

# Solving DSGE Models with a Nonlinear Moving Average \*

Hong Lan <sup>†</sup>

Alexander Meyer-Gohde<sup>§</sup>

This Version: July 10, 2012

## Abstract

We propose a nonlinear infinite moving average as an alternative to the standard state space policy function for solving nonlinear DSGE models. Perturbation of the nonlinear moving average policy function provides a direct mapping from a history of innovations to endogenous variables, decomposes the contributions from individual orders of uncertainty and nonlinearity, and enables familiar impulse response analysis in nonlinear settings. When the linear approximation is saddle stable and free of unit roots, higher order terms are likewise saddle stable and first order corrections for uncertainty are zero. We derive the third order approximation explicitly and examine the accuracy of the method using Euler equation tests.

*JEL classification:* C61, C63, E17

*Keywords:* Perturbation; nonlinear impulse response; DSGE; solution methods; Volterra series

---

\*We are grateful to Michael Burda and Lutz Weinke, as well as participants of the CEF 2012 and of research seminars at the Bundesbank and 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.

<sup>†</sup>Humboldt-Universität zu Berlin, Institut für Wirtschaftstheorie II, Spandauer Straße 1, 10178 Berlin, Germany; Tel.: +49-30-2093 1466; Email: [lanhong@cms.hu-berlin.de](mailto:lanhong@cms.hu-berlin.de)

<sup>§</sup>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](mailto:alexander.meyer-gohde@wiwi.hu-berlin.de)

# 1 Introduction

Solving models with a higher than first order degree of accuracy is an important challenge for DSGE analysis with the growing interest in nonlinearities. 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 impulse response functions (IRFs). For economists interested in studying the transmission of shocks, our method offers new insight into the propagation mechanism of nonlinear DSGE models.

The nonlinear moving average policy function chooses as its state variable basis the infinite history of past shocks.<sup>1</sup> 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.<sup>2</sup> Our infinite dimensional approach is longstanding in linear models and delivers the same solution as state space methods for linear models.<sup>3</sup> For the nonlinear focus of this paper, however, it provides a different solution. Deriving the direct mapping from shocks to endogenous variables of interest—a Volterra series expansion—facilitates familiar impulse response analysis and makes clear the caveats introduced by nonlinearity. These include history dependence, asymmetries, a breakdown of superposition and scale invariance, as well as the potential for harmonic distortion.<sup>4</sup>

As highlighted by Gomme and Klein (2011) in their second order approximation, deriving per-

---

<sup>1</sup>Kalman’s (1980) “external” or “empirical” approach to system theory in contrast to the “internal” or “state-variable” approach of the state space methods more familiar to DSGE practitioners. See Woodford (1986) for a theoretical foundation of nonlinear DSGE solutions in this space of infinite sequences of innovations.

<sup>2</sup>See Collard and Juillard (2001b), Collard and Juillard (2001a), Jin and Judd (2002), Schmitt-Grohé and Uribe (2004), Anderson, Levin, and Swanson (2006), Lombardo and Sutherland (2007), and Kim, Kim, Schaumburg, and Sims (2008). Recent work of Aruoba, Bocola, and Schorfheide (2011) links their quadratic autoregressive (QAR) time series model within a DSGE context to the Volterra series expansion that we use as our solution basis. Whereas we solve a DSGE model directly with a Volterra series, Aruoba, Bocola, and Schorfheide (2011) compare a reduced form Volterra time series with a state space DSGE policy function. Their focus is on estimation, which is beyond the scope of this paper.

<sup>3</sup>Compare, e.g., the state space representations of Uhlig (1999), Klein (2000), or Sims (2001) with the infinite moving-average representations of Muth (1961), Whiteman (1983) or Taylor (1986).

<sup>4</sup>See Priestly (1988), Koop, Pesaran, and Potter (1996), Potter (2000), and Gourieroux and Jasiak (2005) for detailed discussions from a time series perspective.

turbation solutions with standard linear algebra increases the transparency of the technique and makes coding the method more straightforward. In that vein, we adapt Vetter's (1973) multidimensional calculus to provide a mechanical system of differentiation that maintains standard linear algebraic structures for arbitrarily high orders of approximation. We implement our approach numerically by providing an add on for the popular Dynare package.<sup>5</sup> We then apply our method to the stochastic growth model of Aruoba, Fernández-Villaverde, and Rubio-Ramírez (2006) for comparability and explore the resulting decomposition of the contributing components of the responses of variables to exogenous shocks. We develop Euler equation error methods for our infinite dimensional policy function and confirm that our moving average solution produces approximations with a degree of accuracy comparable to state space solutions of the same order of approximation presented in Aruoba, Fernández-Villaverde, and Rubio-Ramírez (2006).<sup>6</sup>

We make two assumptions on the characteristic equation of the first order (i.e., linear) approximation: it is saddle stable and it is free of unit roots. The first is the standard Blanchard and Kahn (1980) assumption and we show that the resulting stability from the first order is passed on to higher order terms. The second is necessary to ensure the boundedness of corrections to constants and essentially embodies the necessary invertibility of a standard state space policy function to yield our infinite moving average. Additionally, these assumptions enable us to show that the derivatives of the moving average policy function first order in the perturbation parameter are uniquely zero, as shown by Jin and Judd (2002) and Schmitt-Grohé and Uribe (2004) for state space policy function.

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 of our nonlinear infinite moving average policy function explicitly out to the third order. We apply our method to a standard stochastic growth model in section 4, a widely used baseline for numerical

---

<sup>5</sup>See Adjemian, Bastani, Juillard, Mihoubi, Perendia, Ratto, and Villemot (2011).

<sup>6</sup>Aruoba, Fernández-Villaverde, and Rubio-Ramírez (2006) also explore several global methods (projection, value function iteration) and our choice allows comparability to these other methods. Our focus, however, is on the alternative basis from the nonlinear moving average for local (perturbation) methods and our analysis proceeds accordingly.

methods in macroeconomics. In section 5, we develop Euler equation error methods for our infinite dimensional solution form and quantify the accuracy of our method. Section 6 concludes.

## 2 Problem Statement and Solution Form

In this section, we introduce the class of models we analyze and the policy function we propose as a solution. Our class of models is a standard system of (nonlinear) second order expectational difference equations. In contrast with the general practice in the literature, however, the solution will be a policy function that directly maps from realizations of the exogenous variables to the endogenous variables of interest. We will first present the model class followed by the solution form and then conclude with the Taylor/Volterra approximation of the solution and the matrix calculus necessary to follow the derivations in subsequent sections.

### 2.1 Model Class

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

$$(1) \quad 0 = E_t[f(y_{t-1}, y_t, y_{t+1}, u_t)], \text{ where } u_t = \sum_{i=0}^{\infty} N^i \varepsilon_{t-i}$$

$f$  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 ( $neq = ny$ ).  $N$  is the  $(nu \times nu)$  matrix of autoregressive coefficients of  $u_t$ , presented here in moving average form. The eigenvalues of  $N$  are assumed all inside the unit circle so that  $u_t$  admits this infinite moving average representation; and  $\varepsilon_t$  is an  $(ne \times 1)$  vector of exogenous shocks of the same dimension ( $nu = ne$ ). Our software add on forces  $N = 0$  to align with Dynare.<sup>7</sup>

---

<sup>7</sup>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 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).

Additionally,  $\varepsilon_t$  is assumed independently and identically distributed such that  $E(\varepsilon_t) = 0$  and  $E(\varepsilon_t^{\otimes [n]})$  exists and is finite for all  $n$  up to and including the order of approximation to be introduced subsequently.<sup>8</sup>

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  $\{\varepsilon_t, \varepsilon_{t-1}, \dots\}$  — the realizations of the exogenous shocks up to (including)  $t$  — with  $\sigma$ , as they are known with certainty at  $t$ . The perturbation parameter does not enter the problem statement explicitly, but only implicitly through the policy functions, and its role will become clear as we introduce the solution form and its approximation.

## 2.2 Solution Form

Let the policy function take the causal one-sided infinite sequence of shocks as its state vector and, following Anderson, Levin, and Swanson (2006, p. 3), let it be time invariant for all  $t$ , analytic and ergodic.<sup>9</sup> The unknown policy function is then given by

$$(2) \quad y_t = y(\sigma, \varepsilon_t, \varepsilon_{t-1}, \dots)$$

Note that  $\sigma$  enters as a separate argument. As the scale of uncertainty changes, so too will the policy function  $y$  itself change. Time invariance and scaling uncertainty give us

$$(3) \quad y_{t-1} = y^-(\sigma, \varepsilon_{t-1}, \varepsilon_{t-2}, \dots)$$

$$(4) \quad y_{t+1} = y^+(\sigma, \tilde{\varepsilon}_{t+1}, \varepsilon_t, \varepsilon_{t-1}, \dots) \text{ where } \tilde{\varepsilon}_{t+1} \equiv \sigma \varepsilon_{t+1}$$

---

<sup>8</sup>The notation  $\varepsilon_t^{\otimes [n]}$  represents Kronecker powers,  $\varepsilon_t^{\otimes [n]}$  is the  $n$ 'th fold Kronecker product of  $\varepsilon_t$  with itself:  $\underbrace{\varepsilon_t \otimes \varepsilon_t \cdots \otimes \varepsilon_t}_{n \text{ times}}$ . For simulations, of course, more specific decisions regarding the distribution of the exogenous processes will have to be made. Kim, Kim, Schaumburg, and Sims (2008, p. 3402) emphasize that distributional assumptions like these are not entirely local assumptions. Dynare (Adjemian, Bastani, Juillard, Mihoubi, Perendia, Ratto, and Villemot 2011) assumes normality of the underlying shocks.

<sup>9</sup>Analyticity is required for the convergence of asymptotic expansion as the order of approximation becomes infinite and ergodicity rules out explosive and nonfundamental solutions.

