Higher-Order Accurate Simulations and Impulse Responses in DSGE Models with the Nonlinear Moving-Average Policy Function*

Hong Lan[†] Alexander Meyer-Gohde[§]

This Version: October 13, 2011

Abstract

We derive a perturbation solution to nonlinear DSGE models using a nonlinear moving average as an alternative to the standard state-space policy function. Our policy function is particularly suited for impulse response analysis and simulations, being a direct mapping from shocks to endogenous variables up to the order of approximation. The policy function eliminates the need to artificially 'prune' simulations to remove specious explosive behavior, as higher-order approximations inherit the stability from the first-order solution. Euler-equation based error tests demonstrate that the method possesses a high degree of accuracy.

JEL classification: C61, C63, E17

Keywords: Perturbation; Nonlinear moving average; DSGE; Solution methods; Pruning

*We are grateful to Michael Burda and Lutz Weinke, as well as participants of research seminars at the HU Berlin for useful comments, suggestions, and discussions. This research was supported by the DFG through the SFB 649 "Economic Risk". Any and all errors are entirely our own.

[†]Humboldt-Universität zu Berlin, Institut für Wirtschaftstheorie II, Spandauer Straße 1, 10178 Berlin, Germany; Email: lanhong@cms.hu-berlin.de

[§]Humboldt-Universität zu Berlin, Institut für Wirtschaftstheorie II, Spandauer Straße 1, 10178 Berlin, Germany; Tel.: +49-30-2093 5720; Fax: +49-30-2093 5696; E-Mail: alexander.meyer-gohde@wiwi.hu-berlin.de

1 Introduction

We introduce a novel policy function, the nonlinear infinite moving average, to perturbation analysis in dynamic macroeconomics. This direct mapping from shocks to endogenous variables neatly dissects the individual contributions of orders of nonlinearity and uncertainty to the IRFs and provides a perturbation foundation for the 'pruning mechanisms' used to avoid specious explosive behavior in simulations. For economists interested in studying the transmission of shocks in a nonlinear DSGE model, our method offers insight hitherto unavailable.

Our nonlinear moving average policy function chooses as its state variable basis the infinite history of past shocks.[1] This infinite dimensional approach in longstanding in linear models and, for this linear case, delivers the same solution as state space methods.[2] In the nonlinear framework we focus on, however, it provides a different solution. The nature of the policy function, mapping from shocks to endogenous variables of interest, directly enables familiar impulse response analysis, though introducing caveats into the analysis (such as history dependence, asymmetries, a breakdown of superposition and scale invariance, as well as the potential for harmonic distortion).

We show that the stability from the first-order solution is passed on to all higher orders of approximation, producing non-explosive simulation and impulses at all orders of approximation without needing to resort to 'pruning.'[3] Indeed, we show that our solution, up to a deterministic trend, is identical to the 'pruning' procedure of Kim, Kim, Schaumburg, and Sims (2008), demonstrating that their procedure has a solid basis in perturbation theory,[4] just from a different perspective, perhaps.

[1]This is the "external" or "empirical" approach to system theory according to Kalman (1980), who lays out the dichotomy to the 'internal" or "state-variable" approach of the state-space methods, now more common to DSGE practitioners. The nonlinear DSGE perturbation literature initiated by Gaspar and Judd (1997), Judd and Guu (1997), and Judd (1998, ch. 13) has thus far operated solely with state-space methods, see Collard and Juillard (2001b), Collard and Juillard (2001a), Jin and Judd (2002), Schmitt-Grohé and Uribe (2004), Lombardo and Sutherland (2007), Kim, Kim, Schaumburg, and Sims (2008), and Anderson, Levin, and Swanson (2006).

[2]Compare, e.g., Uhlig (1999), Klein (2000), or Sims (2001) with the infinite moving-average representations of Muth (1961), Whiteman (1983) or Taylor (1986). Meyer-Gohde (2010) draws this connection explicitly.

[3]Lombardo and Sutherland (2007) and Lombardo (2010) develop a recursively linear higher-order perturbation state-space method that is, in a sense, naturally pruned, and thus similar to the method we develop here.

[4]Den Haan and De Wind (2010) criticize 'pruning' mechanisms and emphasize that they can be distortive. In Lombardo (2010, p. 9) assessment, pruning "is a work-around to an intrinsic problem of perturbation methods." Our

Contrary to their approach, however, our method extends straightforwardly out to higher orders and we provide the associated endogenous third order ‘pruning’ algorithm explicitly.

Our approach completes the result that terms linear in the perturbation parameter are zero (e.g., Schmitt-Grohé and Uribe (2004) theorem 1): the zero solution of the associated homogenous equations is the unique solution if the first order fundamental polynomial is saddle stability and free of unit roots. We provide explicit calculations out to the third order, adapting Vetter’s (1973) multivariate calculus to extending Lombardo and Sutherland’s (2007) and Gomme and Klein’s (2011) use of linear algebra out past the second order. We implement our approach numerically by providing an add on for the popular Dynare package[5] and show how the Volterra representation of the approximated nonlinear infinite moving average solution allows for a decomposition of the contributing components from all orders to the responses of variables to exogenous shocks. We develop Euler equation error methods for our infinite dimensional policy function and confirm that our method produces accurate approximations.

The rest of the paper is organized as follows. The model and the nonlinear infinite moving average policy function are presented in section 2. In section 3, we develop the numerical perturbation solution of the nonlinear infinite moving average form of the policy function explicitly out to the third order with a matrix calculus that avoids tensor notation. We then apply our method to various incarnations of the stochastic growth model in section 4, starting from the full depreciation and log preferences case with a known analytical solution and finishing with a time-varying volatility version that demonstrates the need for nonlinear methods. We reformulate our nonlinear moving average solution into a traditional state-space solution in section 5, deriving a perturbation-based ‘pruning’ solution. In section 6, we develop Euler-equation-error methods for our infinite-dimensional solution form and use the model of Aruoba, Fernández-Villaverde, and Rubio-Ramírez (2006) for comparability to quantify the accuracy of our method. Finally, section 7 concludes.

method supports the argument that the problem is merely that of the chosen policy function.

[5]See Adjemian, Bastani, Juillard, Mihoubi, Perendia, Ratto, and Villemot (2011).

2 Problem Statement and Solution Form

In this section, we introduce the class of models we analyze and the form of the solution we seek. Our class of models generally follows that used by Dynare.[6] In contrast with the general practice in the literature, the solution we seek is a policy function that is a direct mapping from realizations of the exogenous variables to the the endogenous variables of interest. We will first present the model class, then move on to the solution form, and then conclude this section with the approximated solution that we will seek numerically and the matrix calculus necessary to follow the derivations.

2.1 Model Class

We analyze a family of discrete-time rational expectations models given by

$$0 = E_t[f(y_{t-1}, y_t, y_{t+1}, u_t)] \tag{1}$$

$$\text{where } u_t = \sum_{i=0}^{\infty} N^i \varepsilon_{t-i} \tag{2}$$

f is an $(neq \times 1)$ vector valued function, continuously n-times (the order of approximation to be introduced subsequently) differentiable in all its arguments; y_t is an $(ny \times 1)$ vector of endogenous variables; the vector of exogenous variables u_t is of dimension $(nu \times 1)$ and it is assumed that there are as many equations as endogenous variables $(neq = ny)$.

The eigenvalues of N are assumed all inside the unit circle so that u_t admits an infinite moving average representation; and ε_t is a $(ne \times 1)$ vector of exogenous shocks of equal dimension $(nu = ne)$. Our software add on forces $N = 0$ to align with Dynare.[7]

Additionally, ε_t is assumed independently and identically distributed with the distribution function $\Phi(z)$, such that $E[\varepsilon_t] = 0$ and $E\left[\varepsilon_t{}^{\otimes[n]}\right]$ exists and is finite for all n up to and including the order of approximation to be introduced subsequently.[8]

[6]See Adjemian, Bastani, Juillard, Mihoubi, Perendia, Ratto, and Villemot (2011).

[7]See again Adjemian, Bastani, Juillard, Mihoubi, Perendia, Ratto, and Villemot (2011). Thus in practice, the economist using Dynare must incorporate any serial correlation into the vector y_t. This choice is not made in the exposition here primarily as the admissibility of serial correlation in the exogenous driving force brings our first order derivation in line with earlier moving average approaches for linear models (e.g., Taylor (1986)).

[8]The notation $\varepsilon_t{}^{\otimes[n]}$ represents Kronecker powers, $\varepsilon_t{}^{\otimes[n]}$ is the n'th fold Kronecker product of ε_t with itself:

As is usual in perturbation methods, we introduce an auxiliary parameter $\sigma \in [0,\ 1]$ to scale the uncertainty in the model. The value $\sigma = 1$ corresponds to the "true" stochastic model under study and $\sigma = 0$ represents the deterministic version of the model. Following Anderson, Levin, and Swanson (2006, p. 4), we do not scale the realizations of the exogenous variable up to (including) t with σ, as the realizations of $\{\varepsilon_t, \varepsilon_{t-1}, \ldots\}$ are known at t (hence, there is no uncertainty to scale).

2.2 Solution Form

The policy function, where we take the state vector to be the causal one-sided infinite sequence of shocks, is assumed time invariant for all t, back at least one period, analytic and ergodic, following Anderson, Levin, and Swanson (2006, p. 3).[9] I.e.

$$y_t = y(\sigma, \varepsilon_t, \varepsilon_{t-1}, \ldots) \tag{3}$$

Note that σ enters as a separate argument: the function y is unknown and will itself generally depend on the scaling on uncertainty in the model. Time invariance and scaling uncertainty gives us

$$y_{t-1} = y^-(\sigma, \varepsilon_{t-1}, \varepsilon_{t-2}, \ldots) \tag{4}$$

$$y_{t+1} = y^+(\sigma, \widetilde{\varepsilon}_{t+1}, \varepsilon_t, \varepsilon_{t-1}, \ldots) \text{ where } \widetilde{\varepsilon}_{t+1} \equiv \sigma\varepsilon_{t+1} \tag{5}$$

Due to the assumption of time invariance, y, y^-, and y^+ are the same function, yet they differ in the timing of their arguments. The term $\sigma\varepsilon_{t+1}$ in (5) is the source of uncertainty, via ε_{t+1}, that we are perturbing with σ. The known function u of the exogenous variable rewritten similarly

$$u_t = u(\sigma, \varepsilon_t, \varepsilon_{t-1}, \ldots) = \sum_{i=0}^{\infty} N^i \varepsilon_{t-i} \tag{6}$$

For notational ease in derivation, we will define the vector of total variables x_t

$$x_t \equiv \begin{bmatrix} y'_{t-1} & y'_t & y'_{t+1} & u'_t \end{bmatrix}' \tag{7}$$

x_t is of dimension $(nx \times 1)$ with $(nx = 3ny + ne)$. With the policy function of the form (3), (4) and

$\varepsilon_t \otimes \varepsilon_t \cdots \otimes \varepsilon_t$. For simulations and the like, of course, more specific decisions regarding the distribution of the exogenous processes will have to be made. Note that Kim, Kim, Schaumburg, and Sims (2008, p. 3402) emphasize that distributional assumptions like these are not entirely local assumptions. Note that Dynare (Adjemian, Bastani, Juillard, Mihoubi, Perendia, Ratto, and Villemot 2011) assumes normality of the underlying shocks.

[9] Analyticity is required for the convergence of asymptotic expansion as the order of approximation becomes infinite and ergodicity rules out explosive and nonfundamental solutions.

(5), plus the function of the exogenous variable (6), we can write x_t as

(8) $$x_t = x(\sigma, \widetilde{\varepsilon}_{t+1}, \varepsilon_t, \varepsilon_{t-1}, \ldots)$$

Function x is also assumed time invariant, analytic and ergodic.

2.3 Approximation: Taylor/Volterra Series Approximation

We will approximate the solution, (3), with a Taylor series approximation around a nonstochastic steady state, $\bar{y}$, which is the solution to the function

(9) $$0 = f(\bar{y}, \bar{y}, \bar{y}, 0) = f(\bar{x})$$

that is, the function f in (1) with all shocks, past and present, set to zero. Furthermore

(10) $$\bar{y} = y(0; 0, \ldots)$$

represents the solution (3) evaluated at the nonstochastic steady state.

Following the general practice in the perturbation literature, we will pin down the approximation of the unknown policy function (3) by successively differentiating (1) and solving the resulting systems for the unknown coefficients. The method is detailed in section 3. Notice that, since f is a vector valued function, successive differentiation of f with respect to its arguments, which are vectors in general, will generate a hypercube of partial derivatives. Unlike much of the previous work in the literature, we will adapt the structure of matrix derivatives defined in Vetter (1970) and Vetter (1973) to unfold the hypercube in accordance with the Kronecker product, so that partial derivatives resulting successive differentiation of function f can be collected in two dimensional matrices. This avoids tensor notation and enables the use of standard linear algebra results in deriving our results.[10]

The formal definition of this matrix derivative structure is in appendix A.1.1. This structure will make the presentation of the solution method more transparent—successive differentiation of f to

[10] A similar approach can be found in Gomme and Klein (2011). They use the matrix derivative structure and the associated chain rule of Magnus and Neudecker (2007, ch. 6), which is another way to unfold a second order tensor. The approach, in contrast to ours, does not appear to be easily adapted to orders higher than two. Lombardo and Sutherland (2007) also derive a second order solution without appealing to tensor notation. While their approach may benefit from their use of the vech operator to eliminate redundant quadratic terms, the absence of a mechanical recipe that can be applied to higher orders would appear to favor our methodology.

the desired order of approximation is a mechanical application of the following theorem

Theorem 2.1. *A Multidimensional Calculus*

1. *Matrix Product Rule:*

$$\mathscr{D}_{\underset{s\times 1}{B}^T}\left\{\underset{p\times u}{F}\underset{u\times q}{G}\right\} = F_B\left(\underset{s\times s}{I}\otimes G\right) + FG_B$$

2. *Matrix Chain Rule:*

$$\mathscr{D}_{B^T}\left\{\underset{p\times q}{A}(\underset{u\times 1}{C}(B))\right\} = A_C\left(C_B\otimes \underset{q\times q}{I}\right)$$

3. *Matrix Kronecker Product Rule:*

$$\mathscr{D}_{B^T}\left\{\underset{p\times q}{F}\otimes \underset{u\times v}{G}\right\} = F_B\otimes G + (F\otimes G_B)K_{q,vs}\left(\underset{s\times s}{I}\otimes K_{v,q}\right)$$

   where $K_{q,vs}$ and $K_{v,q}$ are $qvs\times qvs$ and $qv\times qv$ commutation matrices.

4. *Vector Chain Rule:*

$$\mathscr{D}_{B^T}\left\{\underset{p\times 1}{A}(\underset{u\times 1}{C}(B))\right\} = A_C C_B$$

Note that $A_B \equiv \mathscr{D}_{B^T}A(B)$ etc. is abbreviated notation to minimize clutter—see appendix A.1.1.

Proof. See appendix A.1.1. □

An M-th order Taylor approximation of the policy function (3) is then

Corollary 2.2. *An M-th order Taylor Approximation of (3)*

$$y_t = \sum_{m=0}^{M}\frac{1}{m!}\sum_{i_1=0}^{\infty}\sum_{i_2=0}^{\infty}\cdots\sum_{i_m=0}^{\infty}\left[\sum_{n=0}^{M-m}\frac{1}{n!}y_{\sigma^n i_1 i_2\cdots i_m}\sigma^n\right](\varepsilon_{t-i_1}\otimes\varepsilon_{t-i_2}\otimes\cdots\varepsilon_{t-i_m}) \tag{11}$$

Proof. See appendix A.1.3. □

This infinite dimensional Taylor approximation, or Volterra series with kernels $\left[\sum_{n=0}^{M-m}\frac{1}{n!}y_{\sigma^n i_1\cdots i_m}\sigma^n\right]$,[11] directly maps the exogenous innovations to endogenous variables up the M-th order. Viewing terms in powers of the perturbation parameter σ as corrections to the kernels of the Volterra series under

[11] See Priestly (1988, pp. 25–26) and Gourieroux and Jasiak (2005) for a representation theorem.

certainty enables a useful classification of the contributions of uncertainty to the model. That is, with the zeroth kernel being constants, the first kernel being linear in the product space of the history of innovations, the second being quadratic in the same, etc., y_{σ^n} represents the n'th order (in σ) constant correction for uncertainty, $y_{\sigma^n i_1}$ the n'th order (in σ) time-varying correction for uncertainty, $y_{\sigma^n i_1 i_2}$ the n'th order (in σ) asymmetric time-varying correction for uncertainty, and so on.

As the notation in (11) is rather dense, consider the case of $M = 2$. That is, the second-order approximation, given by

$$y_t = \bar{y} + y_\sigma \sigma + \frac{1}{2} y_{\sigma^2} \sigma^2 + \sum_{i=0}^{\infty} (y_i + y_{i\sigma}\sigma)\, \varepsilon_{t-i} + \frac{1}{2} \sum_{j=0}^{\infty} \sum_{i=0}^{\infty} y_{ji} (\varepsilon_{t-j} \otimes \varepsilon_{t-i}) \tag{12}$$

Here, $\bar{y}$, the deterministic steady state, represents the rest point in the absence of uncertainty regarding future shocks; $\sum_{i=0}^{\infty} y_i \varepsilon_{t-i}$ and $\frac{1}{2} \sum_{j=0}^{\infty} \sum_{i=0}^{\infty} y_{ji} (\varepsilon_{t-j} \otimes \varepsilon_{t-i})$ capture the first- and second-order responses of the deterministic (i.e., without uncertainty regarding future shocks) system; $y_\sigma \sigma$ and $\frac{1}{2} y_{\sigma^2} \sigma^2$ are the first- and second-order (in σ) corrections due to the presence of uncertainty to the deterministic steady state (i.e., the second-order accurate stochastic steady state); and $\sum_{i=0}^{\infty} y_{i\sigma} \sigma \varepsilon_{t-i}$ is the first-order (in σ) correction for uncertainty concerning future shocks of the first-order response to the history of shocks. The first-order (in σ) corrections will turn out to be zero in this case, a familiar result from state-space analyses.[12] For the case of $M = 2$, the task at hand is to pin down numerical values for $\bar{y}$, y_i, y_σ, y_{ji}, $y_{i\sigma}$, and y_{σ^2} using the information in (1). We will provide derivations out to $M = 3$ in the next section, providing some additional novelty as explicit derivations of third-order approximations are still rather rare in the literature.[13]

3 Numerical Solution of the Perturbation Approximation

It this section, we lay out the method for solving for the coefficients of the approximated solution. We begin with the first-order approximation and proceed to second and higher-order terms. Solving

[12] See Jin and Judd (2002), Schmitt-Grohé and Uribe (2004), and Kim, Kim, Schaumburg, and Sims (2008).

[13] See Andreasen (forthcoming) for a notable extension of Schmitt-Grohé and Uribe's (2004) method out to the third order. The author's appendix with third-order term occupying almost two pages highlights the advantage of our notation.

for the first-order terms is primarily an application of methods well known in the literature and, similarly to existing state-space methods, solving for higher-order terms operates successively on terms from lower orders with linear methods. Specifically, underlying the expansion in present and past shocks at all orders is a system of difference equations with an identical homogenous component. In contrast to previous methods, we prove the uniqueness of the zero solution for terms of first order in the perturbation parameter through our assumptions of the saddle stability of and rule out unit roots in the fundamental equation of the first order approximation and relate the assumptions to the less easily invertibility assumptions of state-space methods.

The method can be outlined as follows.[14] The equilibrium condition of the model (1)

$$0 = E_t[f(y^-(\sigma,\varepsilon_{t-1},\varepsilon_{t-2},\ldots),y(\sigma,\varepsilon_t,\varepsilon_{t-1},\ldots),y^+(\sigma,\widetilde{\varepsilon}_{t+1},\varepsilon_t,\varepsilon_{t-1},\ldots),u(\sigma,\varepsilon_t,\varepsilon_{t-1},\ldots))] \tag{13}$$

f is a function with arguments σ, ε_t, ε_{t-1}, ... At each order of approximation, we take the collection of derivatives of f from the previous order (for the first-order, start with the function f itself) and

1. differentiate each of the derivatives of f from the previous order with respect to each of its arguments (i.e., σ, ε_t, ε_{t-1}, $\ldots$)

2. evaluate the partial derivatives of f and of y at the nonstochastic steady state

3. apply the expectations operator and evaluate using the given moments

4. set the resulting expression to zero and solve for the unknown partial derivatives of y.

Note, firstly, that the set of f derivatives obtained in step (1) are symbolic and will be differentiated in step (1) of the next higher order and, secondly, the partial derivatives of y obtained in set (4) are numeric, valid at the nonstochastic steady state, will be used (if necessary) in step (2) of the next higher order, and—most importantly—constitute the necessary partial derivatives for the Taylor-Volterra approximation of the policy function y at the nonstochastic steady state.

[14]See Anderson, Levin, and Swanson (2006, p. 9) for a similar outline in their state-space context.

3.1 First-Order Approximation

We proceed by, first, differentiating f with respect to the series of present and past shocks and solving for the resulting infinite-moving average coefficients of y_t and, second, proceeding to the perturbation parameter σ.

Even in the first-order case, the problem is infinite dimensional owing to the infinite moving average representation of the solution. As explained by Taylor (1986, p. 2003) for the linear problem, however, the original set of stochastic difference equations in y_t become deterministic difference equations in the moving-average coefficients of y_t. This motivates our choice of beginning with the unknown terms in the history of shocks and then turning to those in σ, as the problem at higher-orders of approximation will inherit a similar structure.

Differentiating f in (13) with respect to some ε_{t-i} yields

$$\mathscr{D}_{\varepsilon_{t-i}^T} f = f_x x_i \tag{14}$$

Evaluating at the nonstochastic steady state ($\bar{y}$) and setting the resulting expression to zero yields

$$\begin{aligned} &E_t \left[\mathscr{D}_{\varepsilon_{t-i}^T} f \right] \Big|_{\bar{y}} = f_{y^-} y_{i-1} + f_y y_i + f_{y^+} y_{i+1} + f_u u_i = 0 \\ &\text{for } i = 0, 1, \ldots, \text{ with } y_{-1} = 0 \end{aligned} \tag{15}$$

yielding a second order linear deterministic difference equation in the matrices y_i—the derivatives of the vector-valued y function with respect to it's $k-1$'th ε element. That is, y_k contains the moving average coefficients of the elements of y_t with respect to the elements of ε_{t-k}. With appropriate initial conditions, all equal to zero, this is an inhomogeneous version of Anderson and Moore's (1985) saddle-point problem, solved in detail by Anderson (2010).

We make two assumptions regarding the difference equation system (15): the system is saddle stable (i.e., the Blanchard and Kahn (1980) conditions are fulfilled)

Assumption 3.1. *Of the* $2ny$ $z \in \mathbb{C}$ *with det* $\left(f_{y^-} + f_y z + f_{y^+} z^2\right) = 0$, *there are exactly* ny *with* $|z| < 1$.

it is hyperbolic (i.e., Klein's (2000) Assumption 4.4 ruling out eigenvalues on the unit circle),

Assumption 3.2. *There is no* $z \in \mathbb{C}$ *with* $|z| = 1$ *and det* $\left(f_{y^-} + f_y z + f_{y^+} z^2\right) = 0$

The first is standard and the second ensures the uniqueness of terms homogenous in σ (as such, this assumption will take on more relevance beginning with first-order term in σ later in this section). The second is our analog to Jin and Judd's (2002, pp. 12–13) solvability constraint and ensures that any constants do not accumulate without bound as the system is solved in accordance with the stability of the manifold. Intuitively from the state-space perspective, unit roots must be ruled out to allow the state-space solution to be recursively solved ('invertiblity') to yield the nonlinear moving average we work with. As in the case of an explosive state-space solution, the impact of an initial condition on the endogenous variables would fail to vanish and constants (i.e., terms involving the perturbation parameter) would fail to converge when solving out a unit-root state-space solution back into the infinite past.[15]

Anderson's (2010, p. 479) method can be applied under our assumptions 3.1 and 3.2 along with the first-order linear autoregressive u_t (i.e., $u_i = N^i$),[16] delivering the stable solution to (15)

(16) $$y_i = \alpha y_{i-1} + \beta_1 u_i, \text{ with } y_{-1} = 0$$

a convergent recursion from which we can recover the linear moving-average terms or y_i's.[17]

Next we differentiate f in (13) with respect to σ

(17) $$\mathscr{D}_\sigma f = f_x \mathscr{D}_\sigma x$$

(18) $$\text{where } \mathscr{D}_\sigma x = x_\sigma + x_{\tilde{\varepsilon}} \varepsilon_{t+1}$$

Evaluating the foregoing at $\bar{y}$ and setting the resulting expression to zero yields

(19) $$E_t(\mathscr{D}_\sigma f)\Big|_{\bar{y}} = (f_{y^-} + f_y + f_{y^+}) y_\sigma = 0$$

[15] Note that Kim, Kim, Schaumburg, and Sims (2008) can provide a state-space solution in the presence of unit roots and even under explosive conditions. Of course, it cannot be 'inverted' to deliver an infinite moving average in the unit-root/explosive case.

