high as 1% of consumption, for high values of β. These qualitative results become more pronounced in Figures 10 - 12, where at the upper extreme ($\sigma_z = 0.04$ and $\gamma = 10$), perturbation errors exceed 10% of consumption, for high values of β. It is this

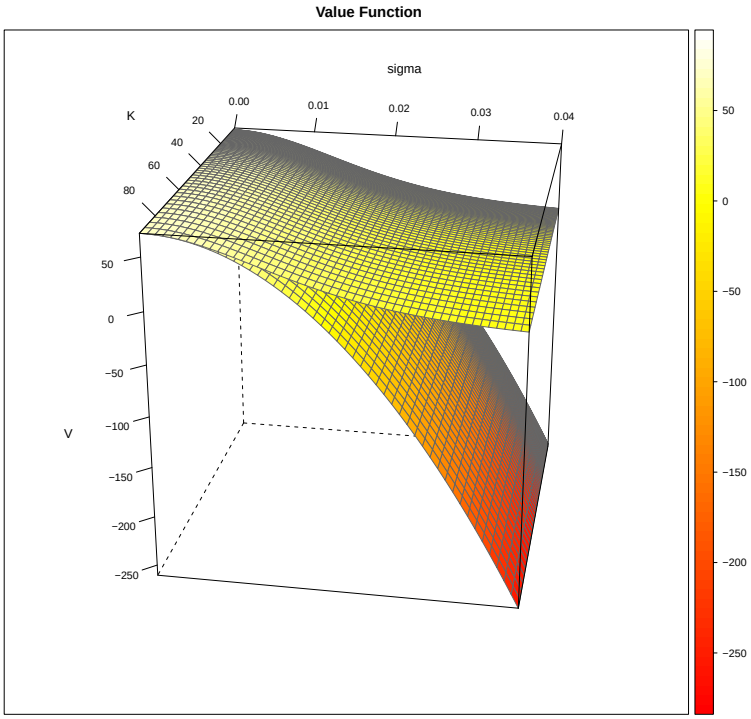

Figure 8: Value function approximations for 5th order Chebyshev projection and 3rd order perturbation, for $\sigma_z \in [0, 0.04]$. The lower surface corresponds to perturbation.

final case that deserves particular attention: models that calibrate to annual, pre-war data generally require high values of σ_z (on the order of 0.04) and β (on the order of 0.998 or above) in order to match output volatility and the level of the risk-free rate. We see that these calibrations, matched with moderate levels of risk aversion (above 5) can lead to poor local approximations. On the other hand, models that calibrate to quarterly, post-war data typically obtain much smaller values of σ_z (on the order of 0.01 or below), and do not suffer from poor local approximations.

The fundamental characteristic driving these results is the curvature of the value function with respect to TFP volatility: as shown in Figure 8, the value function can exhibit a high degree of curvature in the direction of σ_z. In cases where the curvature is extreme, a local method such as perturbation will have difficulty approximating the function at points far from the deterministic steady-state (the locus of approximation), a result which is corroborated by Figures 9 - 12. This is especially relevant for models that require a high calibrated