The notation,  $y$ ,  $y^-$ , and  $y^+$ , is adopted so that we can keep track of the source (through  $y_t$ ,  $y_{t-1}$ , and  $y_{t+1}$  respectively) of any given partial derivative of the policy function. Due to the assumption of time invariance,  $y$ ,  $y^-$ , and  $y^+$  are the same function differing only in the timing of their arguments. The importance of discriminating among these functions will become clear in the next section. The term  $\sigma \varepsilon_{t+1}$  in (4) is the source of uncertainty, via  $\varepsilon_{t+1}$ , that we are perturbing with  $\sigma$ . The known function  $u$  of the exogenous variable is rewritten similarly as

$$(5) \quad u_t = u(\sigma, \varepsilon_t, \varepsilon_{t-1}, \dots) = \sum_{i=0}^{\infty} N^i \varepsilon_{t-i}$$

For notational ease in derivation, we will define vector  $x_t$ , containing the complete set of variables

$$(6) \quad x_t \equiv [y'_{t-1} \quad y'_t \quad y'_{t+1} \quad u'_t]'$$

$x_t$  is of dimension  $(nx \times 1)$  with  $(nx = 3ny + ne)$ . With the policy function of the form (2), (3) and (4), plus the function of the exogenous variable (5), we can write  $x_t$  as

$$(7) \quad x_t = x(\sigma, \tilde{\varepsilon}_{t+1}, \varepsilon_t, \varepsilon_{t-1}, \dots)$$

Following from the assumptions on  $y$  and  $u$ ,  $x$  is likewise time invariant, analytic and ergodic.

### 2.3 Approximation: Taylor/Volterra Series Approximation

We will approximate the solution, (2), as a Taylor series in the infinite state vector (i.e., a Volterra series) expanded around a nonstochastic steady state,  $\bar{x}$ , which is the solution to the function

$$(8) \quad 0 = f(\bar{y}, \bar{y}, \bar{y}, 0) = f(\bar{x})$$

that is, the function  $f$  in (1) with all shocks, past and present, set to zero, and all uncertainty regarding the future eliminated ( $\sigma = 0$ ). Furthermore

$$(9) \quad \bar{y} = y(0, 0, \dots)$$

represents the solution (2) evaluated at the nonstochastic steady state.

Following general practice in the perturbation literature, we pin down the approximation of the unknown policy function (2) by successively differentiating (1) and solving the resulting systems for the unknown coefficients. The algorithm 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. We adapt the structure of matrix derivatives defined in Vetter (1973) to unfold the hypercube conformable to the Kronecker product, collecting partial derivatives from successive differentiation of  $f$  in two dimensional matrices. This allows us to avoid tensor notation and use standard linear algebra.

A similar approach can be found in Gomme and Klein (2011). They use the matrix derivative structure and associated chain rule of Magnus and Neudecker (2007, ch. 6) to unfold a three dimensional cube of second partial derivatives. The approach does not appear to be easily adaptable to orders of approximation higher than two, as Magnus and Neudecker (2007) do not provide methods that go beyond the second differential. Lombardo and Sutherland (2007) also derive a second order solution without appealing to tensor notation and benefit from their use of the vech operator to eliminate redundant quadratic terms. Our approach, however, provides a mechanical recipe applicable to higher orders, extending the ideas of these existing approaches past the second order.

The formal definition of our matrix derivative structure is in the Appendix. This structure will make the presentation of the solution method more transparent—successive differentiation of  $f$  to the desired order of approximation is a mechanical application of the following theorem

**Theorem 2.1.** *A Multidimensional Calculus*

*For the matrix-valued functions  $F$ ,  $G$ ,  $A$ , and  $H$  and vector-valued functions  $J$  and  $C$  there exists an operator  $\mathcal{D}_x^T$  indicating differentiation with respect to the transpose of the column vector  $x$ . Unless indicated otherwise, all matrices and vectors are understood to be functions of the  $s \times 1$  column vector  $B$  and we leave this dependency implicit.*

1. *Matrix Product Rule:*

$$\mathcal{D}_{B^T} \left\{ \begin{matrix} F & G \\ p \times qq \times u \end{matrix} \right\} = F_B \left( \begin{matrix} I & \otimes & G \\ s \times s \end{matrix} \right) + FG_B$$

2. *Matrix Chain Rule:*

$$\mathcal{D}_{B^T} \left\{ A \begin{pmatrix} C \\ \end{pmatrix} \right\} = A_C \left( C_B \otimes I_{q \times q} \right)$$

3. *Matrix Kronecker Product Rule:*

$$\mathcal{D}_{B^T} \left\{ F \otimes H \right\} = F_B \otimes H + (F \otimes H_B) \left( K_{q,s} \otimes I_{v \times v} \right)$$

where  $K_{q,s}$  is a  $qs \times qs$  commutation matrix (see Magnus and Neudecker (1979)).

4. *Vector Chain Rule:*

$$\mathcal{D}_{B^T} \left\{ J \begin{pmatrix} C \\ \end{pmatrix} \right\} = A_C C_B$$

where  $F_B \equiv \mathcal{D}_{B^T} F$  etc. has been used as abbreviated notation to minimize clutter.

*Proof.* See Appendix. □

By adapting the abbreviated notation from above and writing  $y_{\sigma^{n_{i_1} i_2 \dots i_m}}$  as the partial derivative, evaluated at the nonstochastic steady state, of  $y$  with respect to  $\sigma$  for  $n$  times and with respect to  $\varepsilon_{t-i_1}^T, \varepsilon_{t-i_2}^T, \dots, \varepsilon_{t-i_m}^T$ , we can then write the  $M$ -th order Taylor approximation of the policy function (2) using the following

**Corollary 2.2.** *An  $M$ -th order Taylor Approximation of (2) is written as*

$$(10) \quad y_t = \sum_{m=0}^M \frac{1}{m!} \sum_{i_1=0}^{\infty} \sum_{i_2=0}^{\infty} \dots \sum_{i_m=0}^{\infty} \left[ \sum_{n=0}^{M-m} \frac{1}{n!} y_{\sigma^{n_{i_1} i_2 \dots i_m}} \sigma^n \right] (\varepsilon_{t-i_1} \otimes \varepsilon_{t-i_2} \otimes \dots \otimes \varepsilon_{t-i_m})$$

*Proof.* See Appendix. □

This infinite dimensional Taylor approximation, or Volterra series with kernels  $\left[ \sum_{n=0}^{M-m} \frac{1}{n!} y_{\sigma^{n_{i_1} i_2 \dots i_m}} \sigma^n \right]$ ,<sup>10</sup> directly maps the exogenous innovations to endogenous variables up the  $M$ -th order. The kernels at  $m$  collects all the coefficients associated with the  $m$ 'th fold Kronecker products of exogenous innovations  $i_1, i_2, \dots$  and  $i_m$  periods ago. For a given set of indices,  $i_1, i_2, \dots$  and  $i_m$ , the sum over  $n$  gathering terms in powers of the perturbation parameter  $\sigma$ , corrects the kernel for uncertainty up

<sup>10</sup>See Priestly (1988, pp. 25–26) and Gourieroux and Jasiak (2005) for a representation theorem.

to the  $n$ -th order, thereby enabling a useful classification of the contributions of uncertainty to the model. That is, we can first decompose the Volterra series into kernels associated with the order of approximation in the state space itself—the zeroth kernel being constants, the first order kernel being linear in the product space of the history of innovations, the second being quadratic in the same, etc. Thereafter, we can decompose each of the kernels into successively higher order corrections for uncertainty according to the order in  $\sigma$ — $y_{\sigma^n}$  represents the  $n$ 'th order correction for uncertainty of the zeroth order kernel,  $y_{\sigma^n i_1}$  the  $n$ 'th order correction for uncertainty of the first order kernel,  $y_{\sigma^n i_1 i_2}$  the  $n$ 'th order correction for uncertainty of the second order kernel, and so on.

For a different perspective, observe that moving to a higher order in (10) comprises two changes: (i) adding a higher order kernel and (ii) opening up all existing kernels to a higher order correction for uncertainty.<sup>11</sup> The change in moving from an  $M - 1$ 'th to  $M$ 'th order approximation is

$$(11) \quad \sum_{m=0}^M \frac{1}{M!} \sum_{i_1=0}^{\infty} \sum_{i_2=0}^{\infty} \cdots \sum_{i_m=0}^{\infty} \left[ \frac{1}{(M-m)!} y_{\sigma^{M-m} i_1 i_2 \cdots i_m} \sigma^{M-m} \right] (\boldsymbol{\varepsilon}_{t-i_1} \otimes \boldsymbol{\varepsilon}_{t-i_2} \otimes \cdots \otimes \boldsymbol{\varepsilon}_{t-i_m})$$

The difference can be written compactly despite the two changes, as change (i) is an  $M$ 'th order kernel with a zeroth order correction for uncertainty (for  $m = M$  above,  $y_{\sigma^{M-m} i_1 i_2 \cdots i_m} \sigma^{M-m} = y_{\sigma^0 i_1 i_2 \cdots i_m} \sigma^0 = y_{i_1 i_2 \cdots i_m}$ ). From (ii) comes then additionally a first order correction for uncertainty in the  $M - 1$ 'th order kernel, a second order uncertainty correction for the  $M - 2$ 'th kernel and so on up to the  $M$ 'th order correction for uncertainty in the constant or zeroth order kernel. The uncertainty correction at a given order directly depends on the moments of future shocks at each order and so (ii) can be interpreted as successively opening each kernel up to higher moments in the distribution of future shocks, while (i) maintains the standard Taylor notion of moving to a higher order polynomial (captured by the kernels in our Volterra series).

As the notation in (10) is rather dense, it is instructive to consider the case of  $M = 2$  (the second-order approximation) given by

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

---

<sup>11</sup>We are grateful to Michael Burda for suggesting this interpretation.

Here,  $\bar{y}$ , the policy function evaluated at the nonstochastic steady state, represents the rest point in the absence of uncertainty regarding future shocks. The terms  $\sum_{i=0}^{\infty} y_i \varepsilon_{t-i}$  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. The constant term has two uncertainty corrections,  $y_{\sigma} \sigma$  and  $\frac{1}{2} y_{\sigma^2} \sigma^2$  the first and second order corrections for uncertainty respectively, leading to the second order accurate stochastic steady state. At second order,  $\sum_{i=0}^{\infty} y_{i\sigma} \sigma \varepsilon_{t-i}$  is the first order correction for uncertainty concerning future shocks of the first order response to the history of shocks. The first order corrections for uncertainty will turn out to be zero in this case, a familiar result from state space analyses.<sup>12</sup> For the case of  $M = 2$ , the task at hand is to pin down numerical values for  $\bar{y}$ ,  $y_i$ ,  $y_{\sigma}$ ,  $y_{ji}$ ,  $y_{i\sigma}$ , and  $y_{\sigma^2}$  using the information in (1). In the next section, we provide explicit derivations to third order, which is novel in the literature.<sup>13</sup>

### 3 Numerical Solution of the Perturbation Approximation

In this section, we lay out the method for solving for the coefficients of the approximated solution. Solving for the first order terms is primarily an application of methods well known in the literature. Similarly to existing state space methods, solving for higher order terms operates successively on terms from lower orders with linear methods. In contrast to state space methods, the system of equations for the coefficients at all orders of approximation is a system of difference equations with identical homogenous components, enabling the stability from the first order to be passed on to higher orders. Terms linear in the perturbation parameter are zero, as is the case with the state space policy function. The moving average solution function, however, requires us to rule out unit roots in the first order approximation along with the standard saddle point assumptions to ensure the boundedness of uncertainty corrections to constants.