[16] Equivalently, Meyer-Gohde (2010) shows how to apply Klein's (2000) QZ algorithm to this deterministic approach to yield the solution above. Note, as discussed by Meyer-Gohde (2010, pp. 986-987), we are working on a deterministic saddle-point problem in the moving-average coefficients and not on a stochastic saddle-point problem in the endogenous variables themselves.

[17] We have tacitly assumed that this solution exists, see Anderson (2010, p. 483) for the details. In Klein's (2000) notation, Z_{11} of the QZ decomposition must be invertible, the added proviso of translatability.

From assumption (3.2), it follows that

$$det(f_{y^-} + f_y + f_{y^+}) \neq 0 \tag{20}$$

and hence

$$y_\sigma = 0 \tag{21}$$

This not only confirms Schmitt-Grohé and Uribe's (2004) Theorem 1 and others, but also provides the conditions under which it applies. Schmitt-Grohé and Uribe (2004, p. 761) note that their equivalent to (19) "is linear and homogeneous" in their equivalent to y_σ and "[t]hus, if a unique solution exists" it must be zero. Our method improves on their conclusions, giving the condition (the absence of unit roots of assumption 3.2) under which this zero solution is indeed unique.

The first order approximation of the policy function (3) therefore takes the form

$$y_t = \bar{y} + \sum_{i=0}^{\infty} y_i \varepsilon_{t-i},\ i = 0, 1, 2, \ldots \tag{22}$$

trivially extending Muth (1961), Taylor (1986), and others to formal perturbation methods. Note that (22) is independent of σ, emphasizing the certainty equivalent nature of the solution.

3.2 Second-Order Approximation

The only source for the second derivatives of the y function with respect to the shocks, $y_{i,j}$'s, is through differentiation of (14) with respect to some ε_{t-j}, which when evaluated at the nonstochastic steady state yields a set of difference equations with homogenous components identical to the homogenous component in (15). For terms involving the perturbation parameter σ, assumption 3.2, which rules out unit roots in the fundamental equation of the first-order approximation, again plays a crucial role. This is natural: the constant correction for uncertainty induced by the potential for future shocks that enters the solution at the second order becomes arbitrarily large as the system approaches a unit-root system from below. We show that the invertibility condition assumed in the solution package Dynare for the uniqueness of this correction for uncertainty is consistent with our assumptions 3.1 and 3.2. Finally, crossterms, $y_{\sigma,i}$'s, are all equal to zero as a direct consequence of assumption (3.2), as the resulting equations are homogenous in $y_{\sigma,i}$ with coefficient matrices $f_{y^-} + f_y + f_{y^+}$ after the result y_σ from the first order has been taken into account, reaffirming the results of Schmitt-Grohé and Uribe (2004) and others.

We first differentiate f with respect to each and every pair of some ε_{t-i} and ε_{t-j}. As Judd (1998, p. 477) points out, the resulting system of equations remains a linear system, only now in the second derivatives that are being sought

$$\mathscr{D}^2_{\varepsilon^T_{t-j}\varepsilon^T_{t-i}} f = f_{x^2}(x_j \otimes x_i) + f_x x_{j,i} \tag{23}$$

Evaluating at the nonstochastic steady state and setting the resulting expression to zero yields

$$\begin{aligned} &E_t(\mathscr{D}^2_{\varepsilon^T_{t-j}\varepsilon^T_{t-i}} f)\Big|_{\bar{y}} = f_{y^-} y_{j-1,i-1} + f_y y_{j,i} + f_{y^+} y_{j+1,i+1} + f_{x^2}(x_j \otimes x_i) = 0 \\ &\text{for } j,i = 0,1,\ldots, \text{ with } y_{j,i} = 0, \text{ for } j,i < 0 \end{aligned} \tag{24}$$

a second order linear deterministic difference equation in $y_{j,i}$. The coefficients on the homogeneous components of the forgoing and (15) are identical. The inhomogeneous components have a first order Markov representation (see the shifting and transition matrices defined in appendix A.2) in the Kronecker product of the first-order coefficients.[18] The resulting expression is

$$\begin{aligned} &f_{y^-} y_{j-1,i-1} + f_y y_{j,i} + f_{y^+} y_{j+1,i+1} + f_{x^2}(\gamma_1 \otimes \gamma_1)(S_j \otimes S_i) = 0 \\ &\text{for } j,i = 0,1,\ldots, \text{ with } y_{j,i} = 0, \text{ for } j,i < 0 \end{aligned} \tag{25}$$

The solution of the forgoing, analogously to the first order, takes the form

$$y_{j,i} = \alpha y_{j-1,i-1} + \beta_2(S_j \otimes S_i), \text{ with } y_{j,i} = 0, \forall j,i < 0 \tag{26}$$

Note that α in this solution is known. It is the same α as in the first order solution (16) due to the fact that the system (24) and (15) have identical homogeneous components. To determine β_2, we substitute (26) in (24), using the shifting matrices and matching coefficients

$$(f_y + f_{y^+}\alpha)\beta_2 + f_{y^+}\beta_2(\delta_1 \otimes \delta_1) = -f_{x^2}(\gamma_1 \otimes \gamma_1) \tag{27}$$

This is a type of Sylvester equation studied in and solved in detail by Kamenik (2005).

[18] Thus, our nonlinear moving average solution parallels nonlinear state space solutions in a manner analogous to the linear case, where the recursion is in the coefficients as opposed to the variables themselves. Instead of products of the state-variables entering into the solution, we have products of the first-order coefficients.

To determine the partial derivatives of y that involving σ, we first differentiate f twice with respect to σ and some ε_{t-i}. The resulting linear system is

$$\mathscr{D}^2_{\sigma\varepsilon^T_{t-i}} f = f_{x^2}(\mathscr{D}_\sigma x \otimes x_i) + f_x \mathscr{D}_\sigma x_i \tag{28}$$

$$\text{where } \mathscr{D}_\sigma x_i = x_{\sigma,i} + x_{\tilde{\varepsilon},i}(\varepsilon_{t+1} \otimes I_{ne}) \tag{29}$$

Note that the additional potential derivative, $\mathscr{D}^2_{\varepsilon^T_{t-i}\sigma} f$, is simply equal to the derivative in the text, $\mathscr{D}^2_{\sigma\varepsilon^T_{t-i}}$.[19] Evaluating (28) at the nonstochastic steady state, taking into account that $E_t\varepsilon_{t+1}$ and $y_\sigma = 0$ and setting the resulting expression to zero yields

$$E_t(\mathscr{D}^2_{\sigma\varepsilon^T_{t-i}} f)\Big|_{\overline{y}} = f_{y^-} y_{\sigma,i-1} + f_y y_{\sigma,i} + f_{y^+} y_{\sigma,i+1} = 0 \tag{30}$$
$$\text{for } i = 0, 1, \ldots, \text{ with } y_{\sigma,-1} = 0$$

The solution of the forgoing, analogously, takes the form

$$y_{\sigma,i} = \alpha y_{\sigma,i-1}, \text{ for } i = 0, 1, \ldots, \text{ with } y_{\sigma,-1} = 0 \tag{31}$$

Combined with the initial condition $y_{\sigma,-1} = 0$, the forgoing delivers

$$y_{\sigma,i} = 0, \text{ for } i = 0, 1, \ldots \tag{32}$$

This result is analogous to Schmitt-Grohé and Uribe's (2004) Theorem 1. Again, we have improved upon their result by showing not only that the zero solution is a solution (the equation is homogenous), but that it is also the unique solution.

[19] Although the derivative operator works on Kronecker products (i.e. $\mathscr{D}^2_{\sigma\varepsilon^T_{t-i}} = \mathscr{D}^2_{\sigma\otimes\varepsilon^T_{t-i}}$) and although the Kronecker product is not generally commutative, σ is a scalar and, thus, commutation is preserved. This result can be seen by exploiting the properties of the commutation matrix $K_{m,n}$ as follows. Take the first term in $\mathscr{D}^2_{\sigma\varepsilon^T_{t-i}}$, for example, and insert the identity matrix: $f_{x^2} I_{nx^2}(\mathscr{D}_\sigma x \otimes x_i)$. This can be rewritten as $f_{x^2} K_{nx,nx} K_{nx,nx}(\mathscr{D}_\sigma x \otimes x_i)$. Pre-multiplying the Kronecker product of a matrix and a column vector (each with nx rows) with $K_{nx,nx}$ reverses their order (see Theorem 3.1.(ix) of Magnus and Neudecker (1979, p. 384)) and, thus, $K_{nx,nx}(\mathscr{D}_\sigma x \otimes x_i) = x_i \otimes \mathscr{D}_\sigma x$. Now $f_{x^2} = \mathscr{D}^2_{x^T\otimes x^T} f$ and post-multiplying a Kronecker product of row vectors each of dimension nx with $K_{nx,nx}$ reverses their order. But the two row vectors are identical, so reversing their order changes nothing: $f_{x^2} = \mathscr{D}^2_{x^T\otimes x^T} f K_{nx,nx} = \mathscr{D}^2_{x^T\otimes x^T} f = f_{x^2}$. Combining the foregoing two yields $f_{x^2}(\mathscr{D}_\sigma x \otimes x_i) = f_{x^2}(x_i \otimes \mathscr{D}_\sigma x)$. Proceeding likewise with the second term in $\mathscr{D}^2_{\sigma\varepsilon^T_{t-i}}$ completes the argument. Accordingly for higher-order derivatives, the order in which derivatives with respect to σ appear is inconsequential as it is a scalar and we choose to have the σ's appear first.

Next we differentiating f twice with respect to σ, the resulting linear system is

$$\mathscr{D}^2_{\sigma^2} f = f_{x^2}(\mathscr{D}_\sigma x \otimes \mathscr{D}_\sigma x) + f_x \mathscr{D}^2_{\sigma^2} x \tag{33}$$

$$\text{where } \mathscr{D}^2_{\sigma^2} x = x_{\sigma^2} + 2x_{\sigma,\widetilde{\varepsilon}}\varepsilon_{t+1} + x_{\widetilde{\varepsilon}^2}(\varepsilon_{t+1}\otimes\varepsilon_{t+1}) \tag{34}$$

Evaluating the foregoing at $\bar{y}$ and setting the resulting expression to zero yields

$$E_t(\mathscr{D}^2_{\sigma^2} f)\Big|_{\bar{y}} = [f_{y^+}y_{0^2} + f_{y^{+2}}(y_0\otimes y_0)]E_t(\varepsilon_{t+1}\otimes\varepsilon_{t+1}) + (f_{y^-} + f_y + f_{y^+})y_{\sigma^2} = 0 \tag{35}$$

therefore we can recover y_{σ^2} by

$$y_{\sigma^2} = -(f_{y^-} + f_y + f_{y^+})^{-1}[f_{y^+}y_{0^2} + f_{y^{+2}}(y_0\otimes y_0)]E_t(\varepsilon_{t+1}\otimes\varepsilon_{t+1}) \tag{36}$$

By assumption, the second moment of the exogeneous variable is known, hence so is $E_t(\varepsilon_{t+1}\otimes\varepsilon_{t+1})$. As the model approaches a unit root from below, the effect of uncertainty becomes unbounded.

This result is novel, giving additional meaning to the invertibility condition of assumption 3.2: from a state-space perspective, the correction for uncertainty will be accumulated forward starting from the nonstochastic steady state; if the state space contains a unit root, this accumulated correction will become unbounded and there will be no finite stochastic steady state to which the model can converge. To recover y_{σ^2}, Dynare[20] requires instead the invertibility of

$$f_{y^+}(I+\alpha) + f_y \tag{37}$$

This condition is not as easily interpretable as our no-unit-roots invertibility condition in assumption 3.2 as explained above. Yet, the two are equivalent as can be seen from the following. Recall that α solves $f_{y^+}\alpha^2 + f_y\alpha + f_{y^-} = 0$ which can be rearranged as $(f_{y^+}\alpha + f_y)\,\alpha = -f_{y^-}$ or

$$\left(f_{y^+}(I+\alpha) + f_y\right)\alpha = f_{y^+}\alpha - f_{y^-} \tag{38}$$

Adding $f_{y^+} + f_y + f_{y^-}$ to both sides gives

$$f_{y^+} + f_y + f_{y^-} + \left(f_{y^+}(I+\alpha) + f_y\right)\alpha = f_{y^+}(I+\alpha) + f_y \tag{39}$$

solving for $f_{y^+}(I+\alpha) + f_y$ yields

$$f_{y^+} + f_y + f_{y^-} = \left(f_{y^+}(I+\alpha) + f_y\right)(I-\alpha) \tag{40}$$

[20] See http://www.jourdan.ens.fr/~michel/presentations/first_second_order.pdf.

and following from the stability of α

$$\left(f_{y^+}+f_y+f_{y^-}\right)(I-\alpha)^{-1}=f_{y^+}(I+\alpha)+f_y \tag{41}$$

and thus the invertibility of the leading parenthetical term on the left hand side (our invertibility condition from assumption 3.2) is equivalent to the condition on the right hand side used by Dynare.

The second order approximation of the policy function (3) therefore takes form

$$y_t=\bar{y}+\sum_{i=0}^{\infty}y_i\varepsilon_{t-i}+\frac{1}{2}y_{\sigma^2}\sigma^2+\frac{1}{2}\sum_{j=0}^{\infty}\sum_{i=0}^{\infty}y_{j,i}(\varepsilon_{t-j}\otimes\varepsilon_{t-i}) \tag{42}$$

In contrast to the first-order approximation, (42) does depend on σ, with the term $\frac{1}{2}y_{\sigma^2}$ correcting the steady state for uncertainty regarding future shocks. As σ goes from 0 to 1 and we transition from the certain to uncertain model, the rest point of the solution transitions from the nonstochastic steady state $\bar{y}$ to the second-order approximation of the stochastic steady state $\bar{y}+\frac{1}{2}y_{\sigma^2}\sigma^2$. As we are interested in this uncertain version, setting σ to one in (42) gives the second order approximation

$$y_t=\bar{y}+\sum_{i=0}^{\infty}y_i\varepsilon_{t-i}+\frac{1}{2}y_{\sigma^2}+\frac{1}{2}\sum_{j=0}^{\infty}\sum_{i=0}^{\infty}y_{j,i}(\varepsilon_{t-j}\otimes\varepsilon_{t-i}) \tag{43}$$

3.3 Third-Order and Higher Approximations

Computing the third-order approximation largely resembles the computation of the second-order approximation. We first differentiate f three times with respect to each and every triplet of the shocks. The resulting system of equations still remains linear in the third derivatives

$$\begin{aligned}\mathscr{D}^3_{\varepsilon^T_{t-k}\varepsilon^T_{t-j}\varepsilon^T_{t-i}}f=&f_{x^3}(x_k\otimes x_j\otimes x_i)+f_{x^2}(x_{k,j}\otimes x_i)\\&+f_{x^2}[x_j\otimes x_{k,i}]K_{ne,ne^2}(I_{ne}\otimes K_{ne,ne})+f_{x^2}(x_k\otimes x_{j,i})+f_x x_{k,j,i}\end{aligned} \tag{44}$$

Evaluating at the nonstochastic steady state and setting the resulting expression to zero yields

$$\begin{aligned}E_t(\mathscr{D}^3_{\varepsilon^T_{t-k}\varepsilon^T_{t-j}\varepsilon^T_{t-i}}f)\Big|_{\bar{y}}=&f_{y^-}y_{k-1,j-1,i-1}+f_y y_{k,j,i}+f_{y^+}y_{k+1,j+1,i+1}\\&+f_{x^3}(x_k\otimes x_j\otimes x_i)+f_{x^2}(x_{k,j}\otimes x_i)\\&+f_{x^2}(x_j\otimes x_{k,i})K_{ne,ne^2}(I_{ne}\otimes K_{ne,ne})+f_{x^2}(x_k\otimes x_{j,i})\\=&0,\ \text{for } k,j,i=0,1,\ldots,\ \text{with } y_{k,j,i}=0,\ \text{for } k,j,i<0\end{aligned} \tag{45}$$

a linear deterministic second order difference equation in the third derivative $y_{k,j,i}$. The homogeneous components in (45) are identical to those in (15) and (24). The inhomogeneous components can again be rearranged to have a first order Markov representation and by using the shifting and transition matrices defined in appendix A.2, we can write

$$f_{y^-}y_{k-1,j-1,i-1}+f_y y_{k,j,i}+f_{y^+}y_{k+1,j+1,i+1}+\begin{bmatrix} f_{x^3} & f_{x^2} & f_{x^2} & f_{x^2}\end{bmatrix}\gamma_3 S_{k,j,i}=0 \tag{46}$$
$$\text{for } k,j,i=0,1,\ldots,\ \text{with } y_{k,j,i}=0,\ \text{for } k,j,i<0$$

The solution of the forgoing, analogously to lower orders, takes the form

$$y_{k,j,i}=\alpha y_{k-1,j-1,i-1}+\beta_3 S_{k,j,i},\ \text{with } y_{k,j,i}=0,\ \text{for } k,j,i<0 \tag{47}$$

By recursively substituting (47) in (45), using the shifting matrices and matching coefficients, we obtain the following Sylvester equation in β_3

$$(f_y+f_{y^+}\alpha)\beta_3+f_{y^+}\beta_3\delta_3=-\begin{bmatrix} f_{x^3} & f_{x^2} & f_{x^2} & f_{x^2}\end{bmatrix}\gamma_3 \tag{48}$$

Now we move on to the partial derivatives of y function involving the perturbation parameter σ. To determine $y_{\sigma,j,i}$, we differentiate f with respect to some ε_{t-i}, ε_{t-j} and σ sequentially

$$\begin{aligned}\mathscr{D}^3_{\sigma\varepsilon^T_{t-j}\varepsilon^T_{t-i}}f=&f_{x^3}(\mathscr{D}_\sigma x\otimes x_j\otimes x_i)+f_{x^2}(\mathscr{D}_\sigma x\otimes x_{j,i})+f_{x^2}(\mathscr{D}_\sigma x_j\otimes x_i)\\ &+f_{x^2}(x_j\otimes\mathscr{D}_\sigma x_i)K_{ne,ne}+f_x\mathscr{D}_\sigma x_{j,i}\end{aligned} \tag{49}$$

$$\text{where } \mathscr{D}_\sigma x_{j,i}=x_{\sigma,j,i}+x_{\tilde{\varepsilon},j,i}(\varepsilon_{t+1}\otimes I_{ne^2}) \tag{50}$$

Evaluating at $\bar{y}$, setting to zero, and noting the results from lower orders yields

$$E_t(\mathscr{D}^3_{\sigma\varepsilon^T_{t-j}\varepsilon^T_{t-i}}f)\Big|_{\bar{y}}=f_{y^-}y_{\sigma,j-1,i-1}+f_y y_{\sigma,j,i}+f_{y^+}y_{\sigma,j+1,i+1}=0 \tag{51}$$
$$\text{for } j,i=0,1,\ldots,\ \text{with } y_{\sigma,j,i}=0,\ \text{for } j,i<0$$

The solution of the forgoing, again analogously to lower ordes, takes the form

$$y_{\sigma,j,i}=\alpha y_{\sigma,j-1,i-1},\ \text{with } y_{\sigma,j,i}=0,\ \text{for } j,i<0 \tag{52}$$

or

$$y_{\sigma,j,i}=0,\ \text{for } j,i=0,1,\ldots \tag{53}$$

confirming Schmitt-Grohé and Uribe's (2004) conjectured generalization of their Theorem 1.

To determine $y_{\sigma^2,i}$, we differentiate f with respect to some ε_{t-i} and σ twice sequentially

(54)

$$\mathscr{D}^3_{\sigma^2\varepsilon^T_{t-i}} f = f_{x^3}(\mathscr{D}_\sigma x \otimes \mathscr{D}_\sigma x \otimes x_i) + f_{x^2}(\mathscr{D}_\sigma x \otimes \mathscr{D}_\sigma x_i) + f_{x^2}(\mathscr{D}^2_{\sigma^2} x \otimes x_i) + f_{x^2}(\mathscr{D}_\sigma x \otimes \mathscr{D}_\sigma x_i) + f_x \mathscr{D}^2_{\sigma^2} x_i$$

(55)

$$\text{where } \mathscr{D}^2_{\sigma^2} x_i = x_{\sigma^2,i} + 2x_{\sigma,\tilde{\varepsilon},i}(\varepsilon_{t+1} \otimes I_{ne}) + x_{\tilde{\varepsilon}^2,i}(\varepsilon_{t+1} \otimes \varepsilon_{t+1} \otimes I_{ne})$$

Evaluating at the nonstochastic steady state ($\bar{y}$) and setting the resulting expression to zero yields

$$E_t(\mathscr{D}^3_{\sigma^2\varepsilon^T_{t-i}} f)\Big|_{\bar{y}} = f_{x^3}\{[(x_{\tilde{\varepsilon}} \otimes x_{\tilde{\varepsilon}})E_t(\varepsilon_{t+1} \otimes \varepsilon_{t+1})] \otimes x_i\} + 2 f_{x^2}(x_{\tilde{\varepsilon}} \otimes x_{\tilde{\varepsilon} i})[E_t(\varepsilon_{t+1} \otimes \varepsilon_{t+1}) \otimes I_{ne}]$$

$$+ f_{x^2}\{(x_{\sigma^2} \otimes x_i) + ([x_{\tilde{\varepsilon}^2} E_t(\varepsilon_{t+1} \otimes \varepsilon_{t+1})] \otimes x_i)\} + f_x\{x_{\sigma^2,i} + x_{\tilde{\varepsilon}^2,i}[E_t(\varepsilon_{t+1} \otimes \varepsilon_{t+1}) \otimes I_{ne}]\}$$

$$= 0, \text{ for } i = 0, 1, \ldots, \text{ with } y_{-1} = 0 \tag{56}$$

which is still a second order deterministic difference equation. The homogeneous components are packed in $x_{\sigma^2,i}$, and they are identical to those in (15) and (24). The inhomogeneous components can again be rearranged to have a first order Markov representation by using the shifting and transition matrices defined in appendix A.2, thus

$$\begin{aligned} & y_{\sigma^2,i-1} + y_{\sigma^2,i} + y_{\sigma^2,i+1} \\ & + \Big\{ \Big[f_{x^3}(\gamma_4\beta_1 \otimes \gamma_4\beta_1 \otimes \gamma_1) + f_{x^2}([\gamma_4\beta_2(S_0 \otimes S_0)] \otimes \gamma_1) + 2 f_{x^2}(\gamma_4\beta_1 \otimes [\gamma_4\beta_2(S_0 \otimes I)]) \\ & + f_x \gamma_4\beta_3\gamma_5(S_0 \otimes S_0 \otimes I) \Big] [E_t(\varepsilon_{t+1} \otimes \varepsilon_{t+1}) \otimes I_{ne}] + f_{x^2}(x_{\sigma^2} \otimes \gamma_1) \Big\} S_i = 0 \\ & \text{for } i = 0, 1, \ldots, \text{ with } y_{-1} = 0 \end{aligned} \tag{57}$$

The solution of the forgoing takes the form

$$y_{\sigma^2,i} = \alpha y_{\sigma^2,i-1} + \beta_\sigma S_i, \text{ with } y_{\sigma^2,-1} = 0 \tag{58}$$

Substituting (58) in (57) and matching coefficients, we obtain a Sylvester equation in β_σ