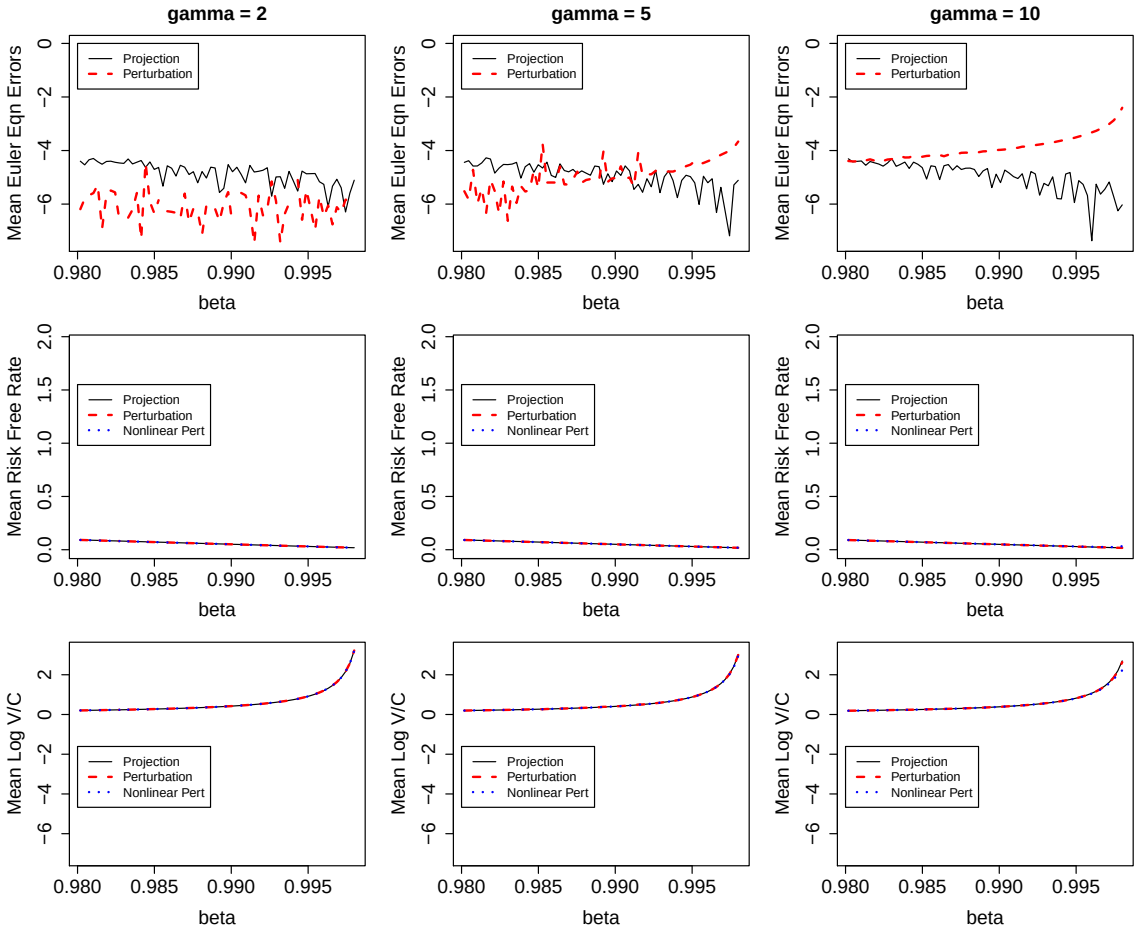

Figure 9: Mean $\log_{10}$ Euler equation errors (first row), mean risk-free rate (second row) and mean log ratio of value function to consumption policy, plotted as functions of β for different values of γ. $\sigma_z = 0.01$ in all cases.

value of σ_z. The parameters γ and β have an effect insofar as they increase the sensitivity of the value function to changes in σ_z; i.e. the sensitivity increases with each of these parameters.

The second and third rows of Figures 9 - 12 depict the mean of R_t^f and the mean of $\log(V_t/C_t)$, respectively, across simulations. As in Section 3.2, we compute these values both nonlinearly, via Equations (17), (25a) and (25b), and directly, via Equations (26a) and (26b). As with the Euler equation errors, the nonlinear perturbation risk-free rate deviates dramatically from that of projection as β, γ and σ_z rise. This discrepancy is most pronounced for $\sigma_z = 0.04$, $\gamma \geq 5$ and $\beta \geq 0.99$. On the other hand, the direct perturbation computation