<sup>12</sup>See Jin and Judd (2002), Schmitt-Grohé and Uribe (2004), and Kim, Kim, Schaumburg, and Sims (2008).

<sup>13</sup>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 one third order term occupying two pages highlights the advantage of our notation.

The method can be outlined as follows.<sup>14</sup> Inserting the policy functions for  $y_{t-1}$ ,  $y_t$ , and  $y_{t+1}$ —equations, (3), (2), and (4) respectively—along with the analogous representation, (5), for the exogenous driving force  $u_t$  into the model (1) yields

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

a function with arguments  $\sigma, \varepsilon_t, \varepsilon_{t-1}, \dots$ . At each order of approximation, we take the collection of derivatives of  $f$  from the previous order (for the first order, we 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.,  $\sigma, \varepsilon_t, \varepsilon_{t-1}, \dots$ )
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  $y$ .

The partial derivatives of  $y$ , obtained in step (4) at each order, constitute the unknown coefficients of the Taylor/Volterra approximation of the policy function  $y$ . They are numeric and used again in step (2) of the next higher order. This introduces the potential for the compounding of numerical errors as we move to higher orders as highlighted by Anderson, Levin, and Swanson (2006). The set of derivatives of  $f$  obtained in step (1), however, are symbolic at each order, limiting the source for potential error compounding to the partial derivatives of the policy function.

### 3.1 First Order Approximation

We are seeking the first order approximation of the policy function (2), evaluated at the nonstochastic steady state ( $\bar{x}$ ), of the form

$$(14) \quad y_t = \bar{y} + y_\sigma \sigma + \sum_{i=0}^{\infty} y_i \varepsilon_{t-i}, \quad i = 0, 1, 2, \dots$$

---

<sup>14</sup>See Anderson, Levin, and Swanson (2006, p. 9) for a similar outline in their state space context.

The task at hand is to pin down the partial derivatives,  $y_\sigma$  and  $y_i$ . 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 the original 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  $\sigma$ , as the problem at higher orders of approximation will inherit a similar structure.

To determine  $y_i$ , we differentiate  $f$  in (13) with respect to some  $\varepsilon_{t-i}$

$$(15) \quad \mathcal{D}_{\varepsilon_{t-i}^T} f = f_x x_i$$

Evaluating this at the nonstochastic steady state ( $\bar{x}$ ) and setting its expectation to zero yields

$$(16) \quad E_t(\mathcal{D}_{\varepsilon_{t-i}^T} f) \Big|_{\bar{x}} = f_{y^-} y_{i-1} + f_{y^0} y_i + f_{y^+} y_{i+1} + f_u u_i = 0$$

for  $i = 0, 1, \dots$ , with  $y_{-1} = 0$

a second order linear deterministic difference equation in the matrices  $y_i$  — the derivatives of the vector valued  $y$  function with respect to its  $i - 1$ 'th  $\varepsilon$  element. That is,  $y_i$  contains the linear moving average coefficients of  $y_t$  with respect to the elements of  $\varepsilon_{t-i}$ . Equation (16) 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 (16).

**Assumption 3.1.** *Saddle stability*

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

**Assumption 3.2.** *No unit roots*

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

The first assumption is standard, fulfilling the Blanchard and Kahn (1980) condition. The second has been found in other analyses, e.g., Klein (2000), and here ensures the solvability of terms homogenous in  $\sigma$  — i.e., uncertainty corrections to the constant. Intuitively from the state space perspective, unit roots must be ruled out to allow the state space solution to be inverted, yielding

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.

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.,  $u_t = N^t$ ),<sup>15</sup> delivering the unique stable solution to (16)

$$(17) \quad y_t = \alpha y_{t-1} + \beta_1 u_t, \text{ with } y_{-1} = 0$$

a convergent recursion from which we can recover the linear moving-average terms or  $y_t$ 's.<sup>16</sup>

To determine  $y_\sigma$ , we differentiate  $f$  in (13) with respect to  $\sigma$

$$(18) \quad \mathcal{D}_\sigma f = f_x \mathcal{D}_\sigma x$$

$$\text{where } \mathcal{D}_\sigma x = x_\sigma + x_{\bar{\varepsilon}} \varepsilon_{t+1}$$

Evaluating this at  $\bar{x}$  and setting its expectation to zero yields

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

as  $E_t(\varepsilon_{t+1}) = 0$ . From assumption (3.2), it follows that

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

and hence

$$(21) \quad y_\sigma = 0$$

The first order correction of the constant for uncertainty is zero, analogous to the result of Jin and Judd (2002) and Schmitt-Grohé and Uribe (2004). This result carries over to our moving average by ruling out unit roots to ensure the invertibility of the state space representation. The result itself reflects the rather obvious fact that opening the expansion to a moment of the future distribution of shocks will change nothing if this moment ( $E_t[\varepsilon_{t+1}]$ ) is exactly zero.

---

<sup>15</sup>Alternatively, one can 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.

<sup>16</sup>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.

Gathering the results of this section, the first order approximation of the policy function (2), which can be thought of as an extension of Muth (1961), Taylor (1986), and others, reduces to

$$(22) \quad y_t = \bar{y} + \sum_{i=0}^{\infty} y_i \varepsilon_{t-i}, \quad i = 0, 1, 2, \dots$$

The certainty equivalence of the first order solution is reflected by the independence of (22) from  $\sigma$ .

### 3.2 Second Order Approximation

Taking the first order results as given, we now move on to the second order approximation of the policy function (2) evaluated at the nonstochastic steady state,  $(\bar{x})$ , of the form

$$(23) \quad y_t = \bar{y} + \frac{1}{2} y_{\sigma^2} \sigma^2 + \sum_{i=0}^{\infty} (y_i + y_{\sigma, i} \sigma) \varepsilon_{t-i} + \frac{1}{2} \sum_{j=0}^{\infty} \sum_{i=0}^{\infty} y_{j, i} (\varepsilon_{t-j} \otimes \varepsilon_{t-i})$$

The task is to pin down the three second order derivatives of the  $y$  function,  $y_{j, i}$ ,  $y_{\sigma^2}$  and  $y_{\sigma, i}$ . The equations governing  $y_{j, i}$  and  $y_{\sigma, i}$  will be difference equations with homogenous components identical to those in (16), with the equation in  $y_{\sigma, i}$  being homogenous in accordance with Schmitt-Grohé and Uribe (2004) and others. The no-unit-root assumption will be crucial again in solving for the term  $y_{\sigma^2}$ , preventing this constant correction for uncertainty induced by the potential for future shocks from becoming arbitrarily large.

We first differentiate (15) with respect to some  $\varepsilon_{t-j}$ , delivering  $y_{j, i}$ , the second derivatives of the  $y$  function with respect to all pairs of  $\varepsilon_{t-i}$  and  $\varepsilon_{t-j}$ . 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

$$(24) \quad \mathcal{D}_{\varepsilon_{t-j} \varepsilon_{t-i}}^2 f = f_{x^2} (x_j \otimes x_i) + f_x x_{j, i}$$

Evaluating at the nonstochastic steady state and setting its expectation to zero

$$(25) \quad E_t(\mathcal{D}_{\varepsilon_{t-j} \varepsilon_{t-i}}^2 f) \Big|_{\bar{x}} = f_y y_{j-1, i-1} + f_y y_{j, i} + f_{y+y_{j+1, i+1}} + f_{x^2} (x_j \otimes x_i) = 0$$

for  $j, i = 0, 1, \dots$ , with  $y_{j, i} = 0$ , for  $j, i < 0$

a second order linear deterministic difference equation in  $y_{j, i}$ . The coefficients on the homogeneous components of the forgoing and (16) are identical. The inhomogeneous components have a first order Markov representation (see the shifting and transition matrices defined in the Appendix) in the

Kronecker product of the first order coefficients.<sup>17</sup> The resulting expression is

$$(26) \quad f_y y_{j-1,i-1} + f_y y_{j,i} + f_y y_{j+1,i+1} + f_{x^2}(\gamma_1 \otimes \gamma_1)(S_j \otimes S_i) = 0$$

for  $j, i = 0, 1, \dots$ , with  $y_{j,i} = 0$ , for  $j, i < 0$

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

$$(27) \quad y_{j,i} = \alpha y_{j-1,i-1} + \beta_2(S_j \otimes S_i), \text{ with } y_{j,i} = 0, \forall j, i < 0$$

Note that  $\alpha$  in this solution is known. It is the same uniquely stable  $\alpha$  as in the first order solution (17) due to the fact that the system (25) and (16) have identical homogeneous components. To determine  $\beta_2$ , we substitute (27) in (25), using the shifting matrices and matching coefficients

$$(28) \quad (f_y + f_y \alpha) \beta_2 + f_y \beta_2 (\delta_1 \otimes \delta_1) = -f_{x^2}(\gamma_1 \otimes \gamma_1)$$

This is a type of Sylvester equation, the solution of which is presented in detail by Kamenik (2005).

Next we pin down  $y_{\sigma,i}$ , the second derivatives of the  $y$  function with respect to  $\varepsilon_{t-i}$  and  $\sigma$  sequentially, by differentiating (15) with respect to  $\sigma$ . The resulting linear system is

$$(29) \quad \mathcal{D}_{\sigma \varepsilon_{t-i}}^2 f = f_{x^2}(\mathcal{D}_{\sigma} x \otimes x_i) + f_x \mathcal{D}_{\sigma} x_i$$

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

Note that the additional potential derivative,  $\mathcal{D}_{\varepsilon_{t-i} \sigma}^2 f$ , is simply equal to the derivative in the text,

$\mathcal{D}_{\sigma \varepsilon_{t-i}}^2 f$ .<sup>18</sup> Evaluating (29) at  $\bar{x}$ , taking expectations, noting that  $y_{\sigma} = 0$ , and setting the resulting

---

<sup>17</sup>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.