(59)

$$(f_y + f_{y^+}\alpha)\beta_\sigma + f_{y^+}\beta_\sigma\delta_1 = -\Big\{ \Big[f_{x^3}(\gamma_4\beta_1 \otimes \gamma_4\beta_1 \otimes \gamma_1) + f_{x^2}([\gamma_4\beta_2(S_0 \otimes S_0)] \otimes \gamma_1)$$

$$+ 2 f_{x^2}(\gamma_4\beta_1 \otimes [\gamma_4\beta_2(S_0 \otimes \delta_1)]) + f_x\gamma_4\beta_3\gamma_5(S_0 \otimes S_0 \otimes \delta_1) \Big] [E_t(\varepsilon_{t+1} \otimes \varepsilon_{t+1}) \otimes I_{ne}] + f_{x^2}(x_{\sigma^2} \otimes \gamma_1) \Big\}$$

To determine y_{σ^3}, we differentiate f with respect to σ three times

$$\mathscr{D}^3_{\sigma^3} f = f_{x^3}(\mathscr{D}_\sigma x \otimes \mathscr{D}_\sigma x \otimes \mathscr{D}_\sigma x) + 2 f_{x^2}(\mathscr{D}_\sigma x \otimes \mathscr{D}^2_{\sigma^2} x) + f_{x^2}(\mathscr{D}^2_{\sigma^2} x \otimes \mathscr{D}_\sigma x) + f_x \mathscr{D}^3_{\sigma^3} x \tag{60}$$

$$\text{where } \mathscr{D}^3_{\sigma^3} x = x_{\sigma^3} + 3x_{\sigma^2,\tilde{\varepsilon}}\varepsilon_{t+1} + 3x_{\sigma,\tilde{\varepsilon}^2}(\varepsilon_{t+1}\otimes\varepsilon_{t+1}) + x_{\tilde{\varepsilon}^3}(\varepsilon_{t+1}\otimes\varepsilon_{t+1}\otimes\varepsilon_{t+1}) \tag{61}$$

Evaluating at the nonstochastic steady state and setting the resulting expression to zero yields

$$\begin{aligned} E_t(\mathscr{D}^3_{\sigma^3} f)\Big|_{\bar{y}} =& f_{x^3}[(x_{\tilde{\varepsilon}}\otimes x_{\tilde{\varepsilon}}\otimes x_{\tilde{\varepsilon}})E_t(\varepsilon_{t+1}\otimes\varepsilon_{t+1}\otimes\varepsilon_{t+1})] + 2 f_{x^2}[E_t(\varepsilon_{t+1}\otimes\varepsilon_{t+1}\otimes\varepsilon_{t+1})(x_{\tilde{\varepsilon}}\otimes x_{\tilde{\varepsilon}^2})] \\ &+ f_{x^2}[(x_{\tilde{\varepsilon}^2}\otimes x_{\tilde{\varepsilon}})E_t(\varepsilon_{t+1}\otimes\varepsilon_{t+1}\otimes\varepsilon_{t+1})] + f_x[y_{\sigma^3} + x_{\tilde{\varepsilon}^3}E_t(\varepsilon_{t+1}\otimes\varepsilon_{t+1}\otimes\varepsilon_{t+1})] \\ =& 0 \end{aligned} \tag{62}$$

Note that, once the third moment of ε_t is introduced, $E_t(\varepsilon_{t+1}\otimes\varepsilon_{t+1}\otimes\varepsilon_{t+1})$ is known. Recovering y_{σ^3} from the forgoing is straightforward under the assumption (3.2). In particular, when ε_t is normally distributed,[21] $E_t(\varepsilon_{t+1}\otimes\varepsilon_{t+1}\otimes\varepsilon_{t+1}) = 0$. Hence

$$y_{\sigma^3} = 0 \tag{63}$$

Therefore, the third order approximation of the policy function (3) takes the form

$$\begin{aligned} y_t =& \bar{y} + \frac{1}{2}y_{\sigma^2}\sigma^2 + \sum_{i=0}^{\infty}\left(y_i + \frac{1}{2}y_{\sigma^2,i}\sigma^2\right)\varepsilon_{t-i} + \frac{1}{2}\sum_{j=0}^{\infty}\sum_{i=0}^{\infty} y_{j,i}(\varepsilon_{t-j}\otimes\varepsilon_{t-i}) \\ &+ \frac{1}{6}\sum_{k=0}^{\infty}\sum_{j=0}^{\infty}\sum_{i=0}^{\infty} y_{k,j,i}(\varepsilon_{t-k}\otimes\varepsilon_{t-j}\otimes\varepsilon_{t-i}) \end{aligned} \tag{64}$$

Again in contrast to the first-order approximation, (64) does depend on σ, with the term $\frac{1}{2}y_{\sigma^2}$ correcting the steady state for uncertainty as in the second-order approximation (42), but now with $\frac{1}{2}y_{\sigma^2,i}\sigma^2$ correcting the first-order kernel for uncertainty; i.e., as σ goes from 0 to 1 and we transition from the certain to uncertain model, we incorporate the additional possibility of a time-varying correction for uncertainty. As we are interested in the original, uncertain formulation, setting σ to one in (64) gives the third-order approximation

$$\begin{aligned} y_t =& \bar{y} + \frac{1}{2}y_{\sigma^2} + \sum_{i=0}^{\infty}\left(y_i + \frac{1}{2}y_{\sigma^2,i}\right)\varepsilon_{t-i} + \frac{1}{2}\sum_{j=0}^{\infty}\sum_{i=0}^{\infty} y_{j,i}(\varepsilon_{t-j}\otimes\varepsilon_{t-i}) \\ &+ \frac{1}{6}\sum_{k=0}^{\infty}\sum_{j=0}^{\infty}\sum_{i=0}^{\infty} y_{k,j,i}(\varepsilon_{t-k}\otimes\varepsilon_{t-j}\otimes\varepsilon_{t-i}) \end{aligned} \tag{65}$$

[21] As is the case in Dynare, see Adjemian, Bastani, Juillard, Mihoubi, Perendia, Ratto, and Villemot (2011).

Higher order approximations of the policy function (3) can be computed using the same steps. Moving through higher-orders of approximation successively, the undetermined partial derivatives of the policy function will be terms of highest order yet considered, ensuring that the leading coefficient matrix is f_x. Thus, for all time varying components, the difference equations in these components will have the same homogenous representation—for non time varying components (i.e. derivatives with respect to σ only), the leading coefficient matrix f_x along with assumption 3.2 ensure the uniqueness of their solution. The inhomogenous elements of the difference equations in the time varying components will be composed of terms of lower order, which are necessarily constants (terms in the given moments and derivatives with respect to σ only) or products of stable recursions (time varying components of lower order). As the latter are likewise stable, we can conclude from assumption 3.1 that the difference equations in all time varying components will be saddle stable; hence, the stability of the first order recursion is passed on to all higher orders.

4 Stochastic Neoclassical Growth Model

In this section, we examine the stochastic neoclassical growth model in several incarnations to demonstrate the techniques developed in the previous sections. This well-studied model has been used in numerous studies comparing numerical techniques and is, thus, the natural choice for a benchmark. We will begin by presenting the general model that encompasses all the various specific cases that we will examine subsequently. Then, starting with log-preferences and full depreciation case with a known solution, we will progress up to time varying volatility version of model used in Aruoba, Fernández-Villaverde, and Rubio-Ramírez (2006).

To that end, consider an infinitely lived representative household seeks to maximize its expected discounted lifetime utility given by

$$E_0\left[\sum_{t=0}^{\infty}\beta^t U\left(C_t, L_t\right)\right] \tag{66}$$

with C_t being consumption, L_t labor, and $\beta \in (0,1)$ the discount factor, subject to

$$C_t + K_t = e^{Z_t} K_{t-1}^{\alpha} L_t^{1-\alpha} + (1-\delta) K_{t-1} \tag{67}$$

where K_t is the capital stock accumulated today for productive purposes tomorrow, Z_t a stochastic productivity process, $\alpha \in [0,1]$ the capital share, and $\delta \in [0,1]$ the depreciation rate. Output Y_t is given by $e^{Z_t} K_{t-1}^{\alpha} L_t^{1-\alpha}$ and investment I_t by $K_t - (1-\delta) K_{t-1}$. Productivity is described by

$$Z_t = \rho_Z Z_{t-1} + e^{\sigma_t} \varepsilon_{Z,t}, \;\; \varepsilon_{Z,t} \sim \mathcal{N}(0,1) \tag{68}$$

with $\rho_Z \in (0,1)$ a persistence parameter, $\varepsilon_{Z,t}$ the innovation to the process, and $exp\sigma_t$ the volatility of the innovations. We allow for time-varying volatility in the form of the following process

$$\sigma_t = (1-\rho_\sigma)\overline{\sigma} + \rho_\sigma \sigma_{t-1} + \tau \varepsilon_{\sigma,t}, \;\; \varepsilon_{\sigma,t} \sim \mathcal{N}(0,1) \tag{69}$$

$\rho_\sigma \in (0,1)$ a persistence parameter, $\overline{\sigma}$ average (log) volatility, $\varepsilon_{\sigma,t}$ innovations to the process, and τ the volatility of the innovations.

The first-order condition include the intertemporal Euler condition equalizing the expected present-discounted utility value of postponing consumption one period to its utility value today

$$U_C(C_t, L_t) = \beta E_t \left[U_C(C_{t+1}, L_{t+1}) \left(\alpha e^{Z_{t+1}} K_t^{\alpha-1} L_{t+1}^{1-\alpha} + 1 - \delta \right) \right] \tag{70}$$

where $U_C(C_t, L_t)$ is the derivative of $U(C_t, L_t)$ with respect to C_t, and the intratemporal condition equalizing the utility cost of marginally increasing labor supply to the utility value of the additional consumption provided therewith

$$-U_L(C_t, L_t) = U_C(C_t, L_t)(1-\alpha) e^{Z_t} K_{t-1}^{\alpha} L_t^{-\alpha} \tag{71}$$

where $U_L(C_t, L_t)$ is the derivative of $U(C_t, L_t)$ with respect to L_t.

4.1 Brock-Mirman Special Case

The first case we will examine is the simple stochastic neoclassical growth model with constant volatility and without a labor-leisure choice under log preferences and complete capital depreciation. This model can be expressed in terms of one endogenous variable, enabling a scalar version of the method to be studied, and possesses a well-known closed-form solution for the state-space policy function. We show how our policy function relates to this well-known state-space version and use our resulting closed-form policy function as a basis for an initial appraisal of our method.

Accordingly, let $U(C_t, L_t)$ in (66) be given by $ln(C_t)$, normalize $L_t = 1$ and set $\delta = 1$ in (67), finally set $\sigma_t = \overline{\sigma}$ in (68). Combining (67) with (70) in this case yields

$$0 = E_t\left[\left(e^{Z_t}K_{t-1}^{\alpha} - K_t\right)^{-1} - \beta\left(e^{Z_{t+1}}K_t^{\alpha} - K_{t+1}\right)^{-1}\left(\alpha e^{Z_{t+1}}K_t^{\alpha-1}\right)\right] \tag{72}$$

This particular case has a well-known closed form solution for the state-space policy function: $K_t = \alpha\beta e^{Z_t}K_{t-1}^{\alpha}$. However, we are interested in its infinite non-linear moving average representation and guess that the logarithm of the solution is linear in the infinite history of technology innovations

$$ln(K_t) = ln(\bar{K}) + \sum_{j=0}^{\infty} b_j \varepsilon_{Z,t-j} \tag{73}$$

Inserting the guess and the infinite moving average representation for Z_t, (72) can be rewritten

$$1 = \alpha\beta E_t\left[\frac{1 - exp\left(\sum_{j=0}^{\infty}\left(\rho^j - b_j + \alpha b_{j-1}\right)\varepsilon_{Z,t-j} - (1-\alpha)\,ln(\bar{K})\right)}{1 - exp\left(\sum_{j=0}^{\infty}\left(\rho^j - b_j + \alpha b_{j-1}\right)\varepsilon_{Z,t+1-j} - (1-\alpha)\,ln(\bar{K})\right)}\right.$$

$$\left. \times exp\left(\sum_{j=0}^{\infty}\left(\rho^j - b_j + \alpha b_{j-1}\right)\varepsilon_{Z,t} - (1-\alpha)\,ln(\bar{K})\right)\right] \tag{74}$$

where $b_{-1} = 0$.

The value and recursion

$$\bar{K} = (\alpha\beta)^{\frac{1}{1-\alpha}}, \;\; b_j = \alpha b_{j-1} + \rho^j, \;\; \text{with } b_{-1} = 0 \tag{75}$$

solve (74) and verify the guess, (73).

Not surprisingly, this solution can also be deduced directly from the known state-space solution. Take logs of $K_t = \alpha\beta e^{Z_t}K_{t-1}^{\alpha}$, yielding $ln(K_t) = ln(\alpha\beta) + Z_t + \alpha ln(K_{t-1})$. Making use of the lag operator, L, and defining $\rho(L) = \sum_{j=0}^{\infty}(\rho L)^j$, the foregoing can be written as $ln(K_t) = (1-\alpha)^{-1}\,ln(\alpha\beta) + (1-\alpha L)^{-1}\rho(L)\,\varepsilon_{Z,t}$ and restating in levels gives

$$K_t = (\alpha\beta)^{\frac{1}{1-\alpha}}\, exp\left((1-\alpha L)^{-1}\rho(L)\,\varepsilon_{Z,t}\right) = (\alpha\beta)^{\frac{1}{1-\alpha}}\, exp\left(\sum_{j=0}^{\infty} b_j \varepsilon_{Z,t-j}\right) \tag{76}$$

where $b(L) = (1-\alpha L)^{-1}\rho(L) = \sum_{j=0}^{\infty} b_j L^j$ as before.

This special case offers a simple check of the numerical approach. We define $\hat{K}_t = ln(K_t)$ and use $K_t = exp(\hat{K}_t)$ to reexpress (72) as[22]

$$(77) \qquad 0 = E_t\left[\left(e^{Z_t+\alpha\hat{K}_{t-1}} - e^{\hat{K}_t}\right)^{-1} - \beta\left(e^{z_{t+1}+\alpha\hat{K}_t} - e^{\hat{K}_{t+1}}\right)^{-1}\left(\alpha e^{Z_{t+1}+(\alpha-1)\hat{K}_t}\right)\right]$$

With this reformulation, the first-order expansion is the true policy rule in this special case. That is (77) can be rewritten as $0 = E_t[f(y_{t-1}, y_t, y_{t+1}, U_t]$ where $y_t = \begin{bmatrix} K_t & Z_t \end{bmatrix}'$ and $U_t = \begin{bmatrix} \varepsilon_{Z,t} \end{bmatrix}$. To check our method, we calculate the kernels of the third order accurate nonlinear moving average solution of (77) out 500 periods, following the parameterization of Hansen's (1985) for the remaining parameters by setting $\alpha = 0.36$, $1/\beta = 1.01$, $\rho = 0.95$, and $\overline{\sigma} = ln(0.00712)$. Our method successfully identifies $y_{j,i}$, $y_{k,j,i}$, and $y_{\sigma^2,i}$ as being zero and the largest absolute difference in y_i from those implied by the analytic solution was 4.3368×10^{-18}. This simple check is far from comprehensive, in this section and especially in section 6 additional and potentially more meaningful measures will be examined. As a first check, this is promising.

4.2 CRRA-Incomplete Depreciation Case

In this case, we relax the complete depreciation and log preferences of the previous section, necessitating an approximation, as no known closed-form solution exists.

Accordingly, let $U(C_t, L_t)$ in (66) be given by $\frac{C_t^{1-\gamma}-1}{1-\gamma}$, normalize $L_t = 1$ and set $\sigma_t = \overline{\sigma}$ in (68).

$$(78) \qquad C_t + K_t = e^{Z_t}K_{t-1}^{\alpha} + (1-\delta)K_{t-1}$$

$$(79) \qquad C_t^{-\gamma} = \beta E_t\left[C_{t+1}^{-\gamma}\left(\alpha e^{Z_{t+1}}K_t^{\alpha-1} + 1 - \delta\right)\right]$$

$$(80) \qquad Z_t = \rho_Z Z_{t-1} + e^{\overline{\sigma}}\varepsilon_{Z,t}$$

Thus $0 = E_t[f(y_{t-1}, y_t, y_{t+1}, U_t]$ where $y_t = \begin{bmatrix} C_t & K_t & Z_t \end{bmatrix}'$ and $U_t = \begin{bmatrix} \varepsilon_{Z,t} \end{bmatrix}$. We reexpress the variables in logs, commensurate with a loglinear approximation. While we maintain $\delta = 0.025$, we set $\gamma = 5$ indicating a substantial departure from log-preferences. Comparing our first-order solution with the impulse response for capital from Uhlig's (1999) exampl0.m for a one standard deviation shock, the largest absolute difference is 9.0093×10^{-15}, confirming the accuracy of the linear terms.

[22]Fernández-Villaverde and Rubio-Ramírez (2006) examine this and other change of variable techniques.

For higher-order approximations, our policy function (3)

$$y_t = y(\sigma, \varepsilon_t, \varepsilon_{t-1}, \varepsilon_{t-2}, \ldots) \tag{81}$$

is particularly suited to enable impulse response analysis. That is, consider a shock in t to an element of ε_t, one measure[23] for the response of y_t through time to this impulse is given by the sequence

$$\begin{aligned} y_t &= y(\sigma, \varepsilon_t, 0, 0, 0, \ldots) \\ y_{t+1} &= y(\sigma, 0, \varepsilon_t, 0, 0, \ldots) \\ y_{t+2} &= y(\sigma, 0, 0, \varepsilon_t, 0, \ldots) \\ &\vdots \end{aligned} \tag{82}$$

[Figure 1 about here.]

Figure 1 depicts the impulse responses and their contributing components from the kernels of different orders for capital and consumption to a positive, one standard deviation shock in $\varepsilon_{Z,t}$.[24] The upper panel displays the impulse responses at first, second, and third order as deviations from their respective (non)stochastic steady states (themselves in the middle right panel) and the first feature to notice is that they are indistinguishable to the eye. This is not surprising, as it is well known that the neoclassical growth model is nearly loglinear. In the middle column of panels in the lower half of each figure, the contributions to the total impulse responses from the second and third-order kernels $y_{i,i}$ and $y_{i,i,i}$ are displayed. Note that these components display multiple 'humps' to either side of the hump in the first-order component (upper-left panel), this is due at least in part to the phenomenon

[23] Note that we are assuming that $y_{t-j} = y(\sigma, 0, 0, \ldots)$, $\forall j > 0$. Fernández-Villaverde, Guerrón-Quintana, Rubio-Ramírez, and Uribe (forthcoming), for example, examine the responses starting from the mean of the ergodic distribution as opposed to the stochastic steady state that we assume. Note that in a nonlinear environment, variables will wander away deterministically from the ergodic mean to the stochastic steady state when the response to a single shock is examined, as the maintenance of variables around the ergodic mean requires the model to be constantly buffeted with shocks. We argue for our measure as it eliminates such deterministic trends in impulse responses.

[24] In terms of the "conceptual difficulties" laid out in Koop, Pesaran, and Potter (1996), we are assuming a particular history of shocks (namely the infinite absence thereof—such interaction will be addressed later), are examining a particular shock realization (positive, one standard deviation: due to the nonlinearity, asymmetries and the absence of scale invariance are a potential confound) and ignore distributional composition issues by examining a realization of a single structural shock irrespective of its potential correlation with other shocks (in this model there is only one shock, so this is moot anyway).

of harmonic distortion discussed in Priestly (1988, p. 27). That both second-order contributions are positive reflects both the fact that, in a stochastic environment, an overaccumulation of capital and, hence, inefficiently high level of consumption is maintained,[25] as can also be seen in the upward correction of the steady states in the rightmost panels, and that the technology shock passes through the exponential function in the production function $e^{Z_t}K_{t-1}^{\alpha}$, adding an additional upward correction to the effect on production of this shock. The lower left panel contains the contributions from $y_{\sigma^2 i}$ the second order (in σ) time-varying correction for risk, this demonstrates an initial wealth effect with consumption increasing and capital decreasing relative to a nonstochastic environment.[26] Nonlinear impulse responses are not scale invariant, as noted also by Fernández-Villaverde, Guerrón-Quintana, Rubio-Ramírez, and Uribe (forthcoming): for example, while the first-order component scales linearly with the magnitude of the shock, the second-order order component scales quadratically. As shocks become larger, a linear approximation would generally not suffice to characterize the dynamics of the model. This is precisely the effect of higher-order terms: as the magnitude of the shock increases, these higher order terms begin to contribute more significantly to the total impulse, attempting to correct the responses for the greater departure from the steady state. For this model, however, one would need to consider shocks of unreasonable magnitude to generate any notable effects from the higher-order terms on the total impulse, reinforcing the conventional wisdom that this model is nearly linear in the variables' logarithms.

[Figure 2 about here.]

In figure 2, the impulse responses to a technology shock with different values (1, 5, and 10) of the CRRA parameter γ are overlayed. Note that for all three values of γ, the first order components dominate. While changes in γ do change the periodicity of the harmonic distortion as well as the shape and sign of the second and third order components, only the constant and time-varying corrections

[25]Overaccumulation and inefficient from the perspective of a nonstochastic environment that is.

[26]Fernández-Villaverde and Rubio-Ramírez (2010) discusses the nonlinear impact of shocks in the production function and similar wealth effects.

for risk display a significant change in magnitude. As γ is increased, the stochastic steady state is associated with higher constant precautionary stocks of capital and the time-varying component displays a magnified wealth effect. At values above 20 (not pictured), the time-varying corrections for risk begin to contribute noticeably to the total impulse, whereas shocks several orders of magnitude larger than a standard deviation are needed to propel the nonlinear kernels to significance.

[Figure 3 about here.]

Figures 3 and 4 draw the second and third order kernels, $y_{j,i}$ and $y_{k,j,i}$, as they depend on differing time separation (potentially $i \neq j \neq k$) of shocks. As likewise discussed in Fernández-Villaverde, Guerrón-Quintana, Rubio-Ramírez, and Uribe (forthcoming), impulse responses are not invariant to the history of shocks. The contributions (though not scaled top the magnitude of the shock) from the second and third-order kernels in the impulse responses in figure 1 can be found along the diagonals of the kernels ($i = j = k$) in 3 and 4. The off diagonal ($i \neq j \neq k$) elements 'correct' for the history of shocks. That is, in addition to the individual second-order contribution that can be found along the diagonals in 3, an off diagonal correction to the second order contribution would be needed for shocks from the past. The deep valleys on either side of both the kernels that bottom out at about fifty periods indicate a substantially persistent nature of the second order kernels.

[Figure 4 about here.]

Additionally, the harmonic distortion mentioned earlier can be seen in the kernels as well. The shapes of the kernels perpendicular to the diagonal have direct analogs in polynomials: on either side of the diagonal of figure 3b, the shape is reminiscent of the parabola of a quadratic equation and the 's' shape of the cubic equation on either side of the diagonal at lower horizons of figure 4b. This bears a word of caution that not too much should be read into the shape itself of the kernels, as they are dictated by the form of the underlying polynomials.

[Figure 5 about here.]

Figure 5 highlights a central component of higher-order impulse responses: the break down of superposition or history dependence of the transfer function. The nonlinear impulse to two shocks at different points in time is not equal to the sum of the individual responses, even after having corrected the individual responses for the higher order. The panels in the figure depict the second-order contributions to the impulse responses of capital and consumption to two positive, one standard deviation technology shocks, spaced 50 periods apart. The dashed line in the top of figure simply adds the individual second order components from each shock together (i.e., presents the total second-order component if superposition were to hold), whereas the solid line additionally contains the second-order cross-component (i.e., presents the true total second-order component). Demonstrating this breakdown of superposition quite vividly, the cross component overwhelms the individual components shortly after the second shocks hits and the second-order contribution to the response of consumption (lower panel) displays a prolonged downward total correction, despite the always positive individual second-order contributions. Although the switch of sign is much briefer for capital (upper panel), the difference from the sum of individual contributions is just as stark and prolonged. In a nonlinear environment, there is no single measure for an impulse response;[27] in starting from the stochastic steady state, however, we remove any deterministic trends in our impulse measure (e.g., starting from the ergodic mean introduces such a trend, see footnote 23).

The standard RBC model is nearly linear and makes much of the analysis here moot. This is, of course, not to be expected for every model and we will now introduce fundamental nonlinearity into the model making the nonlinear analysis essential in understanding the mechanisms at work.