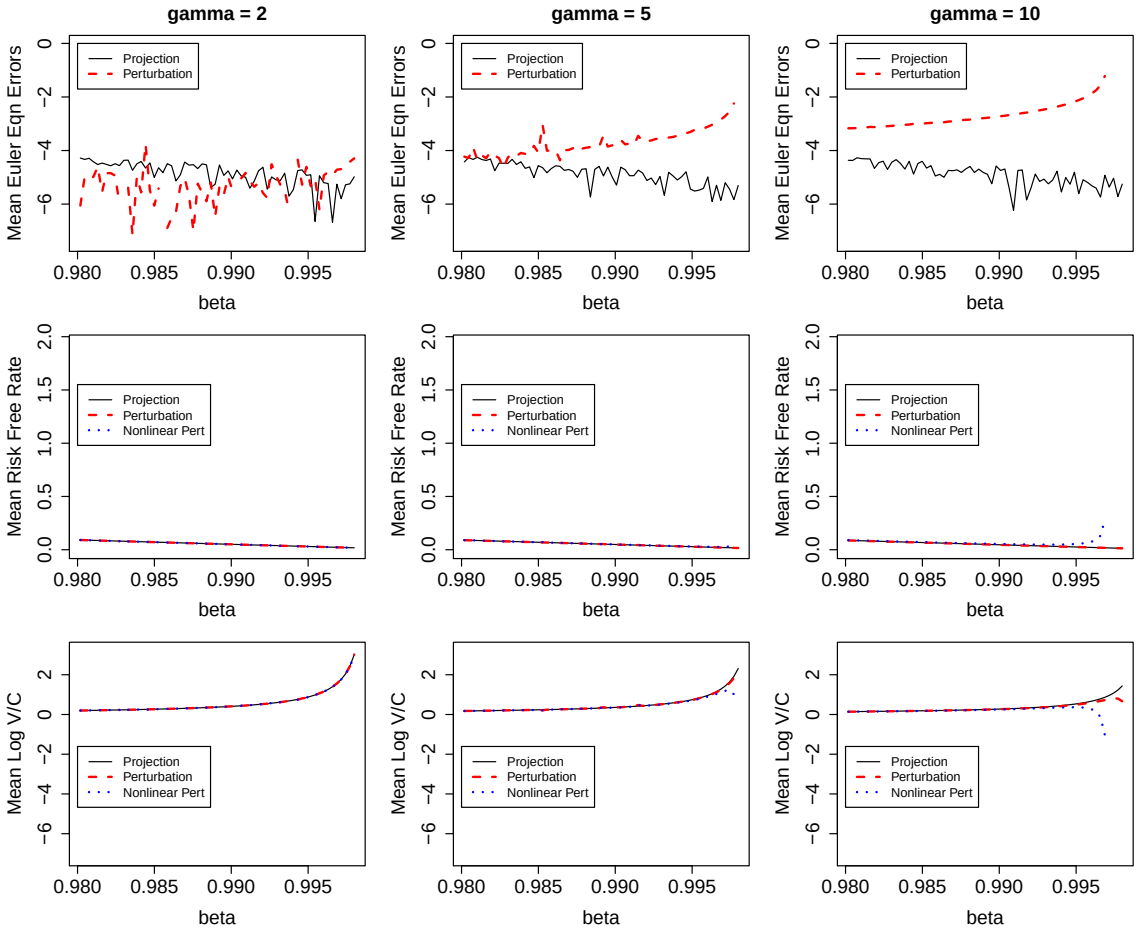

Figure 10: Mean $\log_{10}$ Euler equation errors (first row), mean risk-free rate (second row) and mean log ratio of value function to consumption policy, plotted as functions of β for different values of γ. $\sigma_z = 0.02$ in all cases.

appears to be quite similar to projection for all parameter values. In truth, the local method exhibits a small amount of divergent behavior as well, but the graphical evidence is washed out by the scale of the nonlinear deviation. The moments in Tables 3 and 4 give an idea of the magnitude of divergence.

The reason for the discrepancy in the risk-free rate computations is the same as for the Euler equation errors: for high values of β, γ and σ_z, perturbation provides a poor approximation to the value function. Since the risk-free rate depends directly on the value function in models with recursive utility (see Equations (25a) and (17)), it is likewise poorly approximated by perturbation, insofar as the value function approximation is poor. This effect is

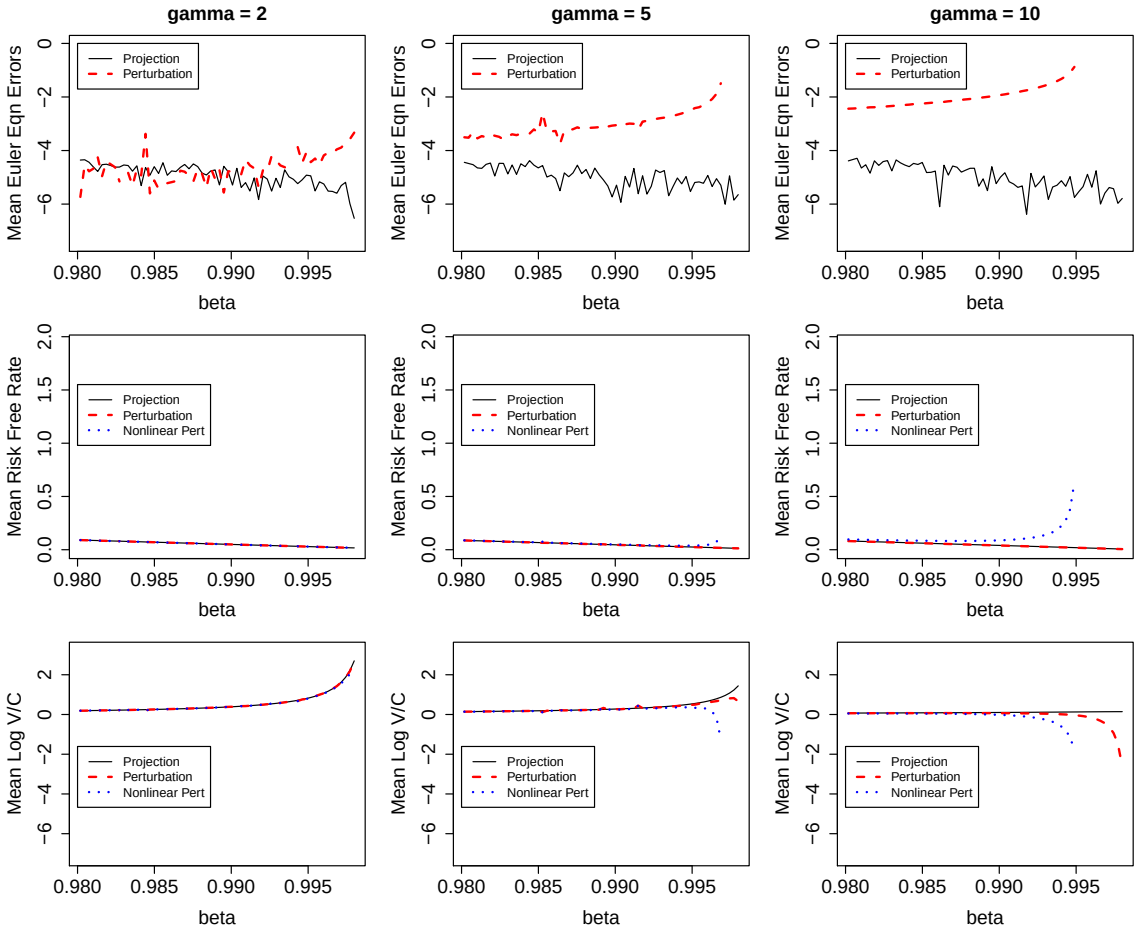

Figure 11: Mean $\log_{10}$ Euler equation errors (first row), mean risk-free rate (second row) and mean log ratio of value function to consumption policy, plotted as functions of β for different values of γ. $\sigma_z = 0.03$ in all cases.

most severe when we compute the risk-free fate nonlinearly. Viewed from the opposite perspective, if our approximations for C_t and V_t were highly accurate, a nonlinear computation of R_t^f would likewise be highly accurate. Thus, the large deviations in Figures 10 - 12 are just further evidence of the poor local approximation of V_t for high β, γ and σ_z. The interesting aspect of our results, though, is that we can ameliorate the effect of the value function by directly computing the risk-free rate via a Taylor expansion. This latter method anchors the risk-free at its deterministic steady state and weakens the computational relationship between R_t^f and V_t.