<sup>18</sup> Although the derivative operator  $\mathcal{D}$  works on Kronecker products (i.e.  $\mathcal{D}_{\sigma \varepsilon_{t-i}}^2 = \mathcal{D}_{\sigma \otimes \varepsilon_{t-i}}^2 f$ ) and although the Kronecker product is not generally commutative,  $\sigma$  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  $\mathcal{D}_{\sigma \varepsilon_{t-i}}^2 f$ , for example, and insert the identity matrix:  $f_{x^2} I_{nx^2}(\mathcal{D}_{\sigma} x \otimes x_i)$ . This can be rewritten as  $f_{x^2} K_{nx,nx} K_{nx,nx}(\mathcal{D}_{\sigma} x \otimes x_i)$ . 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,  $K_{nx,nx}(\mathcal{D}_{\sigma} x \otimes x_i) = x_i \otimes \mathcal{D}_{\sigma} x$ . Now  $f_{x^2} = \mathcal{D}_{x^T \otimes x^T}^2 f$  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:  $f_{x^2} = \mathcal{D}_{x^T \otimes x^T}^2 f K_{nx,nx} = \mathcal{D}_{x^T \otimes x^T}^2 f = f_{x^2}$ . Combining the two yields  $f_{x^2}(\mathcal{D}_{\sigma} x \otimes x_i) = f_{x^2}(x_i \otimes \mathcal{D}_{\sigma} x)$ . Proceeding likewise with the second term in  $\mathcal{D}_{\sigma \varepsilon_{t-i}}^2 f$  completes the argument. Accordingly for higher-order derivatives, the order in which derivatives with respect to  $\sigma$  appear is inconsequential as it is a scalar and we choose to have the  $\sigma$ 's appear first.

expression to zero yields

$$(30) \quad E_t(\mathcal{D}_{\sigma \varepsilon_{t-i}}^2 f) \Big|_{\bar{x}} = f_{y^-} y_{\sigma, i-1} + f_y y_{\sigma, i} + f_{y^+} y_{\sigma, i+1} = 0$$

for  $i = 0, 1, \dots$ , with  $y_{\sigma, -1} = 0$

The unique stable solution takes the form

$$(31) \quad y_{\sigma, i} = \alpha y_{\sigma, i-1}, \text{ for } i = 0, 1, \dots, \text{ with } y_{\sigma, -1} = 0$$

as the system at hand is identical to the homogenous component of the first order system (16).

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

$$(32) \quad y_{\sigma, i} = 0, \text{ for } i = 0, 1, \dots$$

Again, we confirm Schmitt-Grohé and Uribe's (2004) and Jin and Judd's (2002) result that terms with a first order uncertainty correction are zero.

Finally, to determine  $y_{\sigma^2}$ , the second derivative of the  $y$  function with respect to  $\sigma$ , we differentiate (18) with respect to  $\sigma$ , the resulting linear system is

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

where  $\mathcal{D}_{\sigma^2}^2 x = x_{\sigma^2} + 2x_{\sigma, \tilde{\varepsilon}} \tilde{\varepsilon}_{t+1} + x_{\tilde{\varepsilon}^2} (\varepsilon_{t+1} \otimes \varepsilon_{t+1})$

Evaluating this at  $\bar{x}$  and setting its expectation to zero yields

$$(34) \quad E_t(\mathcal{D}_{\sigma^2}^2 f) \Big|_{\bar{x}} = [f_{y^+} y_{0^2} + f_{y^+2} (y_0 \otimes y_0)] E_t(\varepsilon_{t+1} \otimes \varepsilon_{t+1}) + (f_{y^-} + f_y + f_{y^+}) y_{\sigma^2} = 0$$

therefore we can recover  $y_{\sigma^2}$  by

$$(35) \quad y_{\sigma^2} = -(f_{y^-} + f_y + f_{y^+})^{-1} [f_{y^+} y_{0^2} + f_{y^+2} (y_0 \otimes y_0)] E_t(\varepsilon_{t+1} \otimes \varepsilon_{t+1})$$

By assumption, the second moment of the exogeneous variable,  $E_t(\varepsilon_{t+1} \otimes \varepsilon_{t+1})$ , is given.

As the model approaches a unit root from below, the effect of uncertainty on the constant 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.

Gathering the results of this section, the second order approximation of the policy function (2) takes the form

$$(36) \quad y_t = \bar{y} + \frac{1}{2}y_{\sigma^2}\sigma^2 + \sum_{i=0}^{\infty} y_i \varepsilon_{t-i} + \frac{1}{2} \sum_{j=0}^{\infty} \sum_{i=0}^{\infty} y_{j,i} (\varepsilon_{t-j} \otimes \varepsilon_{t-i})$$

In contrast to the first order approximation, (36) does depend on  $\sigma$ , with the term  $\frac{1}{2}y_{\sigma^2}$  correcting the nonstochastic steady state for uncertainty regarding future shocks. As  $\sigma$  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  $\sigma$  to one in (36) gives the second order approximation

$$(37) \quad y_t = \bar{y} + \frac{1}{2}y_{\sigma^2} + \sum_{i=0}^{\infty} y_i \varepsilon_{t-i} + \frac{1}{2} \sum_{j=0}^{\infty} \sum_{i=0}^{\infty} y_{j,i} (\varepsilon_{t-j} \otimes \varepsilon_{t-i})$$

### 3.3 Third Order and Higher Approximations

Given the results from lower orders, including that terms linear in the perturbation parameter are zero, the third order approximation of the  $y$  function we are seeking takes the form

$$(38) \quad y_t = \bar{y} + \frac{1}{2}y_{\sigma^2}\sigma^2 + \frac{1}{6}y_{\sigma^3}\sigma^3 + \sum_{i=0}^{\infty} \left( y_i + \frac{1}{2}y_{\sigma^2,i}\sigma^2 \right) \varepsilon_{t-i} + \frac{1}{2} \sum_{j=0}^{\infty} \sum_{i=0}^{\infty} (y_{j,i} + y_{\sigma,j,i}\sigma) (\varepsilon_{t-j} \otimes \varepsilon_{t-i}) \\ + \frac{1}{6} \sum_{k=0}^{\infty} \sum_{j=0}^{\infty} \sum_{i=0}^{\infty} y_{k,j,i} (\varepsilon_{t-k} \otimes \varepsilon_{t-j} \otimes \varepsilon_{t-i})$$

The task at hand is to pin down some third derivatives of the  $y$  function, including  $y_{k,j,i}$ ,  $y_{\sigma^2,j}$ ,  $y_{\sigma,j,i}$  and  $y_{\sigma^3}$ . Computing these derivatives largely resembles the computation of the second derivatives in the previous section. We relegate the details to the Appendix and focus on the results here.

To determine  $y_{k,j,i}$ , we differentiate (24) with respect to some shocks  $\varepsilon_{t-k}$ , delivering the third derivatives of the  $y$  function with respect to all triplets of the shocks. The resulting system, evaluated at  $\bar{x}$  and in expectation, of equations is a linear deterministic second order difference equation in  $y_{k,j,i}$ . The homogeneous components in (A-17) are identical to those in (16) and (25) and the

inhomogeneous components can again be rearranged to have a first order Markov representation

$$(39) \quad E_t(\mathcal{D}_{\epsilon_{t-k}^T \epsilon_{t-j}^T \epsilon_{t-i}^T}^3 f) \Big|_{\bar{x}} = f_y y_{k-1, j-1, i-1} + f_y y_{k, j, i} + f_y y_{k+1, j+1, i+1} + [f_{x^3} \quad f_{x^2} \quad f_{x^2} \quad f_{x^2}] \gamma_3 S_{k, j, i} = 0$$

for  $k, j, i = 0, 1, \dots$ , with  $y_{k, j, i} = 0$ , for  $k, j, i < 0$

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

$$(40) \quad y_{k, j, i} = \alpha y_{k-1, j-1, i-1} + \beta_3 S_{k, j, i}, \text{ with } y_{k, j, i} = 0, \text{ for } k, j, i < 0$$

and  $\beta_3$  can be solved for by, again, formulating an appropriate Sylvester equation.

To determine  $y_{\sigma, j, i}$ , we differentiate (24) with respect to  $\sigma$ , evaluate at  $\bar{x}$ , take expectations, set the resulting expression to zero, and recall the results from lower orders, yielding

$$(41) \quad E_t(\mathcal{D}_{\sigma \epsilon_{t-j}^T \epsilon_{t-i}^T}^3 f) \Big|_{\bar{x}} = f_y y_{\sigma, j-1, i-1} + f_y y_{\sigma, j, i} + f_y y_{\sigma, j+1, i+1} = 0$$

for  $j, i = 0, 1, \dots$ , with  $y_{\sigma, j, i} = 0$ , for  $j, i < 0$

or

$$(42) \quad y_{\sigma, j, i} = 0, \text{ for } j, i = 0, 1, \dots$$

again confirming Schmitt-Grohé and Uribe's (2004) and Jin and Judd's (2002) result that terms with a first order uncertainty are zero.

To determine  $y_{\sigma^2, i}$ , we differentiate (29) with respect to  $\sigma$ , evaluate at  $\bar{x}$ , take expectations, set the resulting expression to zero, and recall the results from lower orders, yielding

$$(43) \quad E_t(\mathcal{D}_{\sigma^2 \epsilon_{t-i}^T}^3 f) \Big|_{\bar{x}} = f_{x^3} \{[(x_{\bar{\epsilon}} \otimes x_{\bar{\epsilon}}) E_t(\epsilon_{t+1} \otimes \epsilon_{t+1})] \otimes x_i\} + 2f_{x^2}(x_{\bar{\epsilon}} \otimes x_{\bar{\epsilon}i}) [E_t(\epsilon_{t+1} \otimes \epsilon_{t+1}) \otimes I_{ne}]$$

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

= 0, for  $i = 0, 1, \dots$ , with  $y_{-1} = 0$

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 (16) and (25). The inhomogeneous components can again be rearranged to have a first order Markov representation and the unique stable solution of the forgoing takes the form

$$(44) \quad y_{\sigma^2, i} = \alpha y_{\sigma^2, i-1} + \beta_{\sigma} S_i, \text{ with } y_{\sigma^2, -1} = 0$$

where  $\beta_\sigma$  can be solved for by, again, formulating an appropriate Sylvester equation.