4.3 Labor Margin and Time-Varying Volatility Case

We move on to a time-varying volatility version of Aruoba, Fernández-Villaverde, and Rubio-Ramírez (2006). This is motivated by the application of numerous solution techniques to the model by Aruoba, Fernández-Villaverde, and Rubio-Ramírez (2006) to which we will return in section 6

[27] See, e.g., Gourieroux and Jasiak (2005), Potter (2000), and Koop, Pesaran, and Potter (1996).

when assessing the accuracy of our method. The incorporation of time-varying volatility introduces a fundamental nonlinearity into the model, whose dynamic consequences we will demonstrate require a third-order approximation to be observed (see also Fernández-Villaverde and Rubio-Ramírez (2010) and Fernández-Villaverde, Guerrón-Quintana, Rubio-Ramírez, and Uribe (forthcoming)).

Let $U(C_t, L_t)$ in (66) now be given by $\frac{\left(C_t^{\theta}(1-L_t)^{1-\theta}\right)^{1-\gamma}}{1-\gamma}$. The nonlinear system is now

$$C_t + K_t = e^{Z_t} K_{t-1}^{\alpha} L_t^{1-\alpha} + (1-\delta) K_{t-1} \tag{83}$$

$$\frac{\left(C_t^{\theta}(1-L_t)^{1-\theta}\right)^{1-\gamma}}{C_t} = \beta E_t \left[\frac{\left(C_{t+1}^{\theta}(1-L_{t+1})^{1-\theta}\right)^{1-\gamma}}{C_{t+1}} \left(\alpha e^{Z_{t+1}} K_t^{\alpha-1} L_{t+1}^{1-\alpha} + 1 - \delta\right) \right] \tag{84}$$

$$\frac{1-\theta}{1-L_t} = \frac{\theta}{C_t}(1-\alpha) e^{Z_t} K_{t-1}^{\alpha} L_t^{-\alpha} \tag{85}$$

$$Z_t = \rho_Z Z_{t-1} + e^{\overline{\sigma_t}} \varepsilon_{Z,t} \sigma_t = (1-\rho_\sigma)\overline{\sigma} + \rho_\sigma \sigma_{t-1} + \tau \varepsilon_{\sigma,t} \tag{86}$$

or $0 = E_t[f(y_{t-1}, y_t, y_{t+1}, U_t]$ where $y_t = \begin{bmatrix} C_t & K_t & L_t & Z_t & \sigma_t \end{bmatrix}'$ and $U_t = \begin{bmatrix} \varepsilon_{Z,t} & \varepsilon_{\sigma,t} \end{bmatrix}'$.

We use Aruoba, Fernández-Villaverde, and Rubio-Ramírez's (2006) baseline paramterization for all parameters, except the volatility process (as it was not present in their formulation), whose values are taken from Fernández-Villaverde and Rubio-Ramírez (2010).

We will begin with the responses to a technology shock before returning to the impulse responses to a volatility shock. The results for technology shocks here largely coincide with those of the simpler model, but a few points are worth highlighting. The nonlinear manner in which the technology shock enters into the production function is pictured in figure 7, the panel labeled as being in levels gives the nonlinear response of e^{Z_t} to a technology shock, which contains positive first, second, and third-order components (note that the risk correction terms are all zero, as this merely a transformation of the known stochastic process for productivity). These positive nonlinear components carry over to production, not pictured, expressed in levels, but are essentially eliminated when production is expressed in logs. The time-varying risk correction to the response of labor, figure 6, demonstrates the wealth effect discussed in the previous section.

[Figure 6 about here.]

[Figure 7 about here.]

Figures 8 through 13 display the impulse responses of the model's variables to a positive one standard deviation shock in the volatility process. All variables are expressed in logs (the responses for variables in levels are essentially identical, save for scale). Note that all components except for the time-varying risk component are zero, consistent with the assessment of Fernández-Villaverde and Rubio-Ramírez (2010) and Fernández-Villaverde, Guerrón-Quintana, Rubio-Ramírez, and Uribe (forthcoming) that a third-order approximation is needed to calculate impulse responses to volatility shocks. As a shock to the volatility process by itself affects only the distribution from which future technology shocks will be drawn, any response is of a precautionary nature. Indeed, a precautionary stock of capital, figure 8, is accumulated by reducing consumption, figure 9, and increasing production, figure 10, through increased labor input, figure 12. As volatility subsides, figure 13, and technology shocks from this more highly dispersed distribution fail to materialize (by assumption, we are looking at the response to the volatility shock by itself), the precautionary stock is drawn down through an increase in consumption and reduction of labor. Agents are able to reduce their labor efforts to levels below their stochastic steady state levels while maintaining an elevated level of production by the persistence of the accumulated stock of capital. Besides being consistent with the precautionary behavior one would expect from risk-averse agents, the response of consumption and production move initially in opposite directions, hinting that some of the excess (in standard RBC models vis-vis post-war US data) correlation between output and consumption can be reduced by simply introducing a time-varying distribution for productivity shocks.[28]

[Figure 8 about here.]

[Figure 9 about here.]

[28] Though a quick comparison with the scales of the responses to technology shocks shows that technology shocks will overwhelm the effects of volatility shocks, minimizing the reduction in correlation between output and consumption that could be gained from introducing time-varying volatility.

[Figure 10 about here.]

[Figure 11 about here.]

[Figure 12 about here.]

[Figure 13 about here.]

In figure 14, the second-order cross kernel of capital, in logs, with respect to both shocks is displayed. That is, the figure depicts the second order correction for a volatility shock and a technology shock occurring at differing points in time. This portion of the kernel to the right of the diagonal displays the correction for the cross effects of a shock in volatility having occurred *after* a shock in technology; the kernel here is, of course, zero as any change in the distribution of technology shocks will have no effect on technology shocks that have already been realized. To the left and on the diagonal, however, the kernel is not zero and is simply the first-order kernel with respect to a technology shock scaled by associated change in volatility. This scale diminishes as the volatility shock wanders further back into the past, following from the stable autoregressive process describing volatility.

[Figure 14 about here.]

In sum, the introduction of fundamental nonlinearities, like the stochastic volatility examined here or risk-sensitive preference as studied in Rudebusch and Swanson (forthcoming), strains the reliability of linear methods for assessing the transmission mechanism of shocks in a model economy. While our policy function enables a straight forward one shock impulse response analysis, the history dependence of the transfer function makes clear that one must be wary of relying solely on this familiar tool. In the next section, we will examine simulations, allowing for a history of shocks to unfold through the endogenous variables.

5 Simulations and Pruning

This section serves two purposes. First, to demonstrate that our infinite MA function is not subject to the explosive behavior that has spurned interest in pruning algorithms and that our second-order solution provides a perturbation basis for the pruning algorithm of Kim, Kim, Schaumburg, and Sims (2008). Second, then, to provide the literature with the approximated state-space policy function implied by our approximation of our nonlinear infinite moving-average policy function. We use our nonlinear moving average solution to then formulate a third-order pruning procedure, the first explicit such formulation to our knowledge.

Perturbation methods with state space policy functions, unfortunately, have tendency of delivering spuriously explosive simulations is well established.[29] Kim, Kim, Schaumburg, and Sims (2008) have offered a solution, termed 'pruning', to alleviate this tendency by removing the offending terms of order higher than the approximation induced by the recursive substitution involved in simulations, impulse responses, and the like. Likewise, the method of Lombardo and Sutherland (2007), cast in terms of perturbation by Lombardo (2010), avoids explosive simulations through the recursive linearity of their solution. Both Den Haan and De Wind (2010) and Lombardo (2010) have criticized 'pruning' as being ad hoc and not a valid perturbation approximation.

To demonstrate, we simulate a slightly modified[30] version of the model of section 4.3 for 500 periods and calculate the first, second, and third order accurate simulations using our method, the second and third order simulations from the standard state space approach, and the second-order 'pruned' solution of Kim, Kim, Schaumburg, and Sims (2008). We modify the model of section 4.3 by scaling up the volatility of the model (increasing the standard deviations of both the shocks by factor of 5), following Lombardo (2010) for demonstrational purposes. The top panel of figures 15 through 18, provide the simulated paths under the different methods. In the figures for the endoge-

[29] See, e.g., Aruoba, Fernández-Villaverde, and Rubio-Ramírez (2006), Kim, Kim, Schaumburg, and Sims (2008), and Den Haan and De Wind (2010).

[30] We add auxiliary equations for output $Y_t = e^{Z_t} K_{t-1}^{\alpha} L_t^{1-\alpha}$ and investment $I_t = K_t - (1-\delta) K_{t-1}$ and express all variables but investment in logs.

nous variables, figures 15 through 17, an explosion under the second order state space method can be observed towards the end of the simulation.[31] Interestingly, the third order state space simulation manages to return to the vicinity of the steady state near the end of the simulation, despite production, figure 16, and investment, figure 17, three-quarters of the way through the simulation having decreased significantly more than under other methods (by two orders of magnitude for investment). Our moving average solutions at both the second and third orders remain in the vicinity of the steady state along with the first order and 'pruned' second order solutions. Due to the time-varying volatility, identical under all methods (see figure 18b), the technology process is approximated, figure 18a, but here all the methods (at a given order of approximation) agree on the approximated solution. The likely culprit for the explosive behavior in the state space solutions is the increase in volatility around the three-quarter mark and subsequent substantial negative deviation of technology.

[Figure 15 about here.]

[Figure 16 about here.]

[Figure 17 about here.]

[Figure 18 about here.]

The middle panel of figures 15 through 17, removes the non-'pruned' state space solutions from the graphs. As the magnitude of the drop in productivity increases with the order of approximation, this is reflected by the depth of the ensuing decline of activity. All variables demonstrate a substantial decline around the three-quarter mark, but of roughly the same order of decline in productivity, but return along with the measure of technology back to the vicinity of the steady state. Thus we can reasonably conclude that the substantial movements in the second and third order moving average solution are not an artefact of the kind motivating 'pruning' algorithms, but reflect the

[31] The simulations for consumption and labor (not pictured) are similar to those for capital and prodution respectively.

underlying movements of the model at their respective orders of approximation. Conspicuously absent from the middle panels is a discernable difference between our second order moving average solution and the 'pruned' second order state space solution. The lower panel of the figures plots the differences between these two methods. Unlike the solutions of Lombardo and Sutherland (2007) and Lombardo (2010), the solution provided by our method does differ from that of Kim, Kim, Schaumburg, and Sims's (2008) 'pruned' solution. However, the difference is deterministic with the two solutions trending smoothly towards each other during the simulation. To illustrate this point, figure 19 displays a simulation for the stopck of capital under an alternate parameterization (standard deviations are returned to their levels in section 4.3 an γ is increased to 50). Here the slow transition of all the state space methods (upper panel) to our moving average solutions is clearly visible.

[Figure 19 about here.]

The observed incomplete similarity of our second order moving average solution with the 'pruned' second order state space solution motivate the following proposition, which reformulates our solution from section 3.2 in terms of a solution that is, in the words of Lombardo (2010), recursively linear in the orders of approximation.

Proposition 5.1. *The second-order infinite moving-average solution*

$$y_t = \bar{y} + \frac{1}{2}y_{\sigma^2} + \sum_{i=0}^{\infty} y_i \varepsilon_{t-i} + \frac{1}{2}\sum_{j=0}^{\infty}\sum_{i=0}^{\infty} y_{j,i}(\varepsilon_{t-j} \otimes \varepsilon_{t-i}) \tag{87}$$

can be written as the 'pruned' state-space solution

$$\left(y_t^{(2)} - \bar{y} - \frac{1}{2}y_{\sigma^2}\right) = \alpha\left(y_{t-1}^{(2)} - \bar{y} - \frac{1}{2}y_{\sigma^2}\right) + \beta_1 u_t + \frac{1}{2}\beta_2\left(\begin{bmatrix} y_{t-1}^{(1)} - \bar{y} \\ u_t \end{bmatrix}^{\otimes[2]}\right) \tag{88}$$

where

$$y_t^{(1)} - \bar{y} = \alpha\left(y_{t-1}^{(1)} - \bar{y}\right) + \beta_1 u_t \tag{89}$$

Proof. See appendix A.3. □

With some algebraic effort, we show that our 'pruned' or 'recursively linear' solution is the same as Kim, Kim, Schaumburg, and Sims's (2008) 'pruned' state space solution up to the constant term.

Corollary 5.2. *The second-order 'pruned' state-space solution (88) can be rearranged to conform with the second-order 'pruned' solution of Kim, Kim, Schaumburg, and Sims (2008, p. 3409). The two solutions differ only with respect to the constant term in* σ^2.

Proof. See appendix A.3. □

The difference in the constant term can be explained as follows. The nonstochastic steady state is not the rest point of the second order state space solution, due to the inclusion of the constant term in σ^2. Starting from nonstochastic steady state, the state space solutions will gradually move towards a new steady state, our stochastic steady state $\bar{y} + \frac{1}{2}y_{\sigma^2}$. Note that our stochastic steady state is the rest point of our moving average and 'pruned' (88) solutions, a form of dynamic self consistency.[32]

It is then straightforward to rearrange our solution from section 3.3 to formulate a third-order 'pruned' solution. While an almost trivial extension, this is a novel contribution in the literature, offering some guidance that Lombardo (2010) stating was missing from pruning procedures.

Proposition 5.3. *The third-order infinite moving-average solution*

$$\begin{aligned} y_t =& \bar{y} + \frac{1}{2}y_{\sigma^2} + \sum_{i=0}^{\infty}\left(y_i + \frac{1}{2}y_{\sigma^2,i}\right)\varepsilon_{t-i} + \frac{1}{2}\sum_{j=0}^{\infty}\sum_{i=0}^{\infty}y_{j,i}(\varepsilon_{t-j}\otimes\varepsilon_{t-i}) \\ &+ \frac{1}{6}\sum_{k=0}^{\infty}\sum_{j=0}^{\infty}\sum_{i=0}^{\infty}y_{k,j,i}(\varepsilon_{t-k}\otimes\varepsilon_{t-j}\otimes\varepsilon_{t-i}) \end{aligned} \tag{90}$$

can be written as the 'pruned' state-space solution

$$\begin{aligned} \left(y_t^{(3)} - \bar{y} - \frac{1}{2}y_{\sigma^2}\sigma^2\right) =& \alpha\left(y_{t-1}^{(3)} - \bar{y} - \frac{1}{2}y_{\sigma^2}\sigma^2\right) + \beta_1 u_t + \frac{1}{2}\beta_\sigma\begin{bmatrix} y_{t-1}^{(1)} - \bar{y} \\ u_t \end{bmatrix} + \frac{1}{6}\beta_{3,1}\left(\begin{bmatrix} y_{t-1}^{(1)} - \bar{y} \\ u_t \end{bmatrix}^{\otimes[3]}\right) \\ &+ \beta_{3,4}\left(\begin{bmatrix} y_{t-1}^{(1)} - \bar{y} \\ u_t \end{bmatrix}\right)\otimes\left(\begin{bmatrix} y_{t-1}^{(2)} - \bar{y} - \frac{1}{2}y_{\sigma^2} - \left(y_{t-1}^{(1)} - \bar{y}\right) \\ \frac{1}{2}\begin{bmatrix} y_{t-1}^{(1)} - \bar{y} \\ u_t \end{bmatrix}^{\otimes[2]} \end{bmatrix}\right) \end{aligned} \tag{91}$$

where $y_t^{(1)} - \bar{y}$ *and* $y_t^{(2)} - \bar{y} - \frac{1}{2}y_{\sigma^2}$ *are as given in proposition 5.1 and* $\beta_{3,1}$ *and* $\beta_{3,4}$ *correspond to the first* $(ny+ne)^3$ *and last* $(ny+ne)\left((ny+ne)^2+ny\right)$ *columns of* β_3 *respectively.*

[32]Evers (2010) uses the term "self consistent" to refer to the relation between the rest point of the approximation and of the original problem, which is different in detail but similar in spirit to the consistency discussed in the text.

Proof. See appendix A.3. □

Our nonlinear moving average solution avoids the explosive behavior of non 'pruned' state space solutions without any appeal to ad hoc algorithms, being a direct perturbation approximation of a simulation: a mapping from a history of shocks to endogenous variables. Higher order 'pruning' algorithms can be derived from our mapping, providing a firm basis in perturbation theory for the indeterministic component of current 'pruning' algorithms. Yet, it is still unclear whether the nonlinear moving average provides accurate approximations and it is to this question that we now turn.

6 Accuracy

In this section, we explore the accuracy of our solution method using Euler-equation-error methods.[33] Beside validating the accuracy of our solution method, we add an Euler-equation-error method for assessing the accuracy of an impulse response, enabling the method to address our infinite-dimensional state space.

We choose to examine our method using the model of Aruoba, Fernández-Villaverde, and Rubio-Ramírez (2006), the constant volatility version of the model examined in section 4.3. From Judd (1992), the idea of the Euler-equation accuracy test in the neoclassical growth model is to find a unit-free measure that expresses the one-period optimization error in relation to current consumption. Accordingly, (84) can be rearranged to deliver the Euler-equation error function as[34]

$$EE\,() = 1 - \frac{1}{C_t}\left(\frac{\beta E_t\left[\frac{\left(C_{t+1}^{\theta}(1-L_{t+1})^{1-\theta}\right)^{1-\gamma}}{C_{t+1}}\left(\alpha e^{Z_{t+1}}K_t^{\alpha-1}L_{t+1}^{1-\alpha}+1-\delta\right)\right]}{(1-L_t)^{(1-\theta)(1-\gamma)}}\right)^{\frac{1}{\theta(1-\gamma)-1}} \tag{92}$$

Deviations in (92) from zero are interpreted by Judd (1992) and many others as the relative optimization error that results from using a particular approximation. Expressed in absolute value and in base 10 logarithms, an error of -1 implies a one dollar error for every ten dollars spent and an

[33] See, e.g., Judd (1992), Judd and Guu (1997), and Judd (1998)

[34] Cf. Aruoba, Fernández-Villaverde, and Rubio-Ramírez (2006, p. 2499).

error of -6 implies a one dollar error for every million dollars spent.

The arguments of $EE\left(\right)$ depend on the state space postulated. Standard state-space methods would choose $EE\left(K_{t-1}, Z_t\right)$ or $EE\left(K_{t-1}, Z_{t-1}, \varepsilon_{Z,t}\right)$. Our nonlinear moving average policy function requires $EE\left(\varepsilon_{Z,t}, \varepsilon_{Z,t-1}, \ldots\right)$, rendering the Euler-equation error function an infinite dimensional measure. In line with our presentation of impulse response functions, we examine the following set of Euler-equation error functions, holding all be one shock constant and moving back in time from t, essentially assessing the one-step optimizing error associated with the impulse response functions.

(93) $EE_t = EE\left(\varepsilon_{Z,t}, 0, 0, \ldots\right),\ EE_{t-1} = EE\left(0, \varepsilon_{Z,t-1}, 0, \ldots\right),\ EE_{t-2} = EE\left(0, 0, \varepsilon_{Z,t-2}, \ldots\right),\ \ldots$

We examine a range of shock values for $\varepsilon_{Z,t-j}$ that covers 10 standard deviations in either direction. This is perhaps excessive given the assumption of normality, but enables us to cover the same range for the technology process examined in Aruoba, Fernández-Villaverde, and Rubio-Ramírez (2006) from a single shock. Figure 20 plots EE_t for first through third order approximations in levels and in logs. The first observation is that higher order in levels performs uniformly better than the preceding order—this result is reassuring, but not a given. As Lombardo (2010, p. 22) remarks, although within the radius of convergence the error in approximation goes to zero as the order of approximation becomes infinite, this does not necessary happen monotonically. Indeed, the difference between the second and third order approximations in logs does not paint as clear a picture as in levels. If we restrict our attention to three standard deviation shocks (± 0.021), the third order level and the second and third order log approximations make mistakes no greater than one dollar for everyone ten million spent, hardly an unreasonable error. Of independent interest is the result that the first order approximation in logs is uniformly superior to the first order approximation in levels, standing in contrast to the result of Aruoba, Fernández-Villaverde, and Rubio-Ramírez (2006) to the contrary. As their focus was on the mapping from capital to errors and ours on shocks to errors, it is possible that a the preferred approximation depends on the dimension under study.

[Figure 20 about here.]

In figure 21, plots of EE_{t-j} for $j = 0, 1, \ldots, 100$ for the first order approximations in both levels and logs are provided. Comparing these two figures—let alone incorporating the associated results for the second and third order (not pictured)—is difficult at best. Thus, to facilitate comparison of the different approximations across the different horizons, two measures that reduce to two dimensions will be examined, namely maximal and average Euler equation errors.

[Figure 21 about here.]

First, we plot the maximal Euler-equation errors over a span of 100 periods in figure 22a. I.e.,

(94) $$\max_{-10e^{\overline{\sigma}} < \varepsilon_{Z,t-j} < 10e^{\overline{\sigma}}} \left(EE_{t-j}\right), \text{ for } j = 0, 1, \ldots, 100$$

where $e^{\overline{\sigma}}$ is the constant standard deviation of the technology shock. The figure tends to reinforce the results from examining only shocks in period t: for the level approximations, moving to a higher order uniformly improves the quality of approximation, a first order approximation in logs is to be preferred over a first-order in levels, and the evidence is inconclusive as to whether a third order in level or a second or third order in logarithm approximation is to be preferred.

[Figure 22 about here.]

In our final measure, we graph average Euler-equation errors over a span of 100 periods in figure 22b. In contrast to state space analyses, this measure is relatively easy to calculate, as we merely need to integrate with respect to the known distribution (in this case normal) of the shocks

(95) $$\int EE_{t-j} dF_{\varepsilon_{Z,t-j}}, \text{ for } j = 0, 1, \ldots, 100$$

Weighting the regions of shock realizations most likely to be encountered as defined by the distribution of shocks, we are not forced to make a choice regarding the range of shock values to consider. Again, we note the uniform improvement with higher order for the level approximations, the superiority of the first order approximation in logs, and the ambiguity regarding third in levels and second and third in logs. The average error using a first order in level approximation is around one dollar for every ten thousand spent regardless of horizon. The second order approximations show an improvement as the horizon increases, whereas the third order approximations tend to be lower at first, rise and then fall again. The third order approximation in both levels and logs are associated with an average error of about one dollar for every billion spent regardless of horizon.

We conclude that the nonlinear moving average policy function can provide competitive approximations of the mapping from shocks to endogenous variables. As was the case with Aruoba, Fernández-Villaverde, and Rubio-Ramírez (2006), however, the perturbation methods here deteriorate (not reported) in their extreme parameterization. As all perturbations, our method remains a local method and is subject to all the limitations and reservations that face such methods.

7 Conclusion

We have derived an approximation of the nonlinear moving average form of the policy function, providing explicit derivations up to third order. Besides enabling familiar impulse response analysis techniques, it passes the stability from the first order approximation to higher orders, producing non explosive simulations and thereby endogenizing the 'pruning procedure'. That is, our nonlinear moving average method provides the direct mapping up to the order of approximation from stochastic input to endogenous variables.

The nonlinear perturbation DSGE literature is still in an early stage of development and our method provides a different, yet—from linear methods—familiar, perspective. Standard state-space perturbation methods provide insight into the nonlinear mapping between endogenous variables through time. Yet when the researcher's interest lies in examining the nonlinear mapping from exogenous shocks to endogenous variables, our method has considerable insight to offer.

References

ADJEMIAN, S., H. BASTANI, M. JUILLARD, F. MIHOUBI, G. PERENDIA, M. RATTO, AND S. VILLEMOT (2011): "Dynare: Reference Manual, Version 4," Dynare Working Papers 1, CEPREMAP.

ANDERSON, G. S. (2010): "A Reliable and Computationally Efficient Algorithm for Imposing the Saddle Point Property in Dynamic Models," *Journal of Economic Dynamics and Control*, 34(3), 472–489.

ANDERSON, G. S., A. LEVIN, AND E. SWANSON (2006): "Higher-Order Pertubation Solutions to Dynamic Discrete-Time Rational Expectations Models," Federal Reserve Bank of San Francisco Working Paper Series 2006-01.

ANDERSON, G. S., AND G. MOORE (1985): "A Linear Algebraic Procedure for Solving Linear Perfect Foresight Models," *Economics Letters*, 17(3), 247–252.

ANDREASEN, M. M. (forthcoming): "On the Effects of Rare Disasters and Uncertainty Shocks for Risk Premia in Non-Linear DSGE Models," *Review of Economic Dynamics*.