The final rows of Figures 9 - 12 lend more insight to the foregoing results. As mentioned

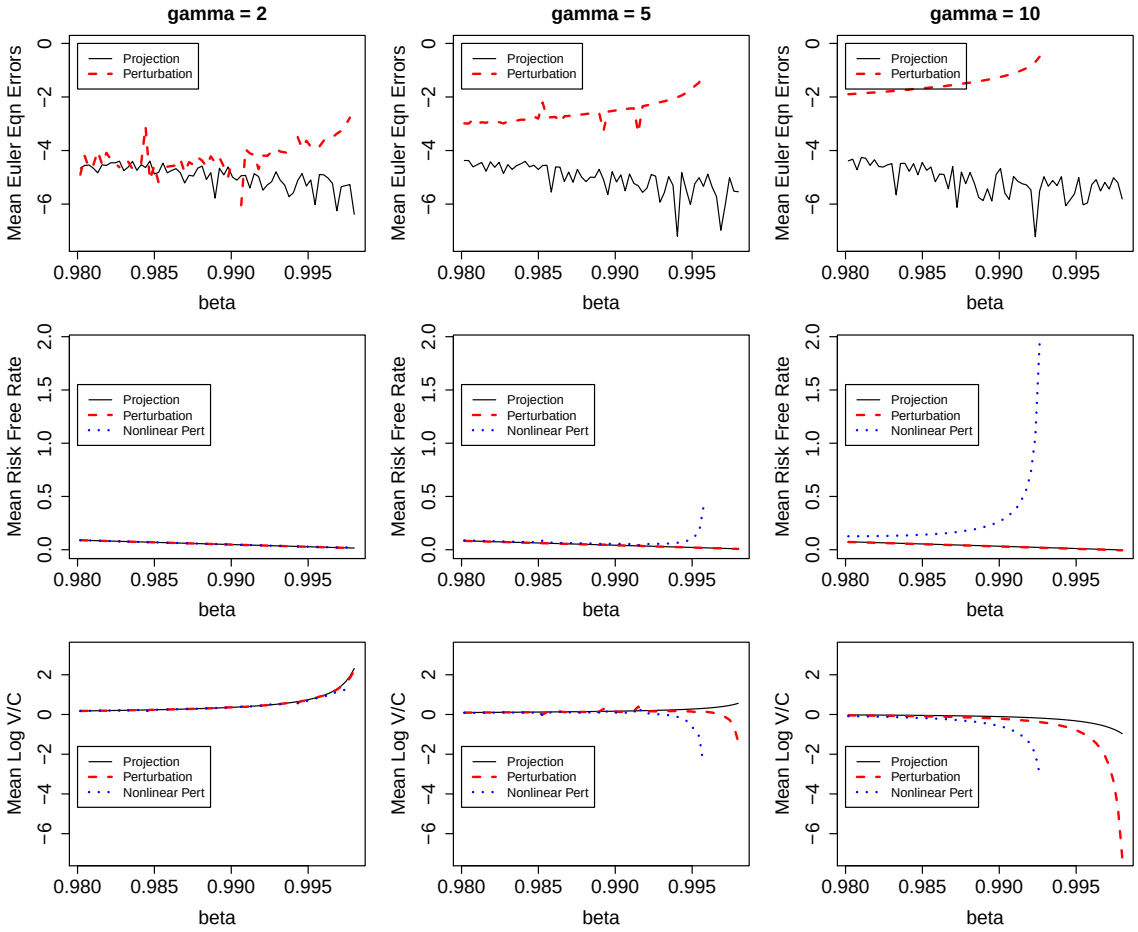

Figure 12: Mean $\log_{10}$ Euler equation errors (first row), mean risk-free rate (second row) and mean log ratio of value function to consumption policy, plotted as functions of β for different values of γ. $\sigma_z = 0.04$ in all cases.

in Section 3.2, we include the log of V_t/C_t in our analysis, since this variable is instrumental in welfare evaluations. Once again, the approximations deteriorate as β, γ and σ_z increase, which we attribute to the poor local approximation of the value function. However, in the case of $\log(V_t/C_t)$ the deviations are more severe (for the direct Taylor expansion method) than for the risk-free. The reason for this is that the risk-free rate depends on the value function in both numerator and denominator (see Equations (17) and (25a)), which mitigates the error propagation of the value function approximation. The same could be true of welfare computations $\log(V_t^*/C_t^*) - \log(V_t^{**}/C_t^{**})$ when V_t^* and V_t^{**} are computed with the same σ_z (i.e., welfare effects are evaluated for a variable other than σ_z). However, since the

approximations of V_t are likely to be very different for different values σ_z (it will be well approximated for low volatility and badly approximated for high volatility), we expect that welfare computations attributed to TFP volatility will be very poorly approximated.