To determine  $y_{\sigma^3}$ , we differentiate (33) with respect to  $\sigma$ , evaluate at  $\bar{x}$ , take expectations, set the resulting expression to zero, and recall the results from lower orders, yielding

$$(45) \quad E_t(\mathcal{D}_{\sigma^3}^3 f) \Big|_{\bar{x}} = f_{x^3} [(x_{\bar{\varepsilon}} \otimes x_{\bar{\varepsilon}} \otimes x_{\bar{\varepsilon}}) E_t(\boldsymbol{\varepsilon}_{t+1} \otimes \boldsymbol{\varepsilon}_{t+1} \otimes \boldsymbol{\varepsilon}_{t+1})] + 2f_{x^2} [E_t(\boldsymbol{\varepsilon}_{t+1} \otimes \boldsymbol{\varepsilon}_{t+1} \otimes \boldsymbol{\varepsilon}_{t+1})(x_{\bar{\varepsilon}} \otimes x_{\bar{\varepsilon}^2})] \\ + f_{x^2} [(x_{\bar{\varepsilon}^2} \otimes x_{\bar{\varepsilon}}) E_t(\boldsymbol{\varepsilon}_{t+1} \otimes \boldsymbol{\varepsilon}_{t+1} \otimes \boldsymbol{\varepsilon}_{t+1})] + f_x [y_{\sigma^3} + x_{\bar{\varepsilon}^3} E_t(\boldsymbol{\varepsilon}_{t+1} \otimes \boldsymbol{\varepsilon}_{t+1} \otimes \boldsymbol{\varepsilon}_{t+1})] = 0$$

as the third moment of  $\boldsymbol{\varepsilon}_t$  is assumed given,  $E_t(\boldsymbol{\varepsilon}_{t+1} \otimes \boldsymbol{\varepsilon}_{t+1} \otimes \boldsymbol{\varepsilon}_{t+1})$  is known. Recovering  $y_{\sigma^3}$  from the forgoing is straightforward under the assumption (3.2). When  $\boldsymbol{\varepsilon}_t$  is normally distributed,<sup>19</sup> however,  $E_t(\boldsymbol{\varepsilon}_{t+1} \otimes \boldsymbol{\varepsilon}_{t+1} \otimes \boldsymbol{\varepsilon}_{t+1}) = 0$ . Hence

$$(46) \quad y_{\sigma^3} = 0$$

Combining, the third order approximation of the policy function (2) takes the form

$$(47) \quad y_t = \bar{y} + \frac{1}{2} y_{\sigma^2} \sigma^2 + \sum_{i=0}^{\infty} \left( y_i + \frac{1}{2} y_{\sigma^2, i} \sigma^2 \right) \boldsymbol{\varepsilon}_{t-i} + \frac{1}{2} \sum_{j=0}^{\infty} \sum_{i=0}^{\infty} y_{j,i} (\boldsymbol{\varepsilon}_{t-j} \otimes \boldsymbol{\varepsilon}_{t-i}) \\ + \frac{1}{6} \sum_{k=0}^{\infty} \sum_{j=0}^{\infty} \sum_{i=0}^{\infty} y_{k,j,i} (\boldsymbol{\varepsilon}_{t-k} \otimes \boldsymbol{\varepsilon}_{t-j} \otimes \boldsymbol{\varepsilon}_{t-i})$$

Again in contrast to the first order approximation, (47) does depend on  $\sigma$ , with the term  $\frac{1}{2} y_{\sigma^2}$  correcting the nonstochastic steady state for uncertainty as in the second order approximation (36), but now with  $\frac{1}{2} y_{\sigma^2, i} \sigma^2$  correcting the first order kernel for uncertainty; i.e., as  $\sigma$  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  $\sigma$  to one in (47) gives the third order approximation

$$(48) \quad y_t = \bar{y} + \frac{1}{2} y_{\sigma^2} + \sum_{i=0}^{\infty} \left( y_i + \frac{1}{2} y_{\sigma^2, i} \right) \boldsymbol{\varepsilon}_{t-i} + \frac{1}{2} \sum_{j=0}^{\infty} \sum_{i=0}^{\infty} y_{j,i} (\boldsymbol{\varepsilon}_{t-j} \otimes \boldsymbol{\varepsilon}_{t-i}) \\ + \frac{1}{6} \sum_{k=0}^{\infty} \sum_{j=0}^{\infty} \sum_{i=0}^{\infty} y_{k,j,i} (\boldsymbol{\varepsilon}_{t-k} \otimes \boldsymbol{\varepsilon}_{t-j} \otimes \boldsymbol{\varepsilon}_{t-i})$$

Higher order approximations of the policy function (2) can be computed using the same steps. Moving through higher orders of approximation successively, the undetermined partial derivatives of the policy function will always be terms of highest order being considered, ensuring that the leading

---

<sup>19</sup>As is the case in Dynare, see Adjemian, Bastani, Juillard, Mihoubi, Perendia, Ratto, and Villemot (2011).

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  $\sigma$  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  $\sigma$  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 two versions of the stochastic neoclassical growth model to demonstrate the method. This model has been used in numerous studies comparing numerical techniques and is a natural benchmark. We begin with the special case of log preferences in consumption and full depreciation that has a known solution to illustrate the relation of the nonlinear moving average to the more familiar state space solution. We then move on to the baseline specification of Aruoba, Fernández-Villaverde, and Rubio-Ramírez's (2006) comprehensive study with inseparable utility to foster comparability with their results. This version of the model lacks a known solution and must be approximated. Using our nonlinear moving average solution, we analyze the contributing elements to the response of the model's endogenous variables to a technology shock and highlight the features of the multidimensional kernels and impulse responses.

The model is populated by an infinitely lived representative household seeking to maximize its expected discounted lifetime utility given by

$$(49) \quad E_0 \left[ \sum_{t=0}^{\infty} \beta^t U(C_t, L_t) \right], \text{ with } U(C_t, L_t) = \frac{(C_t^\theta (1 - L_t)^{1-\theta})^{1-\gamma}}{1 - \gamma}$$

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

$$(50) \quad C_t + K_t = e^{Z_t} K_{t-1}^\alpha L_t^{1-\alpha} + (1 - \delta) K_{t-1}$$

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

$$(51) \quad Z_t = \rho_Z Z_{t-1} + \varepsilon_{Z,t}, \quad \varepsilon_{Z,t} \sim \mathcal{N}(0, \sigma_Z^2)$$

with  $\rho_Z \in (0, 1)$  a persistence parameter,  $\varepsilon_{Z,t}$  the innovation to the process, and  $\sigma_Z$  the standard deviation of the innovations.

The solution is characterized by the intertemporal Euler condition equalizing the expected present-discounted utility value of postponing consumption one period to its utility value today

$$(52) \quad \frac{\left(C_t^\theta (1 - L_t)^{1-\theta}\right)^{1-\gamma}}{C_t} = \beta E_t \left[ \frac{\left(C_{t+1}^\theta (1 - L_{t+1})^{1-\theta}\right)^{1-\gamma}}{C_{t+1}} \left(\alpha e^{Z_{t+1}} K_t^{\alpha-1} L_{t+1}^{1-\alpha} + 1 - \delta\right) \right]$$

and the intratemporal condition equalizing the utility cost of marginally increasing labor supply to the utility value of the additional consumption provided therewith

$$(53) \quad \frac{1 - \theta}{1 - L_t} = \frac{\theta}{C_t} (1 - \alpha) e^{Z_t} K_{t-1}^\alpha L_t^{-\alpha}$$

plus the budget constraint (50) and the technology shock (51). Collecting the four equations into a vector of functions, the set of equilibrium conditions can be written  $0 = E_t[f(y_{t-1}, y_t, y_{t+1}, u_t)]$  where  $y_t = [C_t \ K_t \ L_t \ Z_t]'$  and  $u_t = [\varepsilon_{Z,t}]$ .

## 4.1 Logarithmic Preferences and Complete Depreciation Special Case

The first case we will examine is the simple stochastic neoclassical growth model 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 example 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 (49) be given by  $\ln(C_t)$ ,<sup>20</sup> normalize  $L_t = 1$  and set  $\delta = 1$  in (50).

Combining (50) with (52) in this case yields

$$(54) \quad 0 = E_t \left[ (e^{Z_t} K_{t-1}^\alpha - K_t)^{-1} - \beta (e^{Z_{t+1}} K_t^\alpha - K_{t+1})^{-1} (\alpha e^{Z_{t+1}} K_t^{\alpha-1}) \right]$$

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 nonlinear moving average representation and guess that the logarithm of the solution is linear in the infinite history of technology innovations

$$(55) \quad \ln(K_t) = \ln(\bar{K}) + \sum_{j=0}^{\infty} b_j \varepsilon_{Z,t-j}$$

Inserting the guess and the infinite moving average representation for  $Z_t$ , (54) can be rewritten

$$(56) \quad 1 = \alpha\beta E_t \left[ \frac{1 - \exp\left(\sum_{j=0}^{\infty} (\rho^j - b_j + \alpha b_{j-1}) \varepsilon_{Z,t-j} - (1 - \alpha) \ln(\bar{K})\right)}{1 - \exp\left(\sum_{j=0}^{\infty} (\rho^j - b_j + \alpha b_{j-1}) \varepsilon_{Z,t+1-j} - (1 - \alpha) \ln(\bar{K})\right)} \right. \\ \left. \times \exp\left(\sum_{j=0}^{\infty} (\rho^j - b_j + \alpha b_{j-1}) \varepsilon_{Z,t} - (1 - \alpha) \ln(\bar{K})\right) \right]$$

where  $b_{-1} = 0$ .

The value and recursion

$$(57) \quad \bar{K} = (\alpha\beta)^{\frac{1}{1-\alpha}}, \quad b_j = \alpha b_{j-1} + \rho^j, \quad \text{with } b_{-1} = 0$$

solve (56) and verify the guess, (55).

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

$$(58) \quad K_t = (\alpha\beta)^{\frac{1}{1-\alpha}} \exp\left((1 - \alpha L)^{-1} \rho(L) \varepsilon_{Z,t}\right) = (\alpha\beta)^{\frac{1}{1-\alpha}} \exp\left(\sum_{j=0}^{\infty} b_j \varepsilon_{Z,t-j}\right)$$

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

---

<sup>20</sup>That is, set  $\theta$  and  $\gamma$  to one, subtracting an appropriate constant and extending the utility function over the removable singularity at  $\gamma = 1$ .

use  $K_t = \exp(\hat{K}_t)$  to reexpress (54) as<sup>21</sup>

$$(59) \quad 0 = E_t \left[ \left( e^{Z_t + \alpha \hat{K}_{t-1}} - e^{\hat{K}_t} \right)^{-1} - \beta \left( e^{z_{t+1} + \alpha \hat{K}_t} - e^{\hat{K}_{t+1}} \right)^{-1} \left( \alpha e^{Z_{t+1} + (\alpha-1)\hat{K}_t} \right) \right]$$

With this reformulation, the first order expansion is the true policy rule in this special case. That is, (59) can be rewritten as  $0 = E_t[f(y_{t-1}, y_t, y_{t+1}, u_t)]$  where  $y_t = [K_t \ Z_t]'$  and  $u_t = [\varepsilon_{Z,t}]$ .