ARUOBA, S. B., J. FERNÁNDEZ-VILLAVERDE, AND J. F. RUBIO-RAMÍREZ (2006): "Comparing Solution Methods for Dynamic Equilibrium Economies," *Journal of Economic Dynamics and Control*, 30(12), 2477–2508.

BLANCHARD, O. J., AND C. M. KAHN (1980): "The Solution of Linear Difference Models under Rational Expectations," *Econometrica*, 48(5), 1305–1311.

BREWER, J. W. (1978): "Kronecker Products and Matrix Calculus in System Theory," *IEEE Transactions on Circuits and Systems*, 25(9), 772–781.

COLLARD, F., AND M. JUILLARD (2001a): "Accuracy of Stochastic Pertubation Methods: The Case of Asset Pricing Models," *Journal of Economic Dynamics and Control*, 25(6-7), 979–999.

——— (2001b): "A Higher-Order Taylor Expansion Approach to Simulation of Stochastic Forward-Looking Models with an Application to a Nonlinear Phillips Curve Model," *Computational Economics*, 17(2-3), 125–39.

DEN HAAN, W. J., AND J. DE WIND (2010): "Nonlinear and stable perturbation-based approximations," Mimeo December.

EVERS, M. P. (2010): "A Self-Consistent Perturbation Procedure For Solving Dynamic Stochastic General Equilibrium Models," Mimeo January.

FERNÁNDEZ-VILLAVERDE, J., P. A. GUERRÓN-QUINTANA, J. RUBIO-RAMÍREZ, AND M. URIBE (forthcoming): "Risk Matters: The Real Effects of Volatility Shocks," *American Economic Review*.

FERNÁNDEZ-VILLAVERDE, J., AND J. RUBIO-RAMÍREZ (2010): "Macroeconomics and Volatility: Data, Models, and Estimation," Mimeo December.

FERNÁNDEZ-VILLAVERDE, J., AND J. F. RUBIO-RAMÍREZ (2006): "Solving DSGE Models with Perturbation Methods and a Change of Variables," *Journal of Economic Dynamics and Control*, 30(12), 2509–2531.

GASPAR, J., AND K. L. JUDD (1997): "Solving Large-Scale Rational-Expectations Models," *Macroeconomic Dynamics*, 1(01), 45–75.

GOMME, P., AND P. KLEIN (2011): "Second-Order Approximation of Dynamic Models without the Use of Tensors," *Journal of Economic Dynamics and Control*, 35(4), 604–615.

GOURIEROUX, C. S., AND J. JASIAK (2005): "Nonlinear Innovations and Impulse Responses with Application to VaR Sensitivity," *Annales d'Economie et de Statistique*, (78).

HANSEN, G. D. (1985): "Indivisible Labor and the Business Cycle," *Journal of Monetary Economics*, 16(3), 309–327.

JIN, H.-H., AND K. L. JUDD (2002): "Pertubation Methods for General Dynamic Stochastic Models," Mimeo April.

JUDD, K. L. (1992): "Projection Methods for Solving Aggregate Growth Models," *Journal of Economic Theory*, 58(2), 410–452.

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

JUDD, K. L., AND S.-M. GUU (1997): "Asymptotic Methods for Aggregate Growth Models," *Journal of Economic Dynamics and Control*, 21(6), 1025–1042.

KALMAN, R. (1980): "A System-Theoretic Critique of Dynamic Economic Models," *International Journal of Policy Analysis and Information Systems*, 4(1), 3–22.

KAMENIK, O. (2005): "Solving SDGE Models: A New Algorithm for the Sylvester Equation," *Computational Economics*, 25(1), 167–187.

KIM, J., S. KIM, E. SCHAUMBURG, AND C. A. SIMS (2008): "Calculating and Using Second-Order Accurate Solutions of Discrete Time Dynamic Equilibrium Models," *Journal of Economic Dynamics and Control*, 32(11), 3397–3414.

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.

KOOP, G., M. H. PESARAN, AND S. M. POTTER (1996): "Impulse Response Analysis in Nonlinear Multivariate Models," *Journal of Econometrics*, 74(1), 119–147.

LOMBARDO, G. (2010): "On Approximating DSGE Models by Series Expansions," Working Paper Series 1264, European Central Bank.

LOMBARDO, G., AND A. SUTHERLAND (2007): "Computing Second-Order-Accurate Solutions for Rational Expectation Models Using Linear Solution Methods," *Journal of Economic Dynamics and Control*, 31(2), 515–530.

MAGNUS, J. R., AND H. NEUDECKER (1979): "The Commutation Matrix: Some Properties and Applications," *The Annals of Statistics*, 7(9), 383–394.

——— (2007): *Matrix Differential Calculus with Applications in Statistics and Econometrics*. John Wiley & Sons, West Sussex, UK, 3d edn.

MEYER-GOHDE, A. (2010): "Linear Rational-Expectations Models with Lagged Expectations: A Synthetic Method," *Journal of Economic Dynamics and Control*, 34(5), 984–1002.

MUTH, J. F. (1961): "Rational Expectations and the Theory of Price Movements," *Econometrica*, 29(3), 315–335.

POTTER, S. M. (2000): "Nonlinear Impulse Response Functions," *Journal of Economic Dynamics and Control*, 24(10), 1425–1446.

PRIESTLY, M. B. (1988): *Non-Linear and Non-Stationary Time Series Analysis*. Academic Press Ltd., London, UK.

RUDEBUSCH, G. D., AND E. T. SWANSON (forthcoming): "The Bond Premium in a DSGE Model with Long-Run Real and Nominal Risks," *AEJ: Macroeconomics*.

SCHMITT-GROHÉ, S., AND M. URIBE (2004): "Solving Dynamic General Equilibrium Models Using a Second-Order Approximation to the Policy Function," *Journal of Economic Dynamics and Control*, 28(4), 755–775.

SIMS, C. A. (2001): "Solving Linear Rational Expectations Models," *Computational Economics*, 20(1-2), 1–20.

TAYLOR, J. B. (1986): "Econometric Approaches to Stabilization Policy in Stochastic Models of Macroeconomic Fluctuations," in *Handbook of Econometrics*, ed. by Z. Griliches, and M. D. Intriligator, vol. 3 of *Handbook of Econometrics*, chap. 34, pp. 1997–2055. Elsevier.

UHLIG, H. (1999): "A Toolkit for Analysing Nonlinear Dynamic Stochastic Models Easily," in *Computational Methods for the Study of Dynamic Economies*, ed. by R. Marimon, and A. Scott, chap. 3, pp. 30–61. Oxford University Press.

VETTER, W. J. (1970): "Derivative Operations on Matrices," *IEEE Transactions on Automatic Control*, 15(2), 241–244.

——— (1973): "Matrix Calculus Operations and Taylor Expansions," *SIAM Review*, 15(2), 352–369.

WHITEMAN, C. H. (1983): *Linear Rational Expectations Models: A User's Guide*. University of Minnesota Press, Minneapolis, MN.

A Appendices

A.1 Matrix Calculus and Taylor Expansion

A.1.1 Matrix Calculus Definition

Definition A.1. *Matrix Derivative and Commutation Matrix*

1. *Matrix Derivative [See Vetter (1970), Vetter (1973) and Brewer (1978).]*

(A-1)
$$\underset{p\times q}{\mathscr{D}_{b_{kl}}A(B)} \equiv \left[\frac{\partial a_{ij}}{\partial b_{kl}}\right] = \begin{bmatrix} \frac{\partial a_{11}}{\partial b_{kl}} & \cdots & \frac{\partial a_{1q}}{\partial b_{kl}} \\ \vdots & & \vdots \\ \frac{\partial a_{p1}}{\partial b_{kl}} & \cdots & \frac{\partial a_{pq}}{\partial b_{kl}} \end{bmatrix}$$

(A-2)
$$\underset{sp\times tq}{\mathscr{D}_{B}A(B)} \equiv \left[\mathscr{D}_{b_{kl}}A(B)\right] = \left[\begin{array}{c:c:c} \mathscr{D}_{b_{11}}A(B) & \cdots & \mathscr{D}_{b_{1t}}A(B) \\ \hdashline \vdots & & \vdots \\ \hdashline \mathscr{D}_{b_{s1}}A(B) & \cdots & \mathscr{D}_{b_{st}}A(B) \end{array}\right]$$

Structures of higher derivatives are thereby uniquely defined

(A-3)
$$\mathscr{D}^n_{B^n}A(B) \equiv \mathscr{D}_B(\mathscr{D}_B(\cdots(\mathscr{D}_B A(B))\cdots))$$

2. *Abbreviated Notation*

(A-4)
$$\underset{sp\times tq}{A_B} \equiv \mathscr{D}_{B^T}A(B) \text{ and } A_{B^n} \equiv \mathscr{D}^n_{(B^T)^n}A(B)$$

where ^{T} *indicates transposition. Additionally,*

(A-5)
$$A_{CB} \equiv \mathscr{D}_{C^T}\left(\mathscr{D}_{B^T}A(B,C)\right) = \mathscr{D}_{C^T B^T}A(B,C)$$

3. *Commutation Matrix* $K_{a,b}$ *[See Magnus and Neudecker's (1979, p. 383) Theorem 3.1.]*

(A-6)
$$\underset{m\times t}{B} \otimes \underset{n\times s}{A} = K_{m,n}\left(A\otimes B\right)K_{t,s}$$

A.1.2 Proof of theorem 2.1

1. Matrix Product Rule: Combine Vetter's (1973, p. 356) transpose and product rules and examine the special case of an underlying vector variable.

2. Matrix Chain Rule: Combine Vetter's (1973, p. 356) transpose and chain rules and examine the special case of an underlying vector variable.

3. Matrix Kronecker Product Rule: Combine Vetter's (1973, p. 356) transpose and Kronecker rules with an underlying vector variable and adopt Magnus and Neudecker's (1979) notation.

4. Vector Chain Rule: The result follows from the Matrix Chain Rule, setting q to one.

A.1.3 Proof of corollary 2.2

From Vetter (1970, p. 243) and, especially, Vetter (1973, pp. 358–363), a multidimensional Taylor expansion using the structure of derivatives (evaluated at $\bar{B}$) in appendix A.1.1 is given by

$$\text{(A-7)} \qquad \underset{(p\times 1)}{M}\Big(\underset{(s\times 1)}{B}\Big) = M(\bar{B}) + \sum_{n=1}^{N} \frac{1}{n!} \mathscr{D}^{n}_{B^{T^n}} M(\bar{B}) \left(B-\bar{B}\right)^{\otimes[n]} + R_{N+1}\left(\bar{B}, B\right)$$

$$\text{(A-8)} \qquad \text{where } R_{N+1}\left(\bar{B}, B\right) = \frac{1}{N!} \int_{\xi=\bar{B}}^{B} \mathscr{D}^{N+1}_{B^{T^{N+1}}} M(\xi) \left(I_s \otimes (B-\xi)^{\otimes[N]}\right) d\xi$$

Differentiating (3) with respect to all its arguments M times, evaluating at the steady state $\bar{y}$, and noting permutations of the order of differentiation, a Taylor approximation is

$$\begin{aligned}
y_t = & \frac{1}{0!}\left(\frac{1}{0!}\bar{y} + \frac{1}{1!}y_{\sigma}\sigma + \frac{1}{2!}y_{\sigma^2}\sigma^2 + \ldots + \frac{1}{M!}y_{\sigma^M}\sigma^M\right) \\
& + \frac{1}{1!}\sum_{i_1=0}\left(\frac{1}{0!}y_{i_1} + \frac{1}{1!}y_{\sigma i_1}\sigma + \frac{1}{2!}y_{\sigma^2 i_1}\sigma^2 + \ldots + \frac{1}{(M-1)!}y_{\sigma^{M-1} i_1}\sigma^{M-1}\right)\varepsilon_{t-i_1} \\
& + \frac{1}{2!}\sum_{i_1=0}\sum_{i_2=0}\left(\frac{1}{0!}y_{i_1 i_2} + \frac{1}{1!}y_{\sigma i_1 i_2}\sigma + \frac{1}{2!}y_{\sigma^2 i_1 i_2}\sigma^2 + \ldots + \frac{1}{(M-2)!}y_{\sigma^{M-2} i_1 i_2}\sigma^{M-2}\right)\varepsilon_{t-i_1}\otimes\varepsilon_{t-i_2} \\
& \vdots \\
& + \frac{1}{M!}\sum_{i_1=0}\sum_{i_2=0}\cdots\sum_{i_m=0}\frac{1}{0!}y_{i_1 i_2 \cdots i_m}\varepsilon_{t-i_1}\otimes\varepsilon_{t-i_2}\otimes\cdots\varepsilon_{t-i_m}
\end{aligned}$$

Writing the foregoing more compactly yields (11) in the text.

A.2 Auxiliary Matrices

A.2.1 Shifting Matrices

$$\text{(A-9)} \quad \delta_1 = \begin{bmatrix} \underset{ny\times ny}{\alpha} & \underset{ny\times ne}{\beta_1} \\ \underset{ne\times ny}{0} & \underset{ne\times ne}{0} \end{bmatrix} \quad \delta_2 = \begin{bmatrix} \alpha & \beta_2 \\ 0 & \delta_1\otimes\delta_1 \end{bmatrix} \quad \delta_3 = \begin{bmatrix} \delta_1\otimes\delta_1\otimes\delta_1 & 0 & 0 & 0 \\ 0 & \delta_2\otimes\delta_1 & 0 & 0 \\ 0 & 0 & \delta_1\otimes\delta_2 & 0 \\ 0 & 0 & 0 & \delta_1\otimes\delta_2 \end{bmatrix}$$

$$\gamma_1 = \begin{bmatrix} \underset{ny\times ny}{I} & \underset{ny\times ne}{0} \\ \alpha & \beta_1 \\ \alpha^2 & \alpha\beta_1 + \beta_1 N \\ \underset{nu\times ne}{0} & \underset{ne\times ne}{I} \end{bmatrix} \gamma_2 = \begin{bmatrix} I & 0 \\ \alpha & \beta_2 \\ \alpha^2 & \alpha\beta_2 + \beta_2(\delta_1 \otimes \delta_1) \\ 0 & 0 \end{bmatrix} \tag{A-10}$$

$$\gamma_3 = \begin{bmatrix} \gamma_1 \otimes \gamma_1 \otimes \gamma_1 & 0 & 0 & 0 \\ 0 & \gamma_2 \otimes \gamma_1 & 0 & 0 \\ 0 & 0 & \gamma_1 \otimes \gamma_2 & 0 \\ 0 & 0 & 0 & \gamma_1 \otimes \gamma_2 \end{bmatrix} \gamma_4 = \begin{bmatrix} \underset{ny\times ny}{0} \\ \underset{ny\times ny}{0} \\ \underset{ny\times ny}{I} \\ \underset{ne\times ny}{0} \end{bmatrix} \tag{A-11}$$

$$\gamma_5 = \begin{bmatrix} I_{(ny+ne)^3} \\ \begin{bmatrix} \underset{ny\times(ny+ne)^2}{0} \\ I_{(ny+ne)^2} \end{bmatrix} \otimes I_{(ny+ne)^2} \\ I_{(ny+ne)^2} \otimes \begin{bmatrix} \underset{ny\times(ny+ne)^2}{0} \\ I_{(ny+ne)^2} \end{bmatrix} K_{(ny+ne),(ny+ne)^2}\left(I_{ne} \otimes K_{(ny+ne),(ny+ne)}\right) \\ I_{(ny+ne)^2} \otimes \begin{bmatrix} \underset{ny\times(ny+ne)^2}{0} \\ I_{(ny+ne)^2} \end{bmatrix} \end{bmatrix} \tag{A-12}$$

A.2.2 State Spaces for the Markov Representation

$$x_i = \gamma_1 S_i,\; S_i = \begin{bmatrix} y_{i-1} \\ u_i \end{bmatrix}, \text{ and } S_{i+1} = \delta_1 S_i \tag{A-13}$$

$$x_{j,i} = \gamma_2 S_{j,i},\; S_{j,i} = \begin{bmatrix} y_{j-1,i-1} \\ S_j \otimes S_i \end{bmatrix}, \text{ and } S_{j+1,i+1} = \delta_2 S_{j,i} \tag{A-14}$$

$$S_{k,j,i} = \begin{bmatrix} S_k \otimes S_j \otimes S_i \\ S_{k,j} \otimes S_i \\ (S_j \otimes S_{k,i}) K_{ne,ne^2}(I_{ne} \otimes K_{ne,ne}) \\ S_k \otimes S_{j,i} \end{bmatrix} \text{ and } S_{k+1,j+1,i+1} = \delta_3 S_{k,j,i} \tag{A-15}$$

A.3 Pruning Proofs

A.3.1 Proof of Proposition 5.1

Denote the first-order solution, (22), as $y_t^{(1)}$ and the second order solution, (43), $y_t^{(2)}$. Evaluating and rearranging $y_t^{(2)} - \alpha y_{t-1}^{(2)}$ yields

$$y_t^{(2)} - \bar{y} - \frac{1}{2} y_{\sigma^2} - \sum_{i=0}^{\infty} y_i \varepsilon_{t-i} - \alpha \left(y_{t-1}^{(2)} - \bar{y} - \frac{1}{2} y_{\sigma^2} - \sum_{i=0}^{\infty} y_i \varepsilon_{t-i-1} \right)$$

$$\text{(A-16)} \qquad = \frac{1}{2}\sum_{j=0}^{\infty}\sum_{i=0}^{\infty} y_{j,i}(\varepsilon_{t-j}\otimes\varepsilon_{t-i}) - \alpha\left[\frac{1}{2}\sum_{j=0}^{\infty}\sum_{i=0}^{\infty} y_{j,i}(\varepsilon_{t-j-1}\otimes\varepsilon_{t-i-1})\right]$$

Denote with LHS and RHS the left- and right-hand sides of (A-16). LHS can be rewritten as

$$\text{(A-17)} \qquad LHS = \left(y_t^{(2)} - \bar{y}\right) - \alpha\left(y_{t-1}^{(2)} - \bar{y}\right) - (I-\alpha)\frac{1}{2}y_{\sigma^2} - \left[\sum_{i=0}^{\infty}(y_i - \alpha y_{i-1})\varepsilon_{t-i}\right]$$

as $y_i = 0$ for $i < 0$. Using (16) and the linearity of u_t gives

$$\text{(A-18)} \qquad LHS = \left(y_t^{(2)} - \bar{y}\right) - \alpha\left(y_{t-1}^{(2)} - \bar{y}\right) - (I-\alpha)\frac{1}{2}y_{\sigma^2} - \beta_1 u_t$$

RHS can be written as

$$\text{(A-19)} \qquad RHS = \frac{1}{2}\left[\sum_{j=0}^{\infty}\sum_{i=0}^{\infty} y_{j,i}(\varepsilon_{t-j}\otimes\varepsilon_{t-i}) - \alpha\sum_{j=0}^{\infty}\sum_{i=0}^{\infty} y_{j-1,i-1}(\varepsilon_{t-j}\otimes\varepsilon_{t-i})\right]$$

as $y_{j,i} = 0_{ny\times ne^2}$ for $i, j < 0$. Bringing the sums together and applying (26)

$$\text{(A-20)} \qquad RHS = \frac{1}{2}\beta_2\sum_{j=0}^{\infty}\sum_{i=0}^{\infty}\left(S_j\otimes S_i\right)\left(\varepsilon_{t-j}\otimes\varepsilon_{t-i}\right)$$

which, using the mixed product rule,[35] can be rewritten as

$$\text{(A-21)} \qquad RHS = \frac{1}{2}\beta_2\sum_{j=0}^{\infty}\sum_{i=0}^{\infty}\left(S_j\varepsilon_{t-j}\right)\otimes\left(S_i\varepsilon_{t-i}\right)$$

Applying the definition $S_i = \begin{bmatrix} y_{i-1}' & u_i' \end{bmatrix}'$

$$RHS = \frac{1}{2}\beta_2\left(\begin{bmatrix}0_{ny\times ne}\\ I_{ne}\end{bmatrix}\varepsilon_t + \begin{bmatrix}y_0\\ N\end{bmatrix}\varepsilon_{t-1} + \begin{bmatrix}y_1\\ N^2\end{bmatrix}\varepsilon_{t-2} + \cdots\right)\otimes\left(\begin{bmatrix}0_{ny\times ne}\\ I_{ne}\end{bmatrix}\varepsilon_t + \begin{bmatrix}y_0\\ N\end{bmatrix}\varepsilon_{t-1} + \begin{bmatrix}y_1\\ N^2\end{bmatrix}\varepsilon_{t-2} + \cdots\right)$$

and from (22)

$$\text{(A-22)} \qquad RHS = \frac{1}{2}\beta_2\left(\begin{bmatrix}y_{t-1}^{(1)} - \bar{y}\\ u_t\end{bmatrix}\right)\otimes\left(\begin{bmatrix}y_{t-1}^{(1)} - \bar{y}\\ u_t\end{bmatrix}\right)$$

Combining LHS with RHS yields (88) in the text.