As a final note, we repeated all of the previous analysis for $\mu = 0$ and $\psi = 0.5$; these changes only caused slight shifts, leaving the results above qualitatively the same. For space considerations, we do not include them. We recognize, however, the particular interplay between β and μ: lowering the growth rate increases the model's ability to tolerate high values of β before local approximations become very poor. That is, we can think of the model as depending on the single growth adjusted subjective discount factor $\beta^* = \beta \exp(\mu)$ – lowering μ or β effectively decreases β^*, and hence the model's sensitivity to σ_z.

3.5.2 den Haan-Marcet Statistic

Since agents in our model have rational expectations, the residual of the pricing equation,

$$u_{t+1} = 1 - M_{t+1} R^e_{t+1} \tag{36}$$

should not be in the time t information set. That is, under the null hypothesis that we have correct solutions for the value and policy functions, $\boldsymbol{\beta} = \mathbf{0}$ in regressions of the form

$$u_{t+1} = \sum_{i=1}^{n} \beta_i x_{i,t} + \zeta_{t+1}, \tag{37}$$

for $t = 1, 2, \ldots, T$, where $x_{i,t}$ represent variables in the time t information set. den Haan and Marcet (1994) suggest testing this hypothesis by constructing a Wald-type statistic

$$DM(n) = \boldsymbol{u}' \boldsymbol{X} \left[\sum_{t=1}^{T} \boldsymbol{x}_t \boldsymbol{x}_t' \hat{\zeta}^2_{t+1} \right]^{-1} \boldsymbol{X}' \boldsymbol{u} \tag{38}$$

where $\boldsymbol{x}_t$ is the vector of time t regressors, $x_{1,t}$, $x_{2,t}$, …, $x_{n,t}$, $\boldsymbol{X}$ is the matrix with rows $\boldsymbol{x}_t'$, and $\hat{\zeta}_{t+1} = u_{t+1} - \boldsymbol{x}_t' \hat{\beta}$. If the null hypothesis is true, $DM(n) \overset{a}{\sim} \chi^2(n)$; however, since the probability of attaining the true solution is zero, we expect that large values of T will force a rejection of the test. To account for this, den Haan and Marcet (1994) compute $DM(n)$ for multiple simulations of $\boldsymbol{u}$ and determine the proportion of times that the statistic falls within certain critical limits of the $\chi^2(n)$ distribution. If the approximate solutions are good, the

proportions within these bounds should be close to the actual area under the $\chi^2(n)$ density function.

In our implementation of the Den Haan-Marcet statistic, we regress the price residuals on a constant and five lags of both log consumption growth and log productivity growth (hence, $n = 11$). We fix $\beta = 0.998$ and simulate 500 data sets of $T = 3000$ quarterly observations (750 years of data) and report the proportion of time that the value from Equation (38) is above or below the 5% points of the $\chi^2(11)$ density in Table 5. This Wald-type diagnostic

	$\sigma_z = 0.01$			$\sigma_z = 0.04$		
	$\gamma = 2$	$\gamma = 5$	$\gamma = 10$	$\gamma = 2$	$\gamma = 5$	$\gamma = 10$
Projection	(0.052, 0.054)	(0.052, 0.052)	(0.050, 0.052)	(0.052, 0.056)	(0.054, 0.062)	(0.050, 0.07)
Perturbation	(0.058, 0.052)	(0.050, 0.058)	(0.006, 0.338)	(0,0)	(0,0)	(0,0)

Table 5: Den Haan-Marcet statistics, computed for 500 simulations of $T = 3000$ quarterly observations. The numbers in the parentheses represent the proportion of times the statistic was below and above, respectively, the 5% and 95% percent points of the $\chi^2(11)$ density. In all cases, $\beta = 0.998$.

corroborates the main result of the paper: in all cases the global Chebyshev projection method provides a very accurate solution to the model, whereas a high order perturbation is only adequate for small values of the TFP volatility or where other model parameters (such as γ) eliminate the sensitivity of the value function to σ_z.

4 Conclusion

We have shown that choice of solution method can be critical for production-based asset pricing models with recursive utility. In particular, local perturbation methods have the potential to be inadequate when TFP volatility is calibrated at high levels and when the risk aversion and the discount factor parameters are sufficiently high to make the value function very sensitive to TFP volatility. A global projection method, on the other hand, does quite well under a variety of circumstances. The reason for this result is that the value function in our model has the potential to be highly curved in the direction of TFP volatility, σ_z. In fact, the degree of curvature can be high enough that a local Taylor approximation of the

value function is only suitable over a very small region around the point of approximation. Since perturbation is equivalent to a local Taylor expansion around the deterministic steady state ($\sigma_z = 0$), we find that in certain cases the resulting solution diverges for large σ_z, even for high order approximations. A global approximation method, however, such as Chebyshev projection, is not susceptible to these issues, since it seeks to minimize an error function at a desired level of the TFP volatility.