To check our method, we calculate the kernels of the third order accurate nonlinear moving average solution of (59) out 500 periods, following the parameterization of Hansen (1985) for the remaining parameters by setting  $\alpha = 0.36$ ,  $1/\beta = 1.01$ ,  $\rho = 0.95$ , and  $\sigma_Z = 0.00712$ . 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  $4.3368 \times 10^{-18}$ . This first check, while encouraging, is far from comprehensive. In section 5, additional and potentially more meaningful measures will be examined.

## 4.2 CRRA-Incomplete Depreciation Case

We now move to the general case of Aruoba, Fernández-Villaverde, and Rubio-Ramírez (2006). Following their parameterization, we relax the complete depreciation and log preferences of the previous section, see table 1. As no known closed form solution exists, we will need an approximation. We reexpress variables in logs, commensurate with a loglinear approximation. This choice is additionally motivated by our results in section 5 that indicate a log specification improves the accuracy of the approximation.

[Table 1 about here.]

For higher-order approximations, our policy function (2),  $y_t = y(\sigma, \varepsilon_t, \varepsilon_{t-1}, \varepsilon_{t-2}, \dots)$ , will straightforwardly enable impulse response analysis. That is, consider a shock in  $t$  to an element of  $\varepsilon_t$ , one

---

<sup>21</sup>See Fernández-Villaverde and Rubio-Ramírez (2006) for more on change of variable techniques such as this.

measure<sup>22</sup> for the response of  $y_t$  through time to this impulse is given by the sequence

$$\begin{aligned}
 y_t &= y(\sigma, \varepsilon_t, 0, 0, 0, \dots) \\
 y_{t+1} &= y(\sigma, 0, \varepsilon_t, 0, 0, \dots) \\
 y_{t+2} &= y(\sigma, 0, 0, \varepsilon_t, 0, \dots) \\
 &\vdots
 \end{aligned}
 \tag{60}$$

[Figure 1 about here.]

Figure 1 depicts the impulse responses and their contributing components from the kernels of different orders for capital, consumption, and labor to a positive, one standard deviation shock in  $\varepsilon_{Z,t}$ .<sup>23</sup> 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 in accordance with the artifact of harmonic distortion discussed in Priestly (1988, p. 27).

The second order contributions of capital and consumption are positive and that of labor is negative. This reflects the combination of a precautionary reaction and nonlinear propagation mechanism

<sup>22</sup> Note that we are assuming that  $y_{t-j} = y(\sigma, 0, 0, \dots)$ ,  $\forall j > 0$ . Fernández-Villaverde, Guerrón-Quintana, Rubio-Ramírez, and Uribe (2011), 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.

<sup>23</sup>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 ignoring 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 technology shocks. A technology shock is associated with a larger capital stock, which enables a larger increase in consumption (but of an order of magnitude smaller than capital in terms of second order contribution, as a precautionary reaction) and a smaller increase in labor (due to the second order downward correction) than the linear model would predict. In the case of a negative technology shock (not pictured), the first order components would simply be their mirror images with opposite sign. The second order contributions, however, would remain entirely unchanged following from the symmetry of the quadratic. In combination, the second order approximation can thus capture time invariant asymmetries in the impulse responses.<sup>24</sup>

The precautionary component can likewise be seen in the upward correction of the steady states in the rightmost panels. In the stochastic steady state, agents face uncertainty regarding future shocks and accumulate a precautionary stock of capital through increased labor efforts and disburse this as increased consumption when shocks fail to manifest themselves. The lower left panel contains the contributions from  $y_{\sigma^2 i}$  the second order (in  $\sigma$ ) time varying correction for risk, this demonstrates an initial wealth effect with consumption increasing and capital and labor decreasing relative to a nonstochastic environment.<sup>25</sup> Nonlinear impulse responses are not scale invariant, as noted also by Fernández-Villaverde, Guerrón-Quintana, Rubio-Ramírez, and Uribe (2011): for example, while the first order component scales linearly with the magnitude of the shock, the second 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

---

<sup>24</sup>Time varying asymmetries would be captured by  $y_{\sigma^2, j, i}$ , require a fourth order approximation as the term  $y_{\sigma, j, i}$  from the third order approximation is zero, see section 3.3.

<sup>25</sup>Fernández-Villaverde and Rubio-Ramírez (2010) discusses the nonlinear impact of shocks in the production function and similar wealth effects.

the conventional wisdom that this model is nearly linear.

[Figure 2 about here.]

In figure 2, the impulse responses to a technology shock with different values (2, 5, and 10) of the CRRA parameter  $\gamma$  are overlaid. Note that for all three values of  $\gamma$ , the first order components dominate. While changes in  $\gamma$  do change the periodicity of the harmonic distortion as well as the shape and sign of some second and third order components, the constant and time varying corrections for risk display a significant change in magnitude. As  $\gamma$  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. Though not very large, the precautionary channel of the second order kernel is highlighted by the experiment, with both the second order contributions of capital and labor increasing minimally and that of consumption decreasing initially. 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 (2011), impulse responses are not invariant to the history of shocks. The third order kernels in figure 4 are four dimensional objects, captured by our use of colors inside a cube; we slice the cube with a diagonal plane whose diagonal ( $i = j = k$ ) is the third order contribution (though not scaled to the magnitude of the shock) in figure 1. The unscaled contribution from the second order kernel in the impulse responses in figure 1 can be found along the diagonals of the kernels ( $i = j$ ) in figures 3. 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 figure 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 the kernels for consumption and capital that bottom out at about fifty periods indicate a substantially persistent nature or ‘memory’ 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 figures 3a and 3b, the shape is reminiscent of the parabola of a quadratic equation and the ‘s’ shape of the cubic equation can be found on either side of the diagonal of figure 4. 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 response 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, consumption, and labor 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, the cross component overwhelms the individual components shortly after the second shocks hits and the second order contributions to the responses of capital (upper panel) and consumption (middle panel) fail to match the peak response

from a single shock, despite the lingering contribution from the initial shock in the same direction. Although the mitigation is much less pronounced for labor (lower panel), the difference from the sum of individual contributions is nonetheless noticeable and prolonged. In a nonlinear environment, there is no single measure for an impulse response;<sup>26</sup> in starting from the stochastic steady state, however, we remove any deterministic trends (as would be present, e.g, when starting from the ergodic mean, see footnote 22) in our impulse response measure at each order of approximation.

## 5 Accuracy

In this section, we explore the accuracy of our solution method using Euler equation error methods.<sup>27</sup> 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 examine our method using the model of Aruoba, Fernández-Villaverde, and Rubio-Ramírez (2006), examined in section 4.2. 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, (52) can be rearranged to deliver the Euler equation error function as<sup>28</sup>

$$(61) \quad EE() = 1 - \frac{1}{C_t} \left( \frac{\beta E_t \left[ \frac{(C_{t+1}^\theta (1-L_{t+1})^{1-\theta})^{1-\gamma}}{C_{t+1}} \left( \alpha e^{Z_{t+1}} K_t^{\alpha-1} L_{t+1}^{1-\alpha} + 1 - \delta \right) \right]}{(1-L_t)^{(1-\theta)(1-\gamma)}} \right)^{\frac{1}{\theta(1-\gamma)-1}}$$

Deviations in (61) 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 error of  $-6$  implies a one dollar error for every million dollars spent.

<sup>26</sup>See, e.g., Gourieroux and Jasiak (2005), Potter (2000), and Koop, Pesaran, and Potter (1996).

<sup>27</sup>See, e.g., Judd (1992), Judd and Guu (1997), and Judd (1998)

<sup>28</sup>Cf. Aruoba, Fernández-Villaverde, and Rubio-Ramírez (2006, p. 2499).

The arguments of  $EE()$  depend on the state space postulated. Standard state space methods would choose  $EE(K_{t-1}, Z_t)$  or  $EE(K_{t-1}, Z_{t-1}, \varepsilon_{Z,t})$ . Our nonlinear moving average policy function requires  $EE(\varepsilon_{Z,t}, \varepsilon_{Z,t-1}, \dots)$ , 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.

$$(62) \quad EE_t = EE(\varepsilon_{Z,t}, 0, 0, \dots), EE_{t-1} = EE(0, \varepsilon_{Z,t-1}, 0, \dots), EE_{t-2} = EE(0, 0, \varepsilon_{Z,t-2}, \dots), \dots$$

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 6 plots  $EE_t$  for first through third order approximations in logs, see section 4.2, and in the variables' original level specification. 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 necessarily happen monotonically. Second, switching to a log specification improves the first and second order approximations uniformly, while for the third order, this is true only for very small and very large shocks. If we restrict our attention to three standard deviation shocks ( $\pm 0.021$ ), the second order log approximation make mistakes no greater than one dollar for everyone ten million spent and the third order level and log approximations no greater than one dollar for everyone one hundred 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). As their focus was on the mapping from capital to errors and ours on shocks to errors, it is possible that the preferred approximation depends on the dimension under study.

[Figure 6 about here.]

In figure 7, plots of  $EE_{t-j}$  for  $j = 0, 1, \dots, 100$  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 7 about here.]

First, we plot the maximal Euler equation errors over a span of 100 periods in figure 8a. I.e.,

$$(63) \quad \max_{-10\sigma_Z < \varepsilon_{Z,t-j} < 10\sigma_Z} (EE_{t-j}), \text{ for } j = 0, 1, \dots, 100$$

where  $\sigma_Z$  is the standard deviation of the technology shock, see table 1. The figure tends to reinforce the results from examining only shocks in period  $t$ : for both the level and log approximations, moving to a higher order uniformly improves the quality of approximation and, at all three orders, moving from a level to a log specification likewise improves the accuracy of the approximation uniformly according to this metric.

[Figure 8 about here.]

In our final measure, we graph average Euler equation errors over a span of 100 periods in figure 8b. In contrast to state space measures that require the ergodic distribution of endogenous state variables, our measure is relatively easy to calculate with a nonlinear moving average policy function, as we merely need to integrate with respect to the known distribution (in this case normal) of the shocks

$$(64) \quad \int EE_{t-j} dF_{\varepsilon_{Z,t-j}}, \text{ for } j = 0, 1, \dots, 100$$

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 and the improvement in the approximation by switching to logs. Though now there is some ambiguity regarding the preferred specification among the third in levels and second and third in logs, with the second order log specification surpassing both third order approximations between 25 and 35 quarters. 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, putting the ambiguity mentioned above in perspective.

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 are all perturbations, our method remains a local method and is subject to all the limitations and reservations that face such methods.