A.3.2 Proof of Corollary 5.2

Note firstly that

$$\begin{aligned}
\begin{bmatrix}y_{t-1}^{(1)} - \bar{y}\\ u_t\end{bmatrix}^{\otimes[2]} &= \left(\begin{bmatrix}0_{ny\times 1}\\ u_t\end{bmatrix} + \begin{bmatrix}y_{t-1}^{(1)} - \bar{y}\\ 0_{ne\times 1}\end{bmatrix}\right)\otimes\left(\begin{bmatrix}0_{ny\times 1}\\ u_t\end{bmatrix} + \begin{bmatrix}y_{t-1}^{(1)} - \bar{y}\\ 0_{ne\times 1}\end{bmatrix}\right)\\
&= \begin{bmatrix}y_{t-1}^{(1)} - \bar{y}\\ 0_{ne\times 1}\end{bmatrix}\otimes\begin{bmatrix}y_{t-1}^{(1)} - \bar{y}\\ 0_{ne\times 1}\end{bmatrix} + \begin{bmatrix}y_{t-1}^{(1)} - \bar{y}\\ 0_{ne\times 1}\end{bmatrix}\otimes\begin{bmatrix}0_{ny\times 1}\\ u_t\end{bmatrix}\\
&\quad + \begin{bmatrix}0_{ny\times 1}\\ u_t\end{bmatrix}\otimes\begin{bmatrix}y_{t-1}^{(1)} - \bar{y}\\ 0_{ne\times 1}\end{bmatrix} + \begin{bmatrix}0_{ny\times 1}\\ u_t\end{bmatrix}\otimes\begin{bmatrix}0_{ny\times 1}\\ u_t\end{bmatrix}
\end{aligned}$$

[35] See, e.g., Brewer (1978, p. 773)

$$
\begin{aligned}
&= \left(\begin{bmatrix} I_{ny} \\ 0_{ne\times ny} \end{bmatrix}^{\otimes[2]} \right) \left(\left[y^{(1)}_{t-1} - \bar{y} \right]^{\otimes[2]} \right) + \left(\begin{bmatrix} 0_{ny\times ne} \\ I_{ne} \end{bmatrix}^{\otimes[2]} \right) \left([u_t]^{\otimes[2]} \right) \\
&+ \left(\begin{bmatrix} I_{ny} \\ 0_{ne\times ny} \end{bmatrix} \otimes \begin{bmatrix} 0_{ny\times ne} \\ I_{ne} \end{bmatrix} \right) \left(\left[y^{(1)}_{t-1} - \bar{y} \right] \otimes [u_t] \right) + \left(\begin{bmatrix} 0_{ny\times ne} \\ I_{ne} \end{bmatrix} \otimes \begin{bmatrix} I_{ny} \\ 0_{ne\times ny} \end{bmatrix} \right) \left([u_t] \otimes \left[y^{(1)}_{t-1} - \bar{y} \right] \right)
\end{aligned} \tag{A-23}
$$

using commutation matrices, the final term in the foregoing can be "commuted" as

$$
K_{nyne,nyne} \left(\begin{bmatrix} I_{ny} \\ 0_{ne\times ny} \end{bmatrix} \otimes \begin{bmatrix} 0_{ny\times ne} \\ I_{ne} \end{bmatrix} \right) \left(\left[y^{(1)}_{t-1} - \bar{y} \right] \otimes [u_t] \right) \tag{A-24}
$$

Thus

$$
\begin{aligned}
\begin{bmatrix} y^{(1)}_{t-1} - \bar{y} \\ u_t \end{bmatrix}^{\otimes[2]} &= \left(\begin{bmatrix} I_{ny} \\ 0_{ne\times ny} \end{bmatrix}^{\otimes[2]} \right) \left(\left[y^{(1)}_{t-1} - \bar{y} \right]^{\otimes[2]} \right) + \left(\begin{bmatrix} 0_{ny\times ne} \\ I_{ne} \end{bmatrix}^{\otimes[2]} \right) \left([u_t]^{\otimes[2]} \right) \\
&+ (I_{nyne} + K_{nyne,nyne}) \left(\begin{bmatrix} I_{ny} \\ 0_{ne\times ny} \end{bmatrix} \otimes \begin{bmatrix} 0_{ny\times ne} \\ I_{ne} \end{bmatrix} \right) \left(\left[y^{(1)}_{t-1} - \bar{y} \right] \otimes [u_t] \right)
\end{aligned} \tag{A-25}
$$

As β_2 is symmetric with respect to the Kronecker operator,[36]

$$
\beta_2 \begin{bmatrix} y^{(1)}_{t-1} - \bar{y} \\ u_t \end{bmatrix}^{\otimes[2]} = \beta_{2,11} \left(\left[y^{(1)}_{t-1} - \bar{y} \right]^{\otimes[2]} \right) + \beta_{2,22} \left([u_t]^{\otimes[2]} \right) + 2\beta_{2,12} \left(\left[y^{(1)}_{t-1} - \bar{y} \right] \otimes [u_t] \right) \tag{A-27}
$$

where $\beta_{2,11} \equiv \beta_2 \left(\begin{bmatrix} I_{ny} \\ 0_{ne\times ny} \end{bmatrix}^{\otimes[2]} \right)$, $\beta_{2,22} \equiv \beta_2 \left(\begin{bmatrix} 0_{ny\times ne} \\ I_{ne} \end{bmatrix}^{\otimes[2]} \right)$, $\beta_{2,12} \equiv \beta_2 \left(\begin{bmatrix} I_{ny} \\ 0_{ne\times ny} \end{bmatrix} \otimes \begin{bmatrix} 0_{ny\times ne} \\ I_{ne} \end{bmatrix} \right)$

Equivalently to Kim, Kim, Schaumburg, and Sims (2008, p. 3409),[37] (88) can be written

$$
\begin{aligned}
\left(y^{(2)}_t - \bar{y} \right) &= \alpha \left(y^{(2)}_{t-1} - \bar{y} \right) + (I - \alpha) \frac{1}{2} y_{\sigma^2} + \beta_1 u_t + \frac{1}{2} \beta_{2,11} \left(\left[y^{(1)}_{t-1} - \bar{y} \right]^{\otimes[2]} \right) \\
&+ \beta_{2,12} \left(\left[y^{(1)}_{t-1} - \bar{y} \right] \otimes [u_t] \right) + \frac{1}{2} \beta_{2,22} \left([u_t]^{\otimes[2]} \right)
\end{aligned} \tag{A-28}
$$

A.3.3 Proof of Proposition 5.3

Denote the first-order solution, (22), as $y^{(1)}_t$; the second-order solution, (43), $y^{(2)}_t$; and the third-order solution, (65), as $y^{(3)}_t$. Evaluating and rearranging $y^{(3)}_t - \alpha y^{(3)}_{t-1}$ yields

$$
y^{(3)}_t - \bar{y} - \frac{1}{2} y_{\sigma^2} - \sum_{i=0}^{\infty} y_i \varepsilon_{t-i} - \frac{1}{2} \sum_{j=0}^{\infty} \sum_{i=0}^{\infty} y_{j,i} (\varepsilon_{t-j} \otimes \varepsilon_{t-i})
$$

---

[36]That is, $\tilde{\beta}_2 \equiv \beta_2 K_{nyne,nyne}$ also solves (27): (recalling $f_{x^2} K_{3ny+ne,3ny+ne} = f_{x^2}$, see footnote 19.)

$$
\begin{aligned}
(f_y + f_{y^+}\alpha)\beta_2 K_{nyne,nyne} + f_{y^+}\beta_2(\delta_1 \otimes \delta_1) K_{nyne,nyne} &= -f_{x^2}(\gamma_1 \otimes \gamma_1) K_{nyne,nyne} \\
(f_y + f_{y^+}\alpha)\tilde{\beta}_2 + f_{y^+}\beta_2 K_{nyne,nyne}(\delta_1 \otimes \delta_1) &= -f_{x^2} K_{3ny+ne,3ny+ne}(\gamma_1 \otimes \gamma_1) \\
(f_y + f_{y^+}\alpha)\tilde{\beta}_2 + f_{y^+}\tilde{\beta}_2(\delta_1 \otimes \delta_1) &= -f_{x^2}(\gamma_1 \otimes \gamma_1)
\end{aligned} \tag{A-26}
$$

[37]Except for the absence of the accumulating term $(I - \alpha)$ modifying the term σ^2 in their version.

$$-\alpha\left(y^{(3)}_{t-1}-\bar{y}-\frac{1}{2}y_{\sigma^2}-\sum_{i=0}^{\infty}y_i\varepsilon_{t-i-1}-\frac{1}{2}\sum_{j=0}^{\infty}\sum_{i=0}^{\infty}y_{j,i}(\varepsilon_{t-j-1}\otimes\varepsilon_{t-i-1})\right)$$
$$=\frac{1}{2}\sum_{i=0}^{\infty}\left(y_{\sigma^2,i}\sigma^2-\alpha y_{\sigma^2,i-1}\sigma^2\right)\varepsilon_{t-i}+\frac{1}{6}\sum_{k=0}^{\infty}\sum_{j=0}^{\infty}\sum_{i=0}^{\infty}\left(y_{k,j,i}-\alpha y_{k-1,j-1,i-1}\right)(\varepsilon_{t-k}\otimes\varepsilon_{t-j}\otimes\varepsilon_{t-i})$$

The left-hand side of the foregoing can be rewritten as

$$y^{(3)}_t-\alpha y^{(3)}_{t-1}-(I-\alpha)\left(\bar{y}-\frac{1}{2}y_{\sigma^2}\right)-\beta_1 u_t+\frac{1}{2}\beta_2\left(\begin{bmatrix}y^{(1)}_{t-1}-\bar{y}\\ u_t\end{bmatrix}^{\otimes[2]}\right)\tag{A-29}$$

the first sum on right-hand side as $\frac{1}{2}\beta_\sigma\begin{bmatrix}y^{(1)}_{t-1}-\bar{y}\\ u_t\end{bmatrix}$ and the final term on the right-hand side as

$$\frac{1}{6}\sum_{k=0}^{\infty}\sum_{j=0}^{\infty}\sum_{i=0}^{\infty}\beta_3 S_{k,j,i}(\varepsilon_{t-k}\otimes\varepsilon_{t-j}\otimes\varepsilon_{t-i})$$
$$=\frac{1}{6}\sum_{k=0}^{\infty}\sum_{j=0}^{\infty}\sum_{i=0}^{\infty}\beta_3\begin{bmatrix}S_k\otimes S_j\otimes S_i\\ S_{k,j}\otimes S_i\\ (S_j\otimes S_{k,i})K_{ne,ne^2}(I_{ne}\otimes K_{ne,ne})\\ S_k\otimes S_{j,i}\end{bmatrix}(\varepsilon_{t-k}\otimes\varepsilon_{t-j}\otimes\varepsilon_{t-i})\tag{A-30}$$

which, noting the properties of the commutation matrices can be written

$$=\frac{1}{6}\sum_{k=0}^{\infty}\sum_{j=0}^{\infty}\sum_{i=0}^{\infty}\beta_3\begin{bmatrix}(S_k\varepsilon_{t-k})\otimes\left(S_j\varepsilon_{t-j}\right)\otimes(S_i\varepsilon_{t-i})\\ \left(S_{k,j}(\varepsilon_{t-k}\otimes\varepsilon_{t-j})\right)\otimes(S_i\varepsilon_{t-i})\\ \left(S_j\varepsilon_{t-i}\right)\otimes\left(S_{k,i}(\varepsilon_{t-k}\otimes\varepsilon_{t-i})\right)\\ (S_k\varepsilon_{t-k})\otimes\left(S_{j,i}(\varepsilon_{t-j}\otimes\varepsilon_{t-i})\right)\end{bmatrix}\tag{A-31}$$

partitioning β_3 comfortably, exploiting the commutation matrix, and noting the results from the proof of proposition 5.1 at the beginning of this appendix

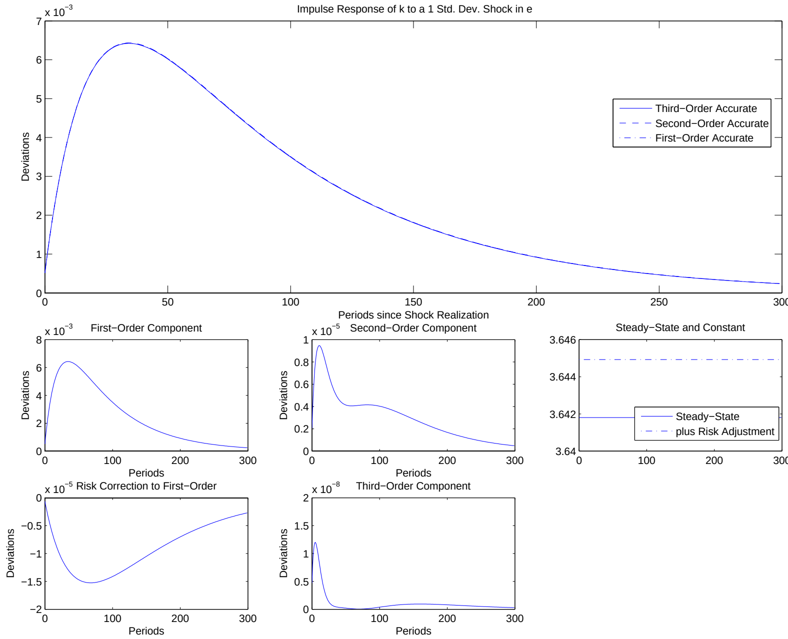
$$=\frac{1}{6}\beta_{3,1}\begin{bmatrix}y^{(1)}_{t-1}-\bar{y}\\ u_t\end{bmatrix}^{\otimes[3]}$$
$$+\frac{1}{3}\left[\beta_{3,2}K_{ny+(ny+ne)^2,ny+ne}+\beta_{3,3}+\beta_{3,4}\right]\begin{bmatrix}y^{(1)}_{t-1}-\bar{y}\\ u_t\end{bmatrix}\otimes\begin{bmatrix}y^{(2)}_{t-1}-\bar{y}-\frac{1}{2}y_{\sigma^2}-\left(y^{(1)}_{t-1}-\bar{y}\right)\\ \frac{1}{2}\begin{bmatrix}y^{(1)}_{t-1}-\bar{y}\\ u_t\end{bmatrix}^{\otimes[2]}\end{bmatrix}\tag{A-32}$$

Finally, noting that $\beta_{3,2}K_{ny+(ny+ne)^2,ny+ne}$, $\beta_{3,3}$, and $\beta_{3,4}$ all solve the same Sylvester equation[38]

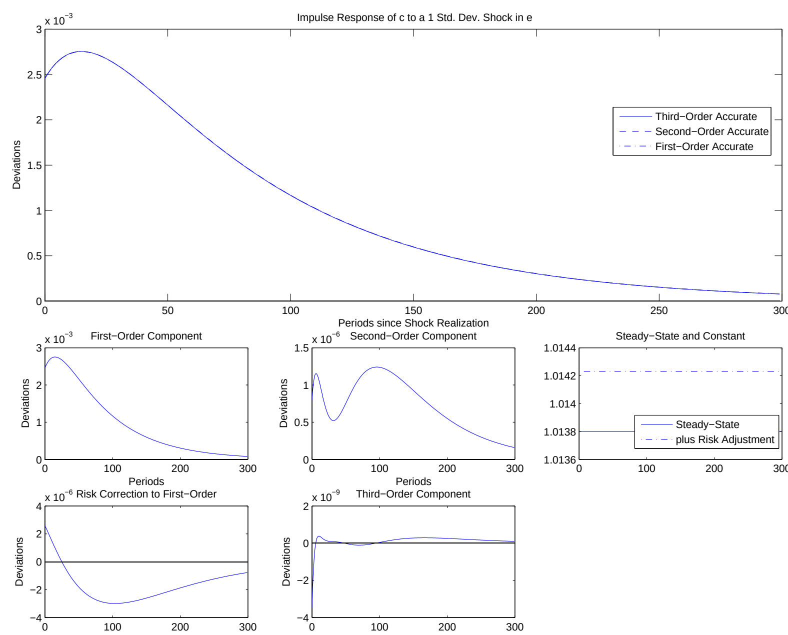
$$=\frac{1}{6}\beta_{3,1}\begin{bmatrix}y^{(1)}_{t-1}-\bar{y}\\ u_t\end{bmatrix}^{\otimes[3]}+\beta_{3,4}\begin{bmatrix}y^{(1)}_{t-1}-\bar{y}\\ u_t\end{bmatrix}\otimes\begin{bmatrix}y^{(2)}_{t-1}-\bar{y}-\frac{1}{2}y_{\sigma^2}-\left(y^{(1)}_{t-1}-\bar{y}\right)\\ \frac{1}{2}\begin{bmatrix}y^{(1)}_{t-1}-\bar{y}\\ u_t\end{bmatrix}^{\otimes[2]}\end{bmatrix}\tag{A-33}$$

Putting all the pieces together yields (91).

[38]See footnote 36, the Sylvester equation for $\beta_{3,2}K_{ny+(ny+ne)^2,ny+ne}$ can be rearranged analogously.

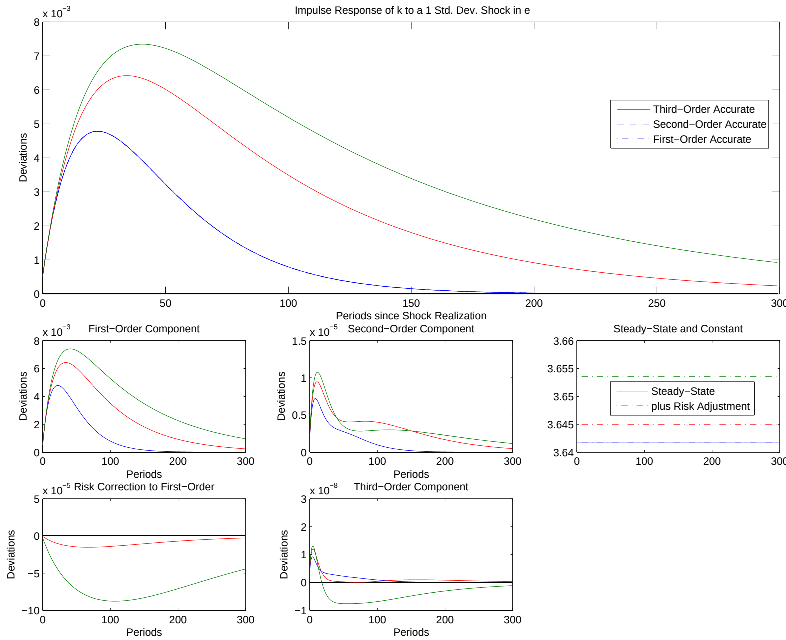


(a) Capital

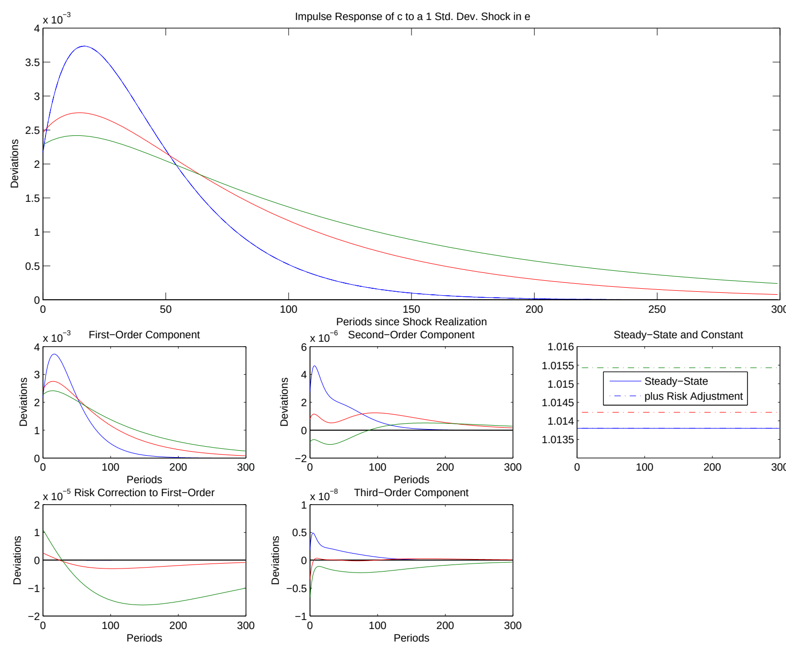


(b) Consumption

Figure 1: Impulse Responses to a Technology Shock, Model of Section 4.2, Variables in Logs

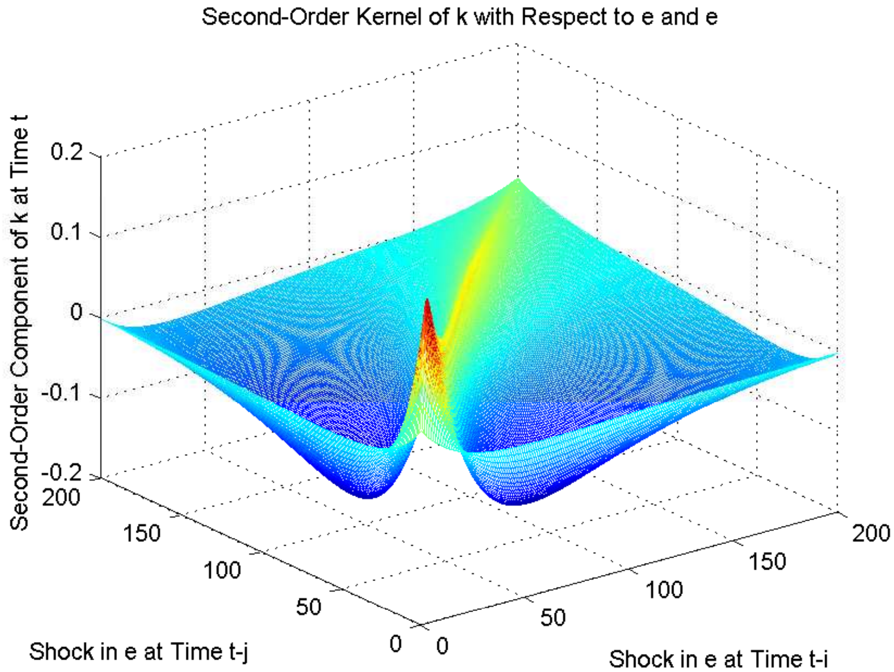


(a) Capital

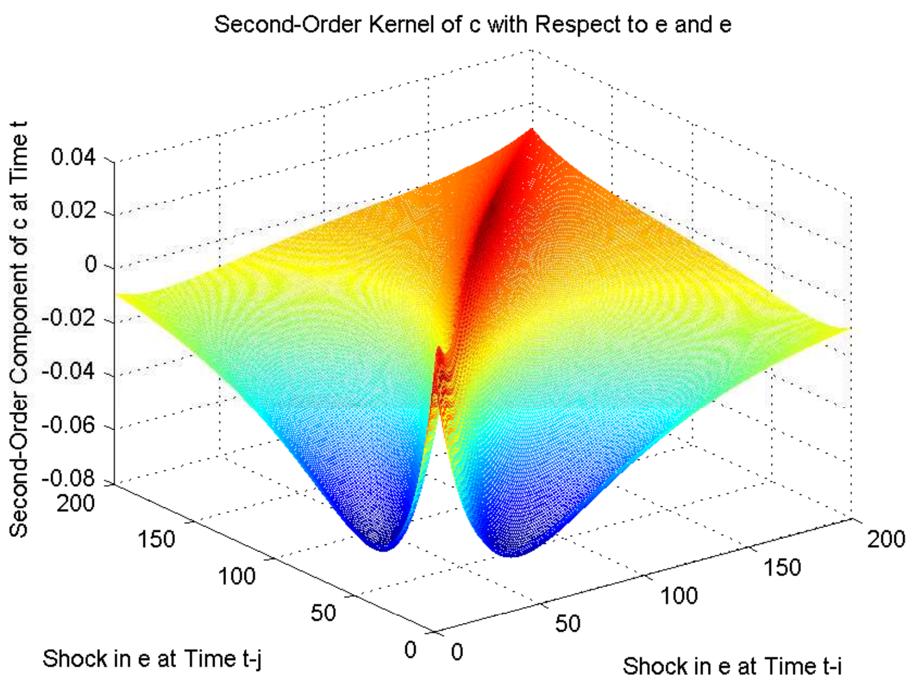


(b) Consumption

Figure 2: Impulse Responses to a Technology Shock, Model of Section 4.2
Blue: $\gamma = 1$, Red $\gamma = 5$, Green $\gamma = 10$