We show that parameter choice and model calibration is pivotal to the results above. For models that calibrate to quarterly, post-war data, typical values of σ_z and β are low enough to eliminate value function sensitivity to TFP volatility, rendering perturbation solutions perfectly acceptable. Caldara et al. (2009) compares perturbation and projection methods for such parameterizations and demonstrates that both are adequate. Our results diverge from those of Caldara et al. (2009) when we consider parameter values that are relevant for models which calibrate to annual, pre-war data. For these latter parameterizations, the quality of high order perturbation methods degrade.

Local approximations of asset prices can be improved by augmenting the system of perturbation conditions and directly computing expansions for the risk-free rate. This method of approximation weakens the dependency of bond prices on the poorly approximated value function, and results in a more accurate solution. The same is less true of welfare costs: while very small improvements are also observed from direct computations, they are not nearly as striking.

As with many macroeconomic models, we find that the stochastic steady-state distribution of capital is far away from the corresponding deterministic stead-state value, $\widehat{K}_{ss}$. One consideration is to incorporate this information in the perturbation approximation by expanding the Taylor polynomial around the mean of the steady-state distribution, rather than $\widehat{K}_{ss}$. However, we emphasize that while the value function is highly curved in the direction of σ_z, *it is relatively linear* in the direction of the $\widehat{K}_{ss}$, and hence a first order shift in the capital direction is likely to have little effect on a problem that is caused in the volatility direction. Similar expansions around different values of σ_z are not possible, as analytic expressions for the derivatives of the perturbation system can only be found at the deterministic steady-state ($\sigma_z = 0$).

In contrast to the value function, the consumption policy is relatively linear in the direction of σ_z and more curved in the direction of capital. As a result, a local Taylor approximation converges over a large range of values of σ_z, and perturbation has little difficulty in providing good approximations of the endogenous choice variable. For this reason, our results would have little bearing on the practitioner who is only interested in macroeconomic quantities: since quantity dynamics are not sensitive to the value function, perturbation delivers adequate solutions, even for fairly large values of σ_z. Aruoba et al. (2006) only considers quantity dynamics and finds that perturbation is competitive with global solution methods.

Our results are important for individuals who are jointly interested in quantity dynamics and other variables such as asset prices and welfare costs . Since asset prices in a recursive utility model depend crucially on the value function, our choice of solution method has an important impact on their moments (risk-free rates, risk premia, their volatilities, etc.) insofar as the method improves the value function approximation. The same is true of other variables that are tightly linked to the value function, such as welfare costs. While we don't emphasize our particular model as a solution to the joint problem of matching macroeconomic and asset pricing data, we feel that extensions of the model have great potential, and that the problems we have uncovered are likely to be present in other production-based models with recursive utility.

Our general caution is for practitioners to be aware of the potential disadvantages of a local approximation method and, when feasible, to compare it to a global method to ensure adequacy. While we find that Chebyshev projection is competitive with perturbation in terms of computing time for the case of a single state variable, such is not likely to be true of models with many more state variables; as the number of variables increases, a global method will suffer from the curse of dimensionality. In cases such as these (see, for example, Rudebusch and Swanson (2008)) perturbation has the benefit of computational simplicity and, hence, is a natural candidate for an estimation procedure. However, for models where perturbation cannot adequately approximate the value function, and where financial moments or welfare costs are of interest, no degree of computational simplicity can compensate for an incorrect solution. For this reason, we suggest using perturbation in cases where solution adequacy

can be verified against a more robust benchmark.

A Adjustment Cost Parameters

We define the parameters α_1 and α_2 of the adjustment cost function. We want to specify a function that does not impose adjustment costs in the deterministic steady state; i.e. a function that satisfies

$$\phi(x_{ss}) = x_{ss} \tag{39}$$

and

$$\phi'(x_{ss}) = 1, \tag{40}$$

where $x_{ss} = \frac{\widehat{K}_{ss}^{\alpha} - \widehat{C}_{ss}}{\widehat{K}_{ss}}$. First, in the deterministic steady state, Equation (20c) becomes

$$\begin{aligned} \widehat{K}_{ss} &= \frac{1}{\widehat{Z}_{ss}} \left((1-\delta)\widehat{K}_{ss} + \phi\left(\frac{\widehat{K}_{ss}^{\alpha} - \widehat{C}_{ss}}{\widehat{K}_{ss}} \right) \widehat{K}_{ss} \right) \\ \Rightarrow x_{ss} &= \phi^{-1}(\widehat{Z}_{ss} - 1 + \delta). \end{aligned} \tag{41}$$

Since $\phi'(x) = \alpha_1 x^{-1/\xi}$, Equation (40) is satisfied if $\alpha_1 = x_{ss}^{1/\xi}$. Substituting this value for α_1, Equation (39) is then satisfied if

$$\begin{aligned} \alpha_2 &= x_{ss} - \frac{\alpha_1}{1 - 1/\xi} x_{ss}^{1-1/\xi} \\ &= x_{ss} - \frac{1}{1 - 1/\xi} x_{ss} \\ &= \frac{-1/\xi}{1 - 1/\xi} x_{ss} \\ &= \frac{1}{1-\xi} x_{ss}. \end{aligned}$$