## 6 Conclusion

We have introduced a nonlinear infinite moving average as an alternative to the standard state space policy function to the dynamical analysis of nonlinear DSGE models. We have derived a perturbation approximation of this policy function, providing explicit derivations up to third order in the form of a Volterra expansion. This direct mapping of the history of shocks into endogenous variables enables familiar impulse response analysis techniques in a nonlinear environment, and provides a convenient decomposition on the mapping from approximation order and uncertainty. We confirm that this approach provides a solution with a degree of accuracy comparable to state space methods by introducing Euler equation error methods for this infinite dimensional mapping.

Although there are a number of DSGE models and applications, for example, welfare analysis, asset pricing and stochastic volatility for which the importance of nonlinear components and uncertainty in the policy function has been proved, the nonlinear components we analyzed in the baseline neoclassical growth model are quantitatively unimportant, this is not surprising as the model is known to be nearly linear. Qualitatively however, the nonlinear contributions to the the mapping from shocks to endogenous variables are economically interpretable, translating, e.g., into precautionary behavior and wealth effects. Likewise, non economically interpretable artifacts of the nonlinear method, such as harmonic distortion are documented as well.

The potential for explosive behavior in the simulation of state space perturbations has lead to the adaptation of ‘pruning’ algorithms, see Kim, Kim, Schaumburg, and Sims (2008), that appear ad-hoc relative to the perturbation solution itself. With our method, however, the stability from the first order solution is passed on to all higher order recursions. This feature of the nonlinear kernels in our moving average solution is consistent with the Volterra operator acting upon the history of shocks being bounded and the existence of an endogenous, perturbation-based ‘pruning’ algorithm derived from inverting our moving average, both of which we study in ongoing research.

The nonlinear perturbation DSGE literature is still in an early stage of development and our method provides a different perspective by mapping directly from the history of shocks. 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 additional 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 Perturbation 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., L. BOCOLA, AND F. SCHORFHEIDE (2011): “A New Class of Nonlinear Time Series Models for the Evaluation of DSGE Models,” Mimeo December.
- 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.
- COLLARD, F., AND M. JUILLARD (2001a): “Accuracy of Stochastic Perturbation 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.
- FERNÁNDEZ-VILLAVERDE, J., P. A. GUERRÓN-QUINTANA, J. RUBIO-RAMÍREZ, AND M. URIBE (2011): “Risk Matters: The Real Effects of Volatility Shocks,” *American Economic Review*, 101(6), 2530–61.
- 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.
- HENDERSON, H. V., AND S. R. SEARLE (1981): “The vec-Permutation Matrix, the vec Operator and Kronecker Products: A Review,” *Linear and Multilinear Algebra*, 9(4), 271–288.
- JIN, H.-H., AND K. L. JUDD (2002): “Perturbation 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.
- 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. (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.
- WOODFORD, M. (1986): “Stationary Sunspot Equilibria: The Case of Small Fluctuations around a Deterministic Steady State,” Mimeo September.

# 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 (1973).]*

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

$$(A-2) \quad \mathcal{D}_B A(B) \equiv \left[ \mathcal{D}_{b_{kl}} A(B) \right]_{sp \times tq} = \begin{bmatrix} \mathcal{D}_{b_{11}} A(B) & \cdots & \mathcal{D}_{b_{1t}} A(B) \\ \vdots & & \vdots \\ \mathcal{D}_{b_{s1}} A(B) & \cdots & \mathcal{D}_{b_{st}} A(B) \end{bmatrix}$$

*Structures of higher derivatives are thereby uniquely defined*

$$(A-3) \quad \mathcal{D}_{B^n} A(B) \equiv \mathcal{D}_B (\mathcal{D}_B (\cdots (\mathcal{D}_B A(B)) \cdots))$$

2. *Abbreviated Notation, where  $T$  indicates transposition,*

$$(A-4) \quad A_B \equiv \mathcal{D}_{B^T} A(B), \quad A_{B^n} \equiv \mathcal{D}_{(B^T)^n} A(B), \quad \text{and} \quad A_{CB} \equiv \mathcal{D}_{C^T} (\mathcal{D}_{B^T} A(B, C)) = \mathcal{D}_{C^T B^T} A(B, C)$$

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

$$(A-5) \quad \begin{matrix} B \\ m \times t \end{matrix} \otimes \begin{matrix} A \\ n \times s \end{matrix} = K_{m,n} (A \otimes B) K_{s,t}$$

### 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. Commute the term  $H_b \otimes F$  and note that

$$(A-6) \quad K_{q,vs} \left( \begin{matrix} I \\ s \times s \end{matrix} \otimes K_{v,q} \right) = \left( K_{q,s} \otimes \begin{matrix} I \\ v \times v \end{matrix} \right) \left( \begin{matrix} I \\ s \times s \end{matrix} \otimes K_{q,v} \right) \left( \begin{matrix} I \\ s \times s \end{matrix} \otimes K_{v,q} \right) = \left( K_{q,s} \otimes \begin{matrix} I \\ v \times v \end{matrix} \right)$$

where the first equality follows from Henderson and Searle (1981, p. 283) and the second from Magnus and Neudecker (1979, p. 383).

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 (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

$$(A-7) \quad M \begin{pmatrix} B \\ (p \times 1) \end{pmatrix} \begin{pmatrix} \\ (s \times 1) \end{pmatrix} = M(\bar{B}) + \sum_{n=1}^N \frac{1}{n!} \mathcal{D}_{B^{Tn}}^n M(\bar{B}) (B - \bar{B})^{\otimes [n]} + R_{N+1}(\bar{B}, B)$$

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

Differentiating (2) 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

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

Writing the foregoing more compactly yields (10) in the text.

## A.2 Auxiliary Matrices

### A.2.1 Shifting Matrices

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

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

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

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

## A.2.2 State Spaces for the Markov Representation

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

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

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

## A.3 Details of Third-Order Derivation

We begin by differentiating  $f$  with respect to each triplet of shocks. The resulting system of equations remains linear in the third derivatives

$$(A-16) \quad \begin{aligned} \mathcal{D}_{\xi_{t-k}^T \xi_{t-j}^T \xi_{t-i}^T}^3 f &= f_{x^3}(x_k \otimes x_j \otimes x_i) + f_{x^2}(x_{k,j} \otimes x_i) \\ &+ f_{x^2}(x_j \otimes x_{k,i})(K_{ne,ne} \otimes I_{ne}) + f_{x^2}(x_k \otimes x_{j,i}) + f_{x^2} x_{k,j,i} \end{aligned}$$

Evaluating this at  $\bar{y}$  and setting its expectation to zero yields

$$\begin{aligned}
E_t(\mathcal{D}_{\varepsilon_{t-k}^T \varepsilon_{t-j}^T \varepsilon_{t-i}^T}^3 f) \Big|_{\bar{x}} &= f_{y^-} y_{k-1, j-1, i-1} + f_y y_{k, j, i} + f_{y^+} y_{k+1, j+1, i+1} \\
&\quad + f_{x^3}(x_k \otimes x_j \otimes x_i) + f_{x^2}(x_{k, j} \otimes x_i) \\
&\quad + f_{x^2}(x_j \otimes x_{k, i})(K_{ne, ne} \otimes I_{ne}) + f_{x^2}(x_k \otimes x_{j, i}) \\
\text{(A-17)} \qquad \qquad \qquad &= 0, \text{ for } k, j, i = 0, 1, \dots, \text{ with } y_{k, j, i} = 0, \text{ for } k, j, i < 0
\end{aligned}$$

a linear deterministic second order difference equation in the third derivative  $y_{k, j, i}$ . The homogeneous components in (A-17) are identical to those in (16) and (25). The inhomogeneous components again have a first order Markov representation. Using the shifting and transition matrices defined in appendix A.2 gives (39) of the main text, whose solution takes the form (40). By recursively substituting (40) in (A-17), using the shifting matrices and matching coefficients, we obtain a Sylvester equation in  $\beta_3$

$$\text{(A-18)} \qquad (f_y + f_{y^+} \alpha) \beta_3 + f_{y^+} \beta_3 \delta_3 = - [f_{x^3} \quad f_{x^2} \quad f_{x^2} \quad f_{x^2}] \gamma_3$$

Now we move on to the partial derivatives of  $y$  function involving the perturbation parameter  $\sigma$ .

To determine  $y_{\sigma, j, i}$ , we differentiate  $f$  with respect to  $\varepsilon_{t-i}$ ,  $\varepsilon_{t-j}$  and  $\sigma$

$$\begin{aligned}
\mathcal{D}_{\sigma \varepsilon_{t-j}^T \varepsilon_{t-i}^T}^3 f &= f_{x^3}(\mathcal{D}_\sigma x \otimes x_j \otimes x_i) + f_{x^2}(\mathcal{D}_\sigma x \otimes x_{j, i}) + f_{x^2}(\mathcal{D}_\sigma x_j \otimes x_i) \\
\text{(A-19)} \qquad \qquad \qquad &\quad + f_{x^2}(x_j \otimes \mathcal{D}_\sigma x_i) K_{ne, ne} + f_x \mathcal{D}_\sigma x_{j, i} \\
&\quad \text{where } \mathcal{D}_\sigma x_{j, i} = x_{\sigma, j, i} + x_{\tilde{\varepsilon}, j, i}(\varepsilon_{t+1} \otimes I_{ne^2})
\end{aligned}$$

Evaluating at  $\bar{y}$ , taking expectations, setting the resulting expression to zero yields, and noting the results from lower orders yields the expression in the text, whose solution, again analogously to lower orders, takes the form  $y_{\sigma, j, i} = \alpha y_{\sigma, j-1, i-1}$ , with  $y_{\sigma, j, i} = 0$ , for  $j, i < 0$  delivering (42) in the main text.