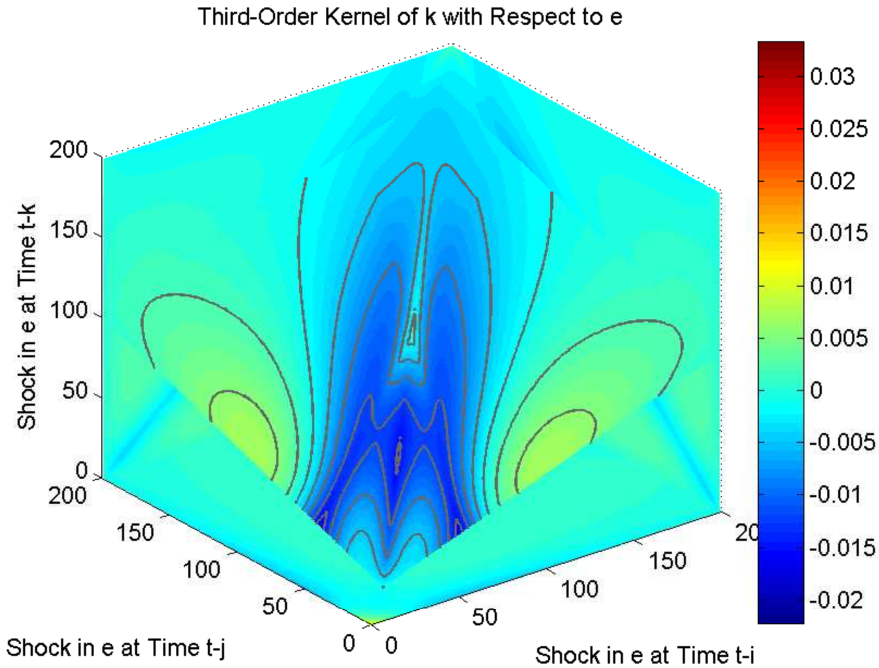


(a) Capital

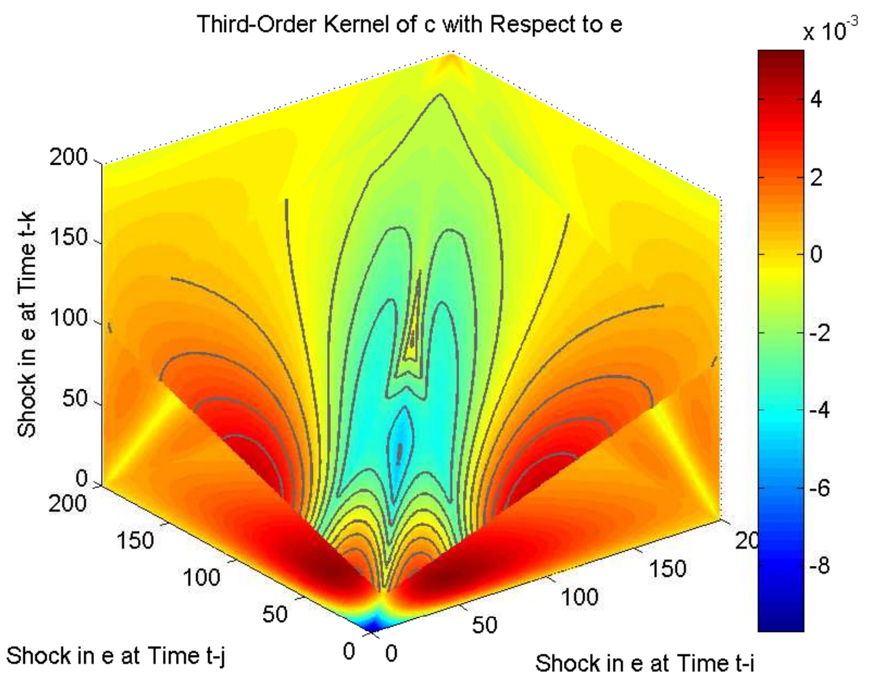


(b) Consumption

Figure 3: Second-Order Kernels, Model of Section 4.2

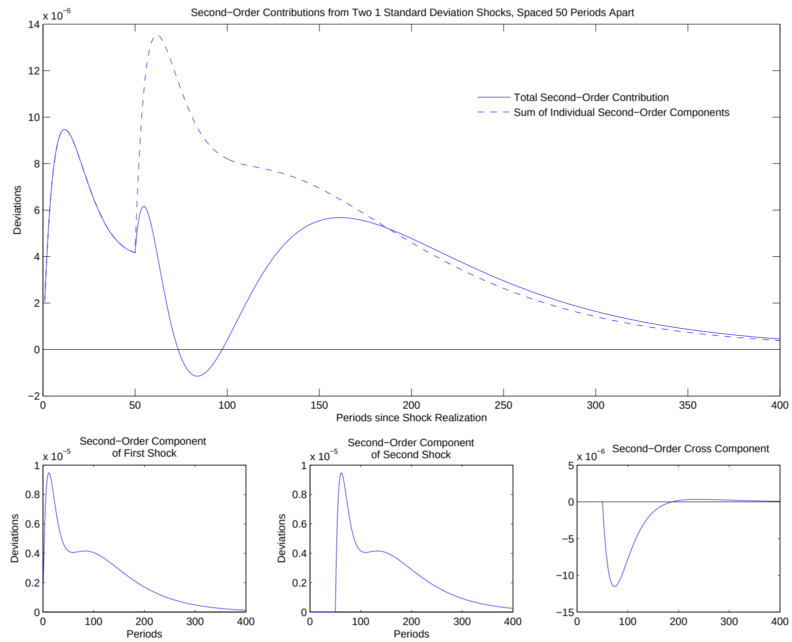


(a) Capital

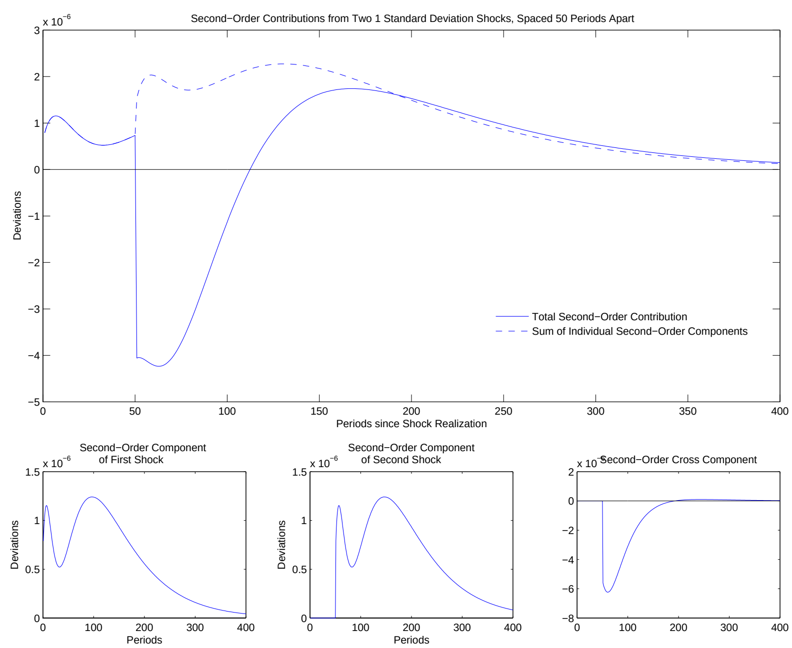


(b) Consumption

Figure 4: Third-Order Kernels, Model of Section 4.2

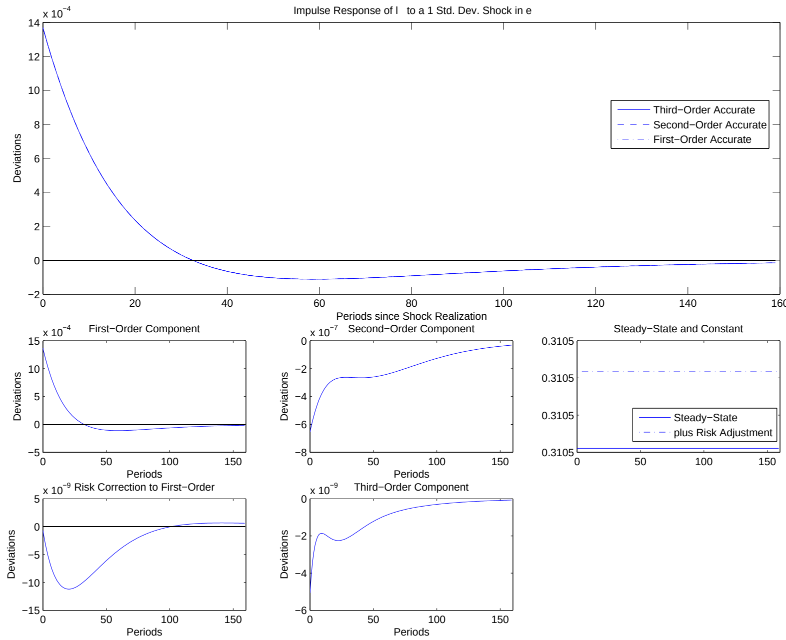


(a) Capital

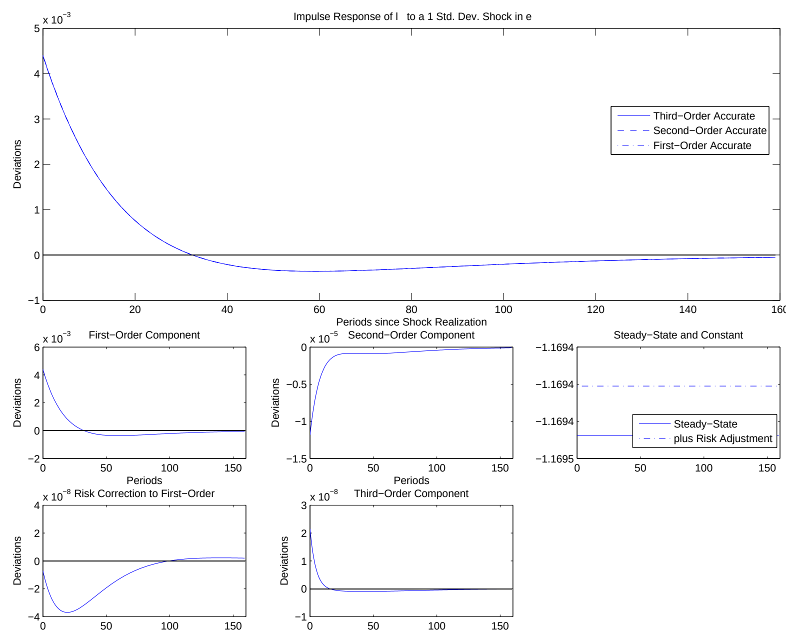


(b) Consumption

Figure 5: Second-Order Contributions to Impulse Responses to a Technology Shock, Model of Section 4.2

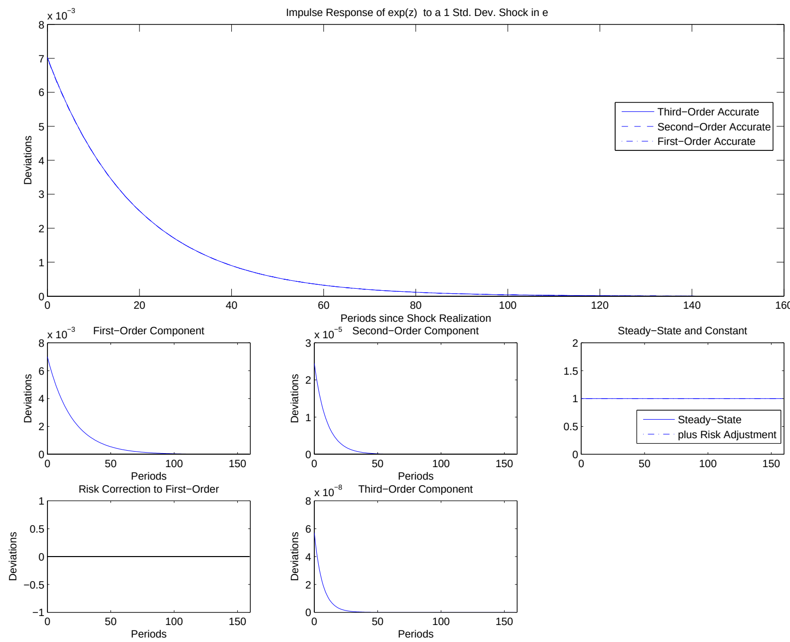


(a) In Levels

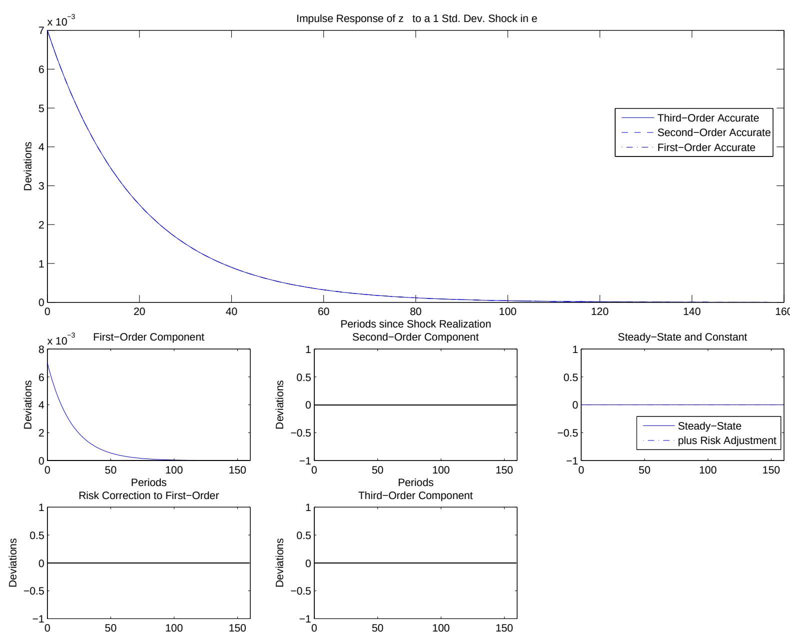


(b) In Logs

Figure 6: Impulse Response of Labor to a Technology Shock



(a) In Levels



(b) In Logs

Figure 7: Impulse Response of e^{z_t} to a Technology Shock

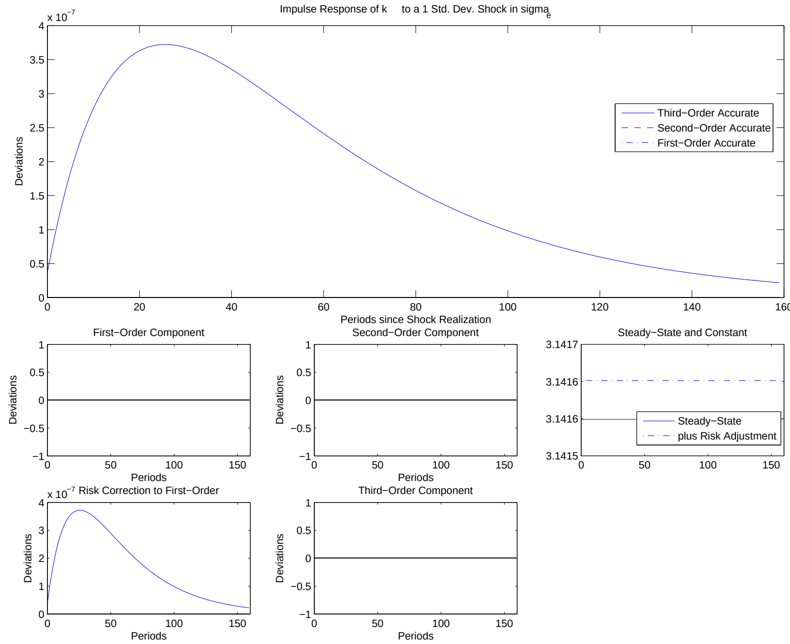


Figure 8: Impulse Response of Capital to a Volatility Shock, in Logs

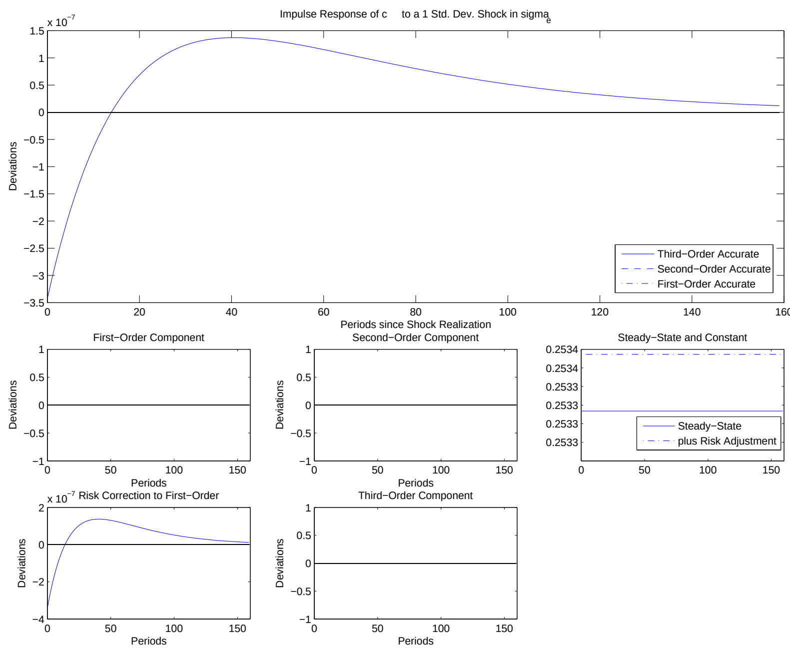


Figure 9: Impulse Response of Consumption to a Volatility Shock, in Logs

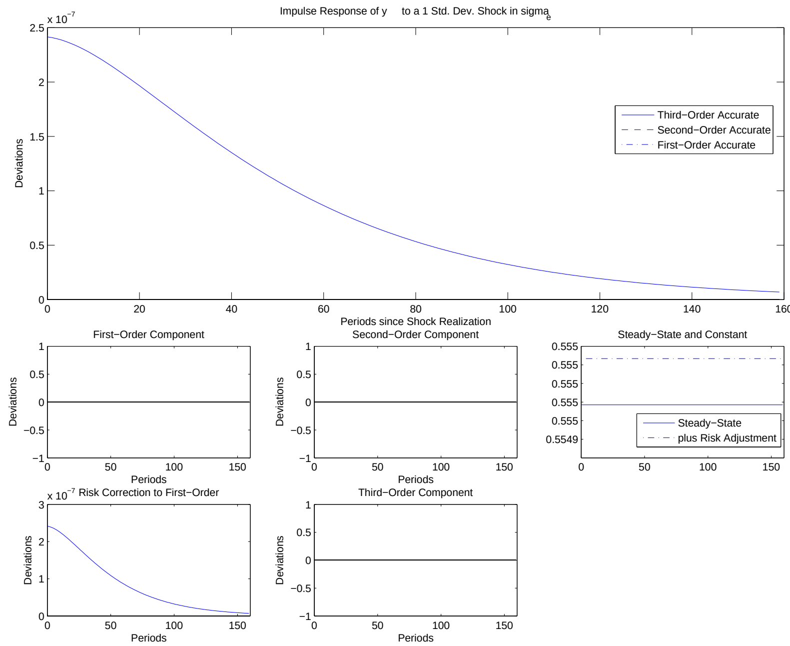


Figure 10: Impulse Response of Production to a Volatility Shock, in Logs

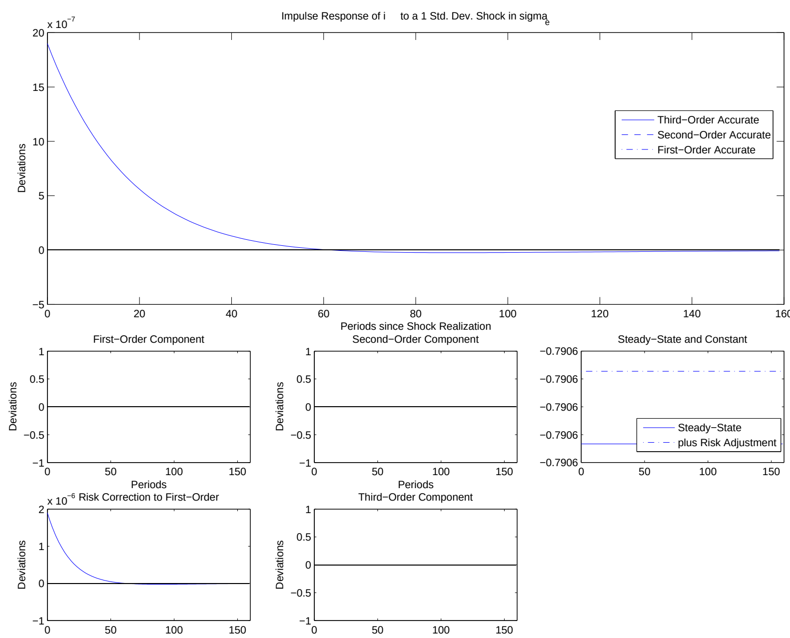


Figure 11: Impulse Response of Investment to a Volatility Shock, in Logs

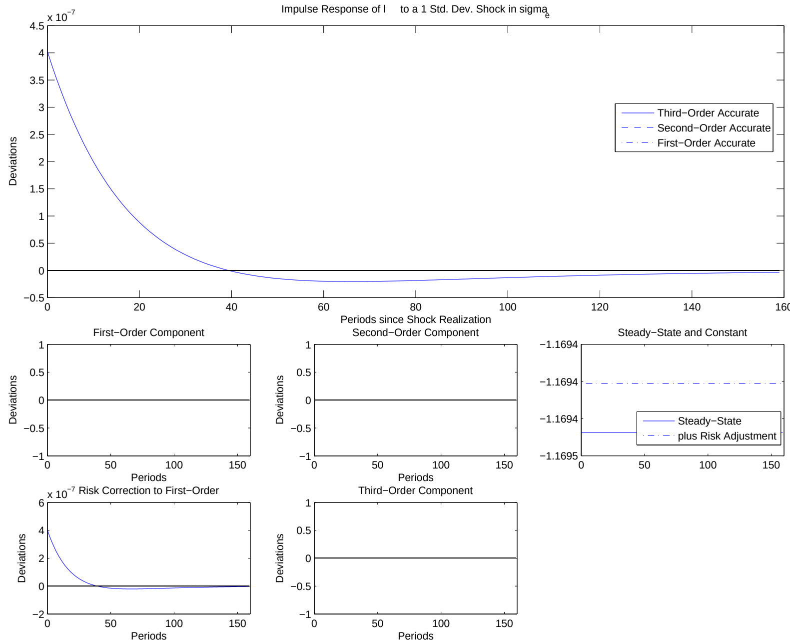


Figure 12: Impulse Response of Labor to a Volatility Shock, in Logs

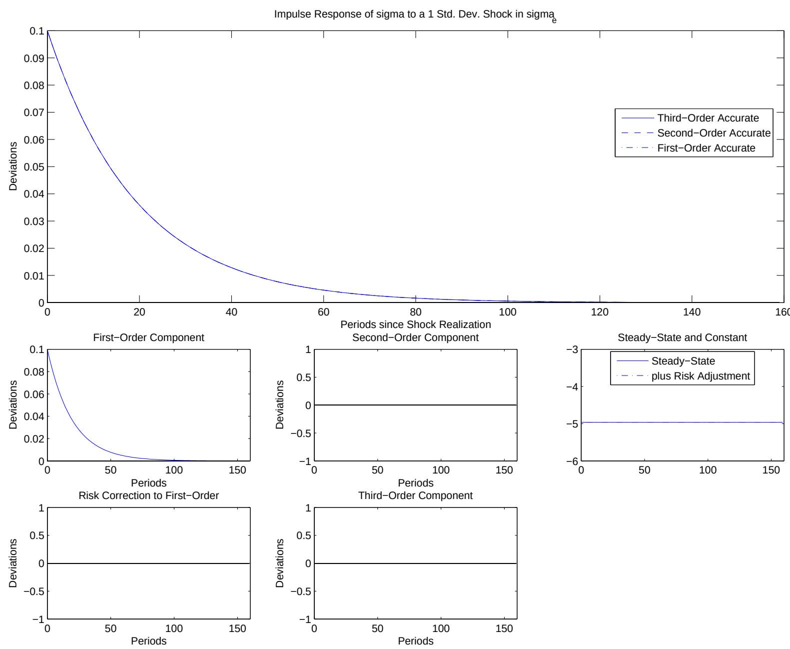


Figure 13: Impulse Response of σ to a Volatility Shock

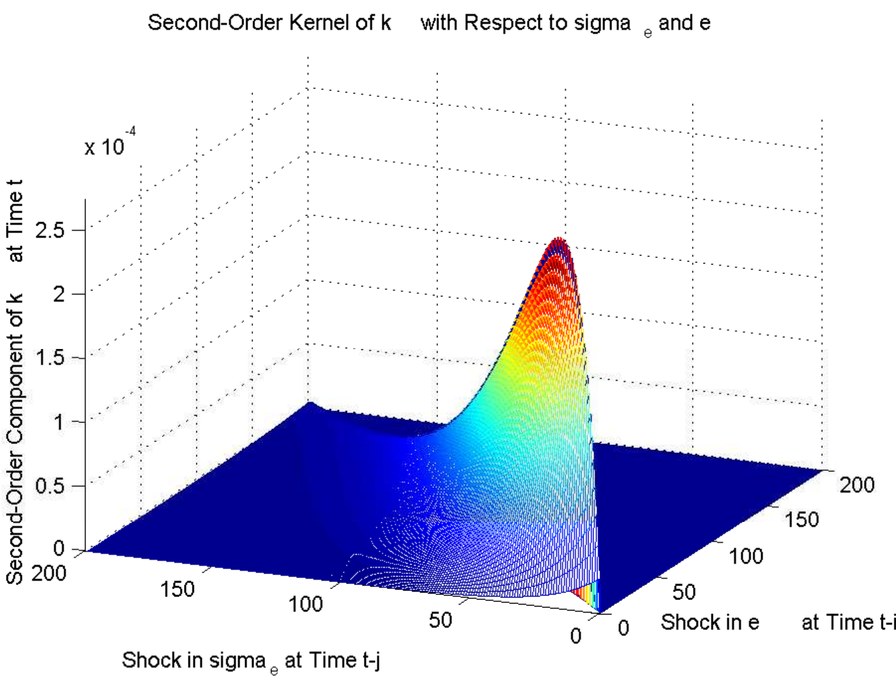
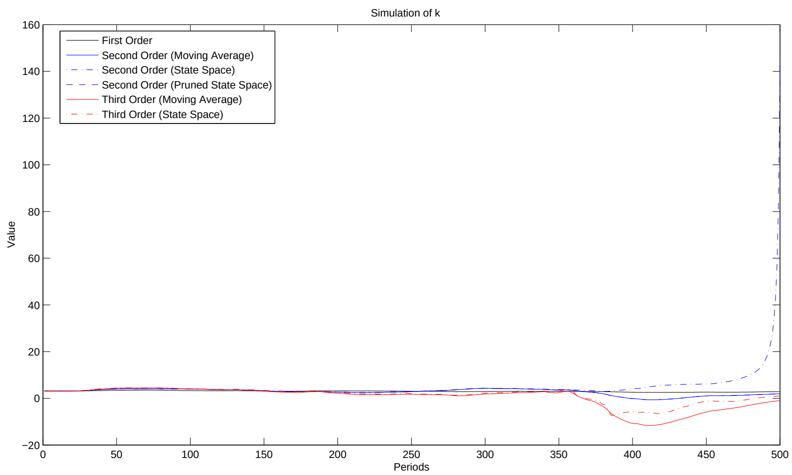
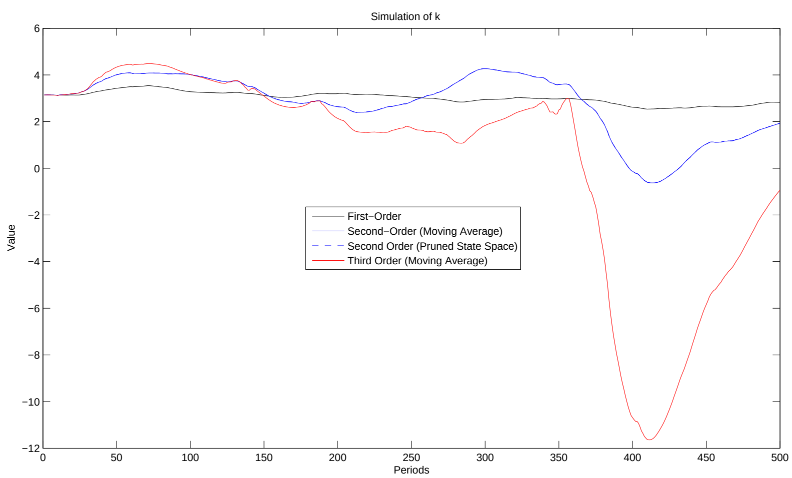