Combining Equations (39) and (41), it's clear that

$$\begin{aligned} x_{ss} &= \widehat{Z}_{ss} - 1 + \delta \\ &= \exp(\mu) - 1 + \delta, \end{aligned}$$

from which we conclude

$$\alpha_1 = (\exp(\mu) - 1 + \delta)^{1/\xi} \tag{42}$$

and

$$\alpha_2 = \frac{1}{1-\xi} \left(\exp(\mu) - 1 + \delta \right). \tag{43}$$

B Derivation of Return on Equity

The Lagrangian for the firm's problem is:

$$\max_{\{I_t, K_{t+1}, H_t\}} \mathbb{E}_0 \Bigg[\sum_{t=0}^{\infty} M_{t+1} \Bigg\{ (Z_t H_t)^{1-\alpha} K_t^{\alpha} - W_t H_t - I_t + \mu_t \left(\phi \left(\frac{I_t}{K_t} \right) K_t + (1-\delta) K_t - K_{t+1} \right) \Bigg\} \Bigg]. \quad (44)$$

The first order condition with respect to I_t is

$$-1 + \mu_t \phi' \left(\frac{I_t}{K_t} \right) = 0,$$

which implies,

$$\mu_t = \frac{1}{\phi' \left(\frac{I_t}{K_t} \right)}. \quad (45)$$

The first order condition with respect to K_{t+1} is

$$-\mu_t + \mathbb{E}_t \left[M_{t+1} \, \alpha (Z_{t+1} H_{t+1})^{1-\alpha} K_{t+1}^{\alpha-1} \right] + \mathbb{E}_t \left[M_{t+1} \, \mu_{t+1} \left((1-\delta) - \phi' \left(\frac{I_{t+1}}{K_{t+1}} \right) \frac{I_{t+1}}{K_{t+1}} + \phi \left(\frac{I_{t+1}}{K_{t+1}} \right) \right) \right] = 0. \quad (46)$$

Using Equation (45) we substitute for $\phi'(I_{t+1}/K_{t+1})$ in Equation (46) and rearrange to get

$$\mu_t = \mathbb{E}_t \left[M_{t+1} \left\{ \frac{\alpha (Z_{t+1} H_{t+1})^{1-\alpha} K_{t+1}^{\alpha} - I_{t+1}}{K_{t+1}} + \mu_{t+1} \left(\phi \left(\frac{I_{t+1}}{K_{t+1}} \right) + 1 - \delta \right) \right\} \right]. \quad (47)$$

Substituting for μ_t and μ_{t+1}, and recognizing $Y_t = (Z_t H_t)^{1-\alpha} K_t^{\alpha}$ and $I_t = Y_t - C_t$, we obtain

$$1 = \mathbb{E}_t \left[M_{t+1} \phi' \left(\frac{I_t}{K_t} \right) \left(\frac{(\alpha - 1) Y_{t+1} + C_{t+1}}{K_{t+1}} + \frac{\phi \left(\frac{I_{t+1}}{K_{t+1}} \right) + 1 - \delta}{\phi' \left(\frac{I_{t+1}}{K_{t+1}} \right)} \right) \right] \quad (48)$$

$$= \mathbb{E}_t \left[M_{t+1} R_{t+1}^I \right], \quad (49)$$

where

$$R_{t+1}^I = \phi' \left(\frac{I_t}{K_t} \right) \left(\frac{(\alpha - 1) Y_{t+1} + C_{t+1}}{K_{t+1}} + \frac{\phi \left(\frac{I_{t+1}}{K_{t+1}} \right) + 1 - \delta}{\phi' \left(\frac{I_{t+1}}{K_{t+1}} \right)} \right). \quad (50)$$

Equation (49) is the standard Euler condition for the return on investment, R_{t+1}^I. Moreover, since the production technology and adjustment costs satisfy constant returns to scale, Restoy and Rockinger (1994) prove $R_{t+1}^E = R_{t+1}^I$, where R_{t+1}^E is the unlevered return on equity.

C Maximization Algorithm

We present the binary search algorithm that we use to select the optimal consumption value in line (2.2) of the projection algorithm. This method exploits the monotonicity of the value function with respect to $\widehat{C}_t$ and converges very quickly. It is performed for each value in the capital grid, $\widehat{K}$, at each step in the value function iteration where maximization is performed (non-Howard steps).

1: Set $\tau^c = 0.000001$, $\varepsilon^c = \frac{\tau^c}{10}$, $\Delta^c = 1$, $l^c = 0$, $c^{min} = 0$ and $c^{max} = \widehat{K}_i^{\alpha}$.
2: **while** $\Delta^c > \tau^c$ **do**
3: $c^1 = \frac{c^{max}+c^{min}}{2}$ and $c^2 = c^1 + \varepsilon^c$
4: **for** $m = 1$ to 2 **do**
5: $\widehat{C} = c^m$
6: steps 2.2 – 2.2 of the projection algorithm
7: $v^i = \left\{(1-\beta)\widehat{C}^{\frac{1-\gamma}{\theta}} + \beta\left(Exp(\widehat{K}_i, \widehat{C})\right)^{\frac{1}{\theta}}\right\}^{\frac{\theta}{1-\gamma}}$
8: **end for**
9: **if** $v^1 > v^2$ **then**
10: $c^{max} = c^1$
11: **else**
12: $c^{min} = c^2$
13: **end if**
14: $\Delta^c = c^{max} - c^{min}$
15: **end while**
16: $\widehat{C}_i^* = c^1$.