To determine  $y_{\sigma^2, i}$ , we differentiate  $f$  with respect to  $\varepsilon_{t-i}$  once and  $\sigma$  twice

$$\begin{aligned}
\mathcal{D}_{\sigma^2 \varepsilon_{t-i}^T}^3 f &= f_{x^3}(\mathcal{D}_\sigma x \otimes \mathcal{D}_\sigma x \otimes x_i) + f_{x^2}(\mathcal{D}_\sigma x \otimes \mathcal{D}_\sigma x_i) + f_{x^2}(\mathcal{D}_{\sigma^2}^2 x \otimes x_i) + f_{x^2}(\mathcal{D}_\sigma x \otimes \mathcal{D}_\sigma x_i) + f_x \mathcal{D}_{\sigma^2}^2 x_i \\
\text{(A-20)} \qquad \qquad \qquad &\quad \text{where } \mathcal{D}_{\sigma^2}^2 x_i = x_{\sigma^2, i} + 2x_{\sigma, \tilde{\varepsilon}, i}(\varepsilon_{t+1} \otimes I_{ne}) + x_{\tilde{\varepsilon}^2, i}(\varepsilon_{t+1} \otimes \varepsilon_{t+1} \otimes I_{ne})
\end{aligned}$$

Evaluating at the nonstochastic steady state ( $\bar{y}$ ), taking expectations, and setting the resulting expression to zero yields the expression in the main text, 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 (16) and (25). 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

$$\begin{aligned}
& f_y y_{\sigma^2,i-1} + f_y y_{\sigma^2,i} + f_y y_{\sigma^2,i+1} \\
& + \left\{ \left[ f_{x^3}(\gamma_4 \beta_1 \otimes \gamma_4 \beta_1 \otimes \gamma_1) + f_{x^2}([\gamma_4 \beta_2(S_0 \otimes S_0)] \otimes \gamma_1) + 2f_{x^2}(\gamma_4 \beta_1 \otimes [\gamma_4 \beta_2(S_0 \otimes \delta_1)]) \right] \right. \\
\text{(A-21)} \quad & \left. + f_x \gamma_4 \beta_3 \gamma_5(S_0 \otimes S_0 \otimes \delta_1) \right] [E_t(\boldsymbol{\varepsilon}_{t+1} \otimes \boldsymbol{\varepsilon}_{t+1}) \otimes I_{ny+ne}] + f_{x^2}(x_{\sigma^2} \otimes \gamma_1) \Big\} S_i = 0 \\
& \text{for } i = 0, 1, \dots, \text{ with } y_{-1} = 0
\end{aligned}$$

The solution of the forgoing takes the form of (44) in the main text. Substituting (44) in (A-21) and matching coefficients, we obtain a Sylvester equation in  $\beta_{\sigma}$

$$\begin{aligned}
\text{(A-22)} \quad & (f_y + f_{y+\alpha})\beta_{\sigma} + f_{y+\alpha}\beta_{\sigma}\delta_1 = - \left\{ \left[ f_{x^3}(\gamma_4 \beta_1 \otimes \gamma_4 \beta_1 \otimes \gamma_1) + f_{x^2}([\gamma_4 \beta_2(S_0 \otimes S_0)] \otimes \gamma_1) \right. \right. \\
& \left. \left. + 2f_{x^2}(\gamma_4 \beta_1 \otimes [\gamma_4 \beta_2(S_0 \otimes \delta_1)]) + f_x \gamma_4 \beta_3 \gamma_5(S_0 \otimes S_0 \otimes \delta_1) \right] [E_t(\boldsymbol{\varepsilon}_{t+1} \otimes \boldsymbol{\varepsilon}_{t+1}) \otimes I_{ny+ne}] + f_{x^2}(x_{\sigma^2} \otimes \gamma_1) \right\}
\end{aligned}$$

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

$$\begin{aligned}
\text{(A-23)} \quad & \mathcal{D}_{\sigma^3}^3 f = f_{x^3}(\mathcal{D}_{\sigma} x \otimes \mathcal{D}_{\sigma} x \otimes \mathcal{D}_{\sigma} x) + 2f_{x^2}(\mathcal{D}_{\sigma} x \otimes \mathcal{D}_{\sigma^2}^2 x) + f_{x^2}(\mathcal{D}_{\sigma^2}^2 x \otimes \mathcal{D}_{\sigma} x) + f_x \mathcal{D}_{\sigma^3}^3 x \\
& \text{where } \mathcal{D}_{\sigma^3}^3 x = x_{\sigma^3} + 3x_{\sigma^2, \tilde{\varepsilon}} \boldsymbol{\varepsilon}_{t+1} + 3x_{\sigma, \tilde{\varepsilon}^2}(\boldsymbol{\varepsilon}_{t+1} \otimes \boldsymbol{\varepsilon}_{t+1}) + x_{\tilde{\varepsilon}^3}(\boldsymbol{\varepsilon}_{t+1} \otimes \boldsymbol{\varepsilon}_{t+1} \otimes \boldsymbol{\varepsilon}_{t+1})
\end{aligned}$$

Evaluating this at the nonstochastic steady state and setting its expectation to zero yields

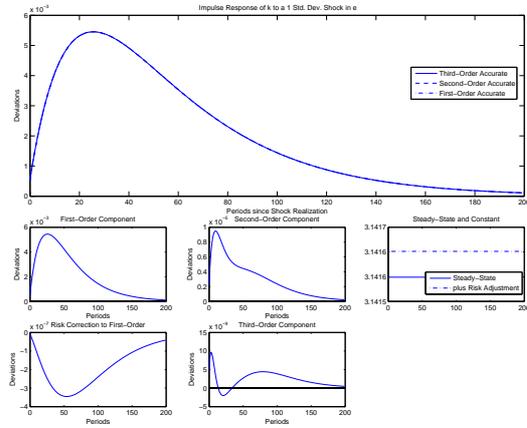
$$\begin{aligned}
& E_t(\mathcal{D}_{\sigma^3}^3 f) \Big|_{\bar{x}} = f_{x^3}[(x_{\tilde{\varepsilon}} \otimes x_{\tilde{\varepsilon}} \otimes x_{\tilde{\varepsilon}}) E_t(\boldsymbol{\varepsilon}_{t+1} \otimes \boldsymbol{\varepsilon}_{t+1} \otimes \boldsymbol{\varepsilon}_{t+1})] + 2f_{x^2}[E_t(\boldsymbol{\varepsilon}_{t+1} \otimes \boldsymbol{\varepsilon}_{t+1} \otimes \boldsymbol{\varepsilon}_{t+1})(x_{\tilde{\varepsilon}} \otimes x_{\tilde{\varepsilon}^2})] \\
& \quad + f_{x^2}[(x_{\tilde{\varepsilon}^2} \otimes x_{\tilde{\varepsilon}}) E_t(\boldsymbol{\varepsilon}_{t+1} \otimes \boldsymbol{\varepsilon}_{t+1} \otimes \boldsymbol{\varepsilon}_{t+1})] + f_x[y_{\sigma^3} + x_{\tilde{\varepsilon}^3} E_t(\boldsymbol{\varepsilon}_{t+1} \otimes \boldsymbol{\varepsilon}_{t+1} \otimes \boldsymbol{\varepsilon}_{t+1})] \\
\text{(A-24)} \quad & = 0
\end{aligned}$$

the expression in the text.

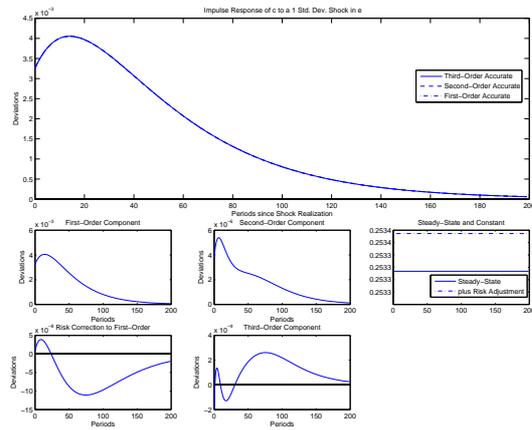
Table 1: Parameter Values for the Model of Section 4.2

Parameter	$\beta$	$\tau$	$\theta$	$\alpha$	$\delta$	$\rho_Z$	$\sigma_Z$
Value	0.9896	2.0	0.357	0.4	0.0196	0.95	0.007

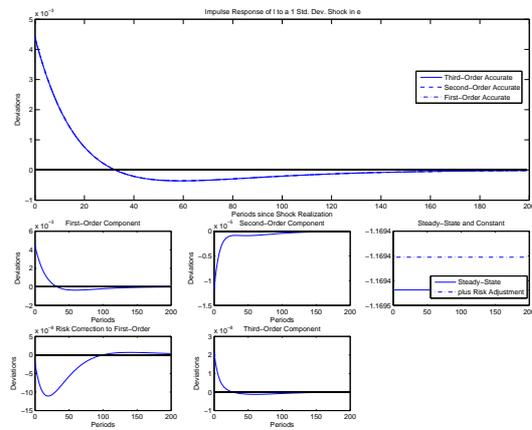
See Aruoba, Fernández-Villaverde, and Rubio-Ramírez (2006).



(a) Capital

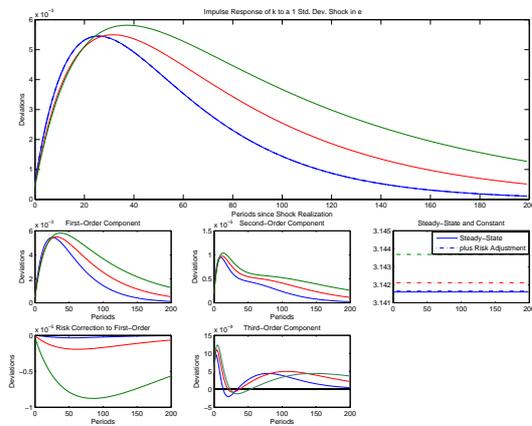


(b) Consumption

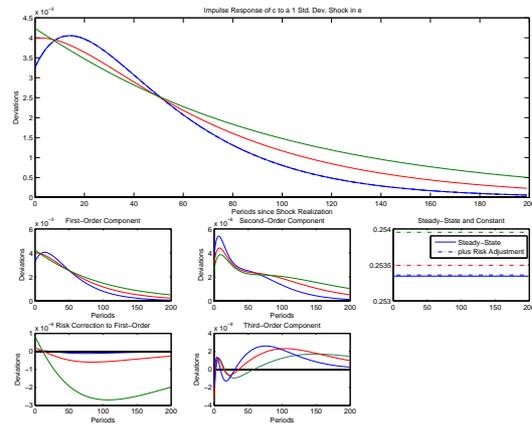


(c) Labor

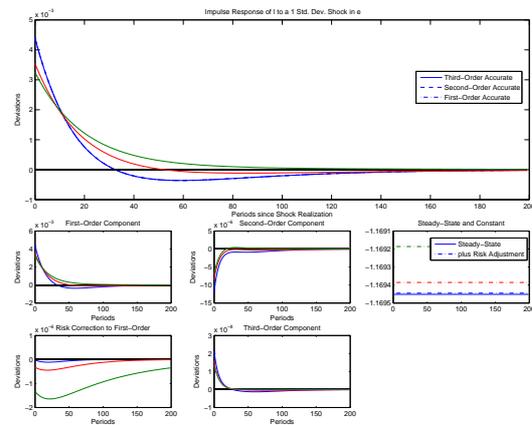
Figure 1: Impulse Responses to a Technology Shock, Model of Section 4.2



(a) Capital