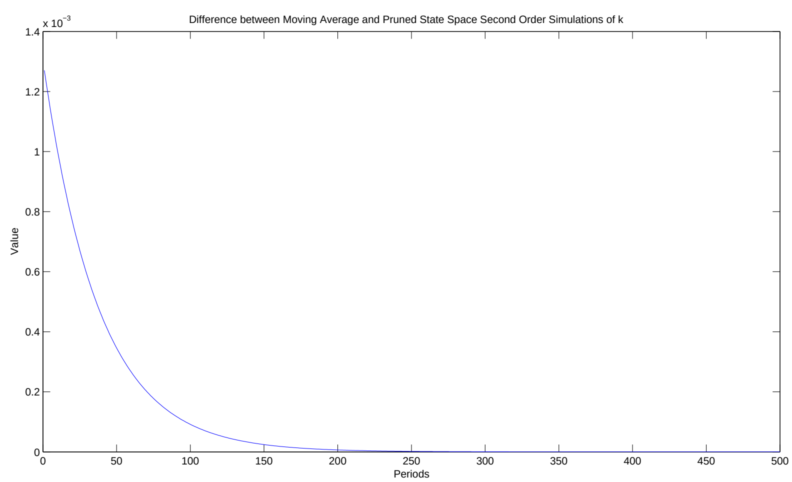
Figure 14: Second Order Cross Kernel of Capital to Volatility and Technology Shocks, in Logs



(a) Including Non-Pruned

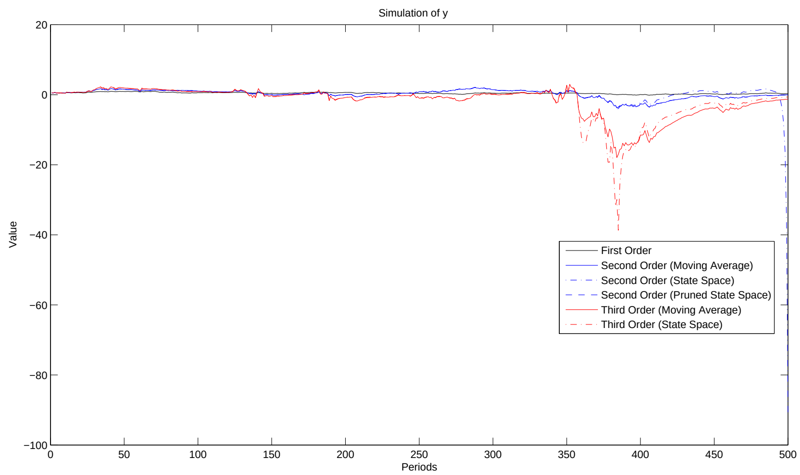


(b) Excluding Non-Pruned



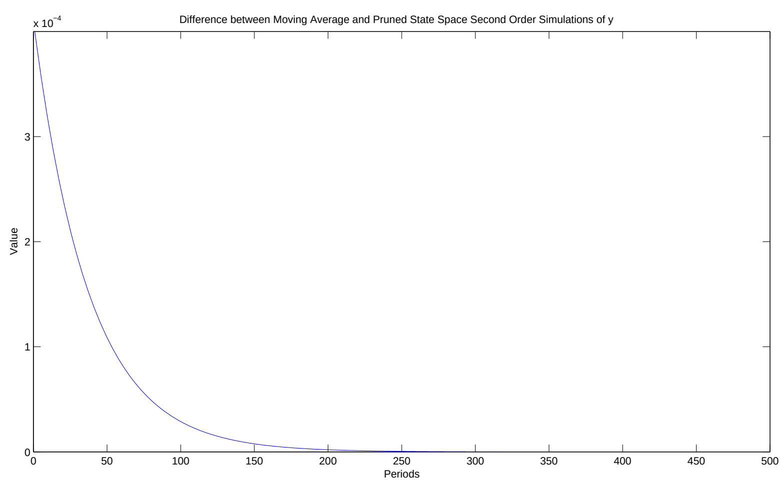
(c) Difference between Moving Average and Pruned

Figure 15: Simulation of Capital, Logarithms



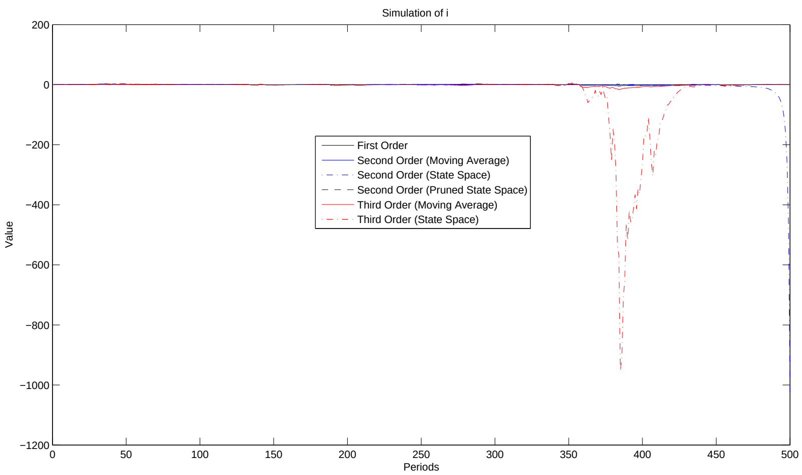
(a) Including Non-Pruned

(b) Excluding Non-Pruned

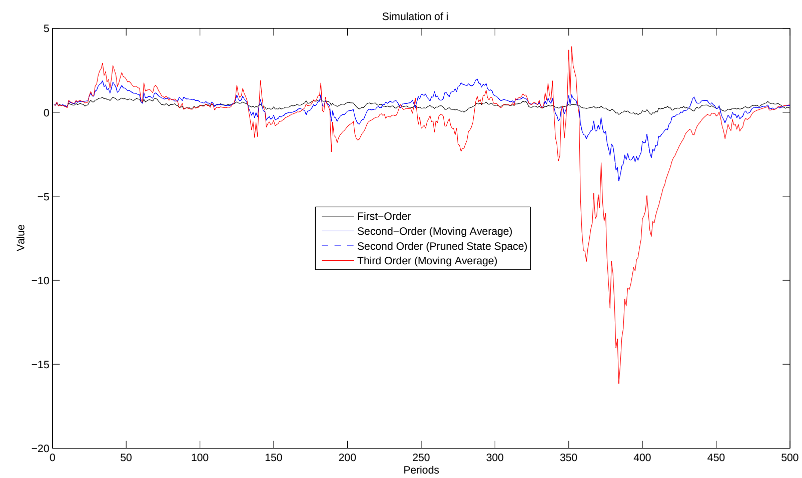


(c) Difference between Moving Average and Pruned

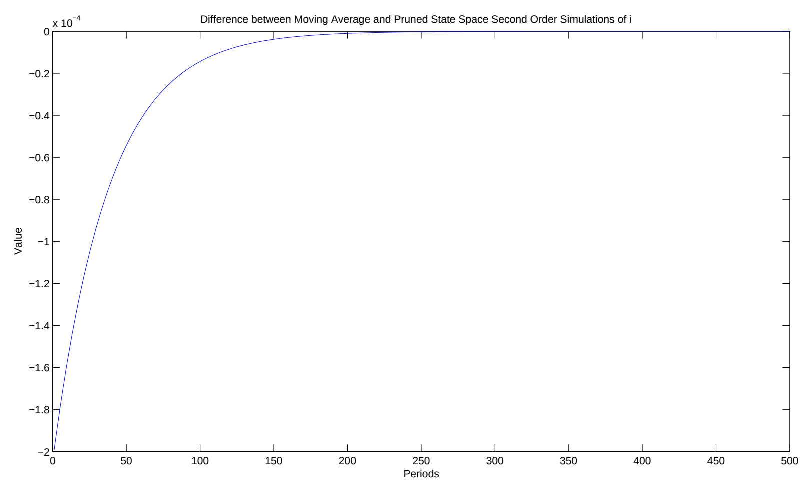
Figure 16: Simulation of Production, Logarithms



(a) Including Non-Pruned

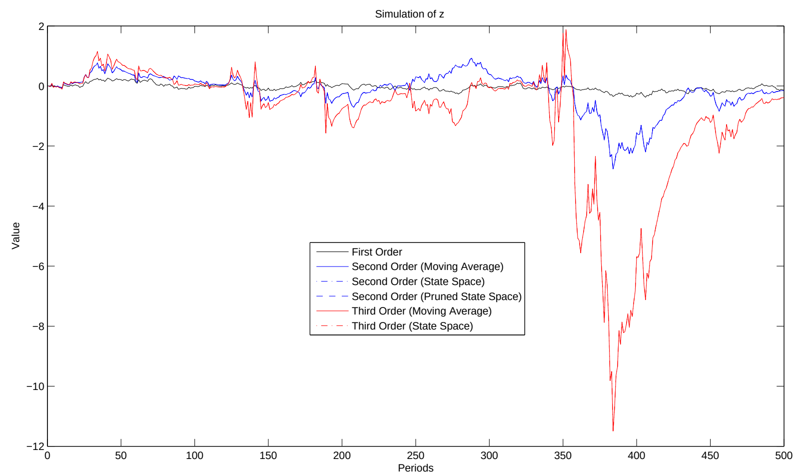


(b) Excluding Non-Pruned

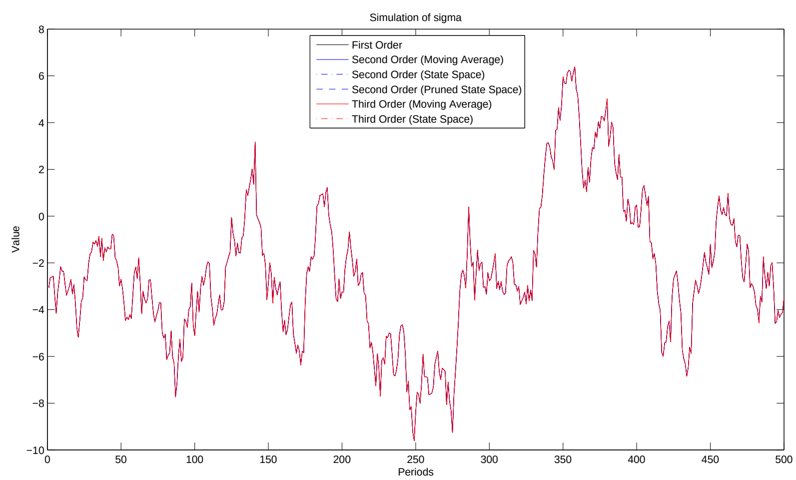


(c) Difference between Moving Average and Pruned

Figure 17: Simulation of Investment, Levels

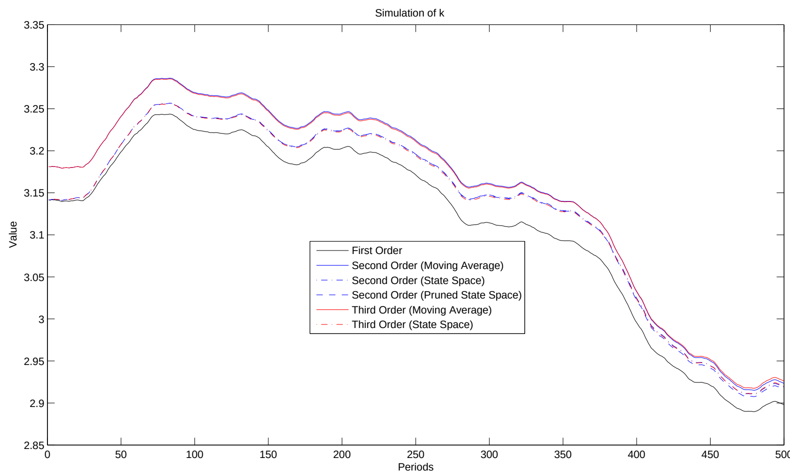


(a) Technology

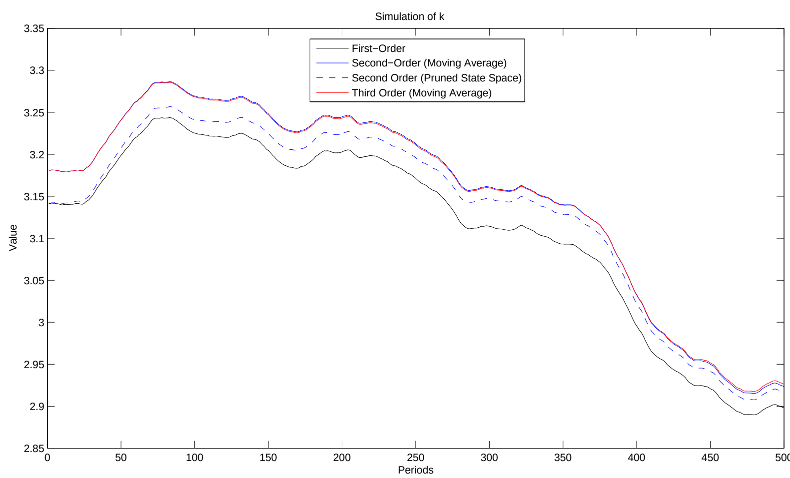


(b) Volatility

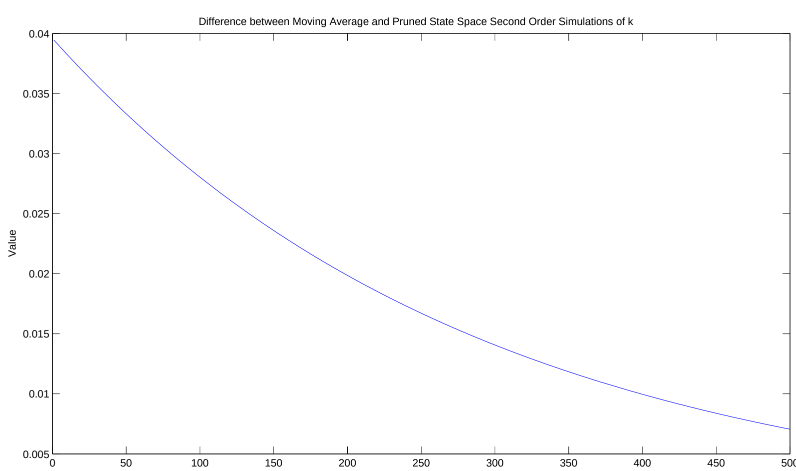
Figure 18: Simulation of Technology and Volatility, Logarithms



(a) Including Non-Pruned



(b) Excluding Non-Pruned



(c) Difference between Moving Average and Pruned

Figure 19: High Risk Aversion and Low Volatility Simulation of Capital, Logarithms

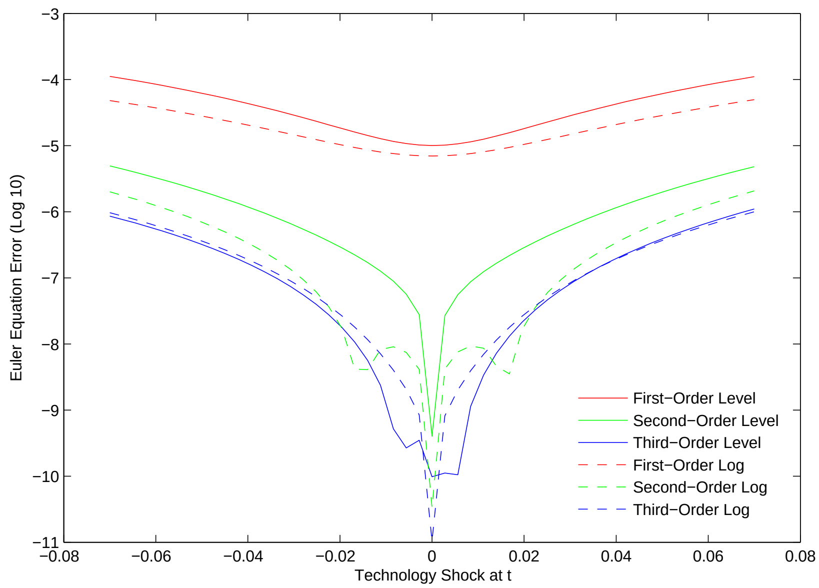
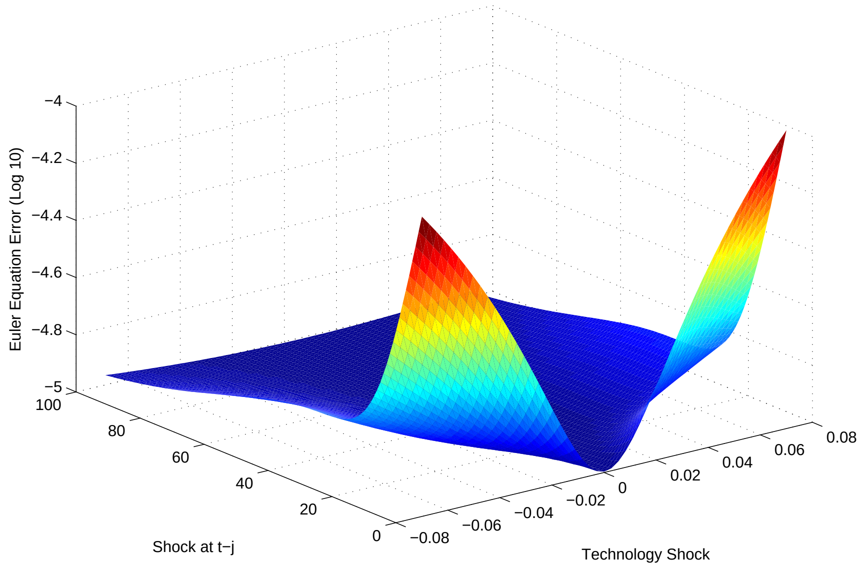
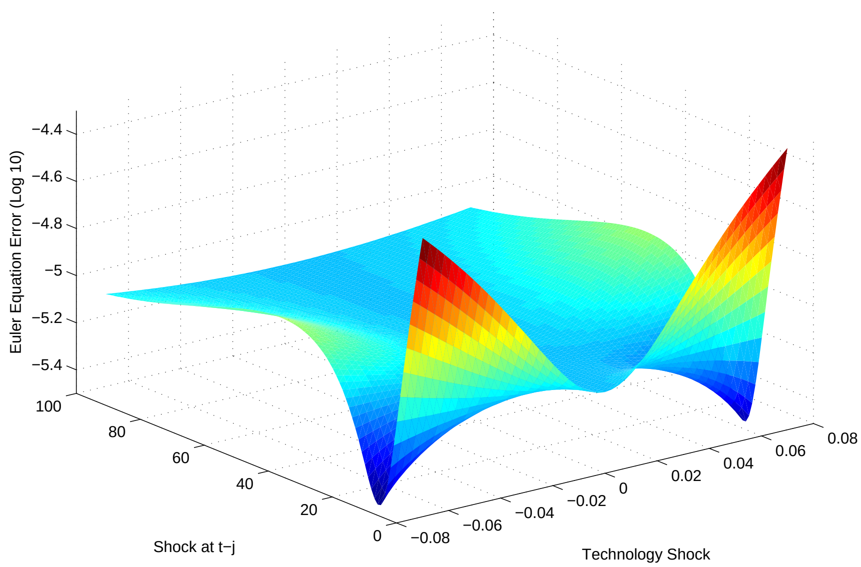


Figure 20: Euler Equation Errors, Shock at Time t, Aruoba, Fernández-Villaverde, and Rubio-Ramírez's (2006) Baseline Case

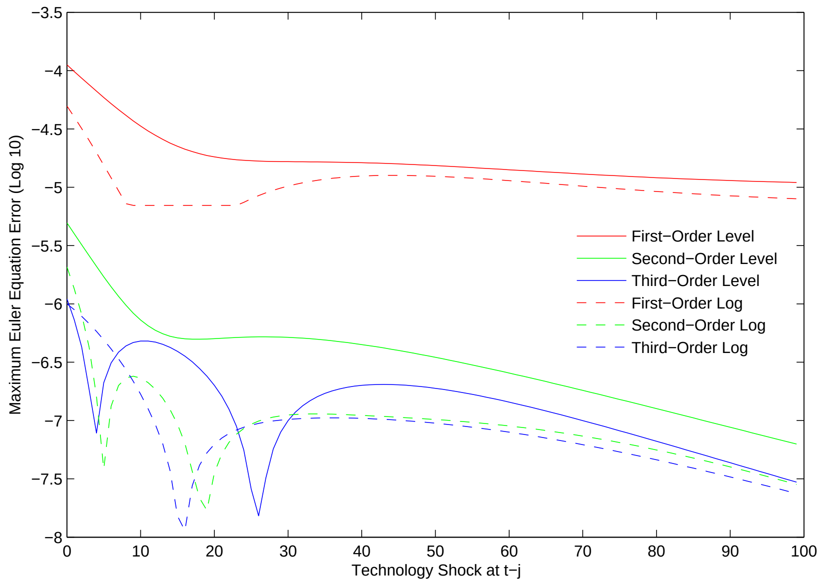


(a) In Levels

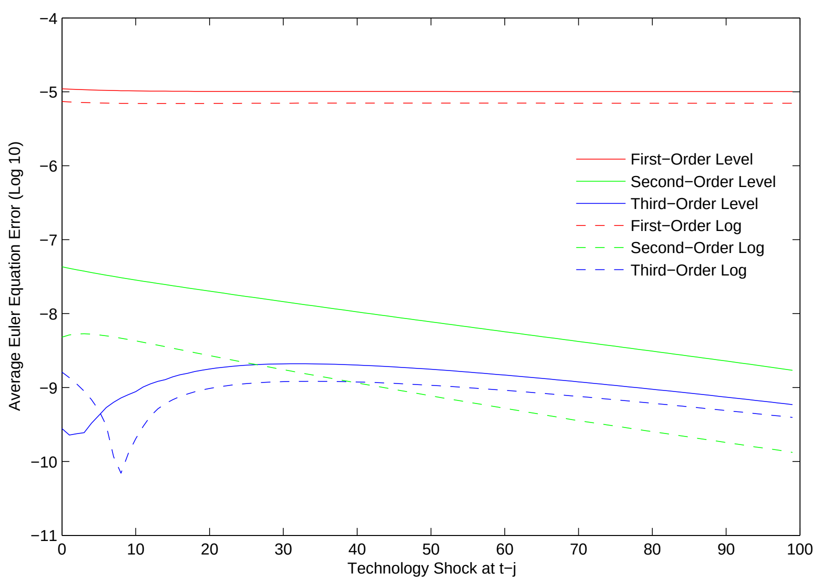


(b) In Logs

Figure 21: Euler Equation Errors, First-Order Approximation, Aruoba, Fernández-Villaverde, and Rubio-Ramírez's (2006) Baseline Case



(a) Maximum Euler Equation Errors



(b) Average Euler Equation Errors

Figure 22: Maximum and Average Euler Equation Errors, Aruoba, Fernández-Villaverde, and Rubio-Ramírez's (2006) Baseline Case