References

Aruoba, S. B., Fernández-Villaverde, J., and Rubio-Ramírez, J. F. (2006), "Comparing solution methods for dynamic equilibrium economies," *Journal of Economic Dynamics and Control*, 30, 2477–2508.

Bansal, R., Kiku, D., and Yaron, A. (2007), "Risks For the Long Run: Estimation and Inference," *Working Paper*.

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

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

Caldara, D., Fernández-Villaverde, J., Rubio-Ramírez, J. F., and Yao, W. (2009), "Computing DSGE Models with Recursive Preferences," *Working Paper*.

Campanale, C., Castro, R., and Clementi, G. L. (2010), "Asset Pricing in a Production Economy with Chew-Dekel Preferences," *Review of Economic Dynamics*, 13, 379–402.

Campbell, J. Y. and Koo, H. K. (1997), "A comparison of numerical and analytic approximate solutions to an intertemporal consumption choice problem," *Journal of Economic Dynamics and Control*, 21, 273–295.

Collard, F. and Juillard, M. (2001), "Accuracy of stochastic perturbation methods: The case of asset pricing models," *Journal of Economic Dynamics and Control*, 25, 979–999.

Croce, M. M. (2006), "Welfare Costs and Long-Run Consumption Risk in a Production Economy," *Working Paper*.

den Haan, W. J. and de Wind, J. (2009), "How well-behaved are higher-order perturbation solutions?" *Working Paper*.

den Haan, W. J. and Marcet, A. (1994), "Accuracy in Simulations," *The Review of Economics Studies*, 61, 3–17.

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

— (1991), "Substitution, Risk Aversion, and the Temporal Behavior of Consumption and Asset Returns: An Empirical Analysis," *Journal of Political Economy*, 99, 263–286.

Fernández-Villaverde, J. and Rubio-Ramírez, J. (2006), "Solving DSGE models with perturbation methods and a change of variables," *Journal of Economic Dynamics and Control*, 30, 2509–2531.

Heer, B. and Maussner, A. (2005), *Dynamic General Equilibrium Modelling*, Berlin: Springer.

Jermann, U. J. (1998), "Asset pricing in production economies," *Journal of Monetary Economics*, 41, 257–275.

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

— (1998), *Numerical Methods in Economics*, Cambidge, MA: MIT Press.

Judd, K. L. and Guu, S.-M. (1997), "Asymptotic methods for aggregate growth models," *Journal of Economic Dynamics and Control*, 21, 1025–1042.

Kaltenbrunner, G. and Lochstoer, L. A. (2010), "Long-Run Risk through Consumption Smoothing," *Review of Financial Studies*, 23, 3190–3224.

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

Lettau, M. and Uhlig, H. (2000), "Can Habit Formation Be Reconciled with Business Cycle Facts?" *Review of Economic Dynamics*, 3, 79–99.

Lucas, R. E. (1978), "Asset Prices in an Exchange Economy," *Econometrica*, 46, 1429–1445.

Mehra, R. and Prescott, E. C. (1985), "The Equity Premium: A Puzzle," *Journal of Monetary Economics*, 15, 145–161.

Miranda, M. J. and Fackler, P. L. (2002), *Applied Computational Economics and Finance*, Cambidge, MA: MIT Press.

Nocedal, J. and Wright, S. J. (2006), *Numerical Optimization*, New York, NY: Springer.

Restoy, F. and Rockinger, G. M. (1994), "On Stock Market Returns and Returns on Investment," *The Journal of Finance*, XLIX, 543–556.

Rudebusch, G. D. and Swanson, E. T. (2008), "The Bond Premium in a DSGE Model with Long-Run Real and Nominal Risks," *Working Paper*.

Schmitt-Grohé, S. and Uribe, M. (2004), "Solving dynamic general equilibrium models using a second-order approximation to the policy function," *Journal of Economic Dynamics and Control*, 28, 755–775.

Swanson, E., Anderson, G., and Levin, A. (2005), "Higher-Order Perturbation Solutions to Dynamic, Discrete-Time Rational Expectations Models," *Working Paper*.

Tauchen, G. (1986), "Finite state markov-chain approximations to univariate and vector autoregressions," *Economics Letters*, 20, 177–181.

van Binsbergen, J. H., Fernández-Villaverde, J., Rubio-Ramírez, J. F., and Koijen, R. S. (2008), "Likelihood Estimation of DSGE Models with Epstein-Zin Preferences," *Working Paper*.

Weil, P. (1989), "The Equity Premium Puzzle and the Risk-Free Rate Puzzle," *Journal of Monetary Economics*, 24, 401–421.

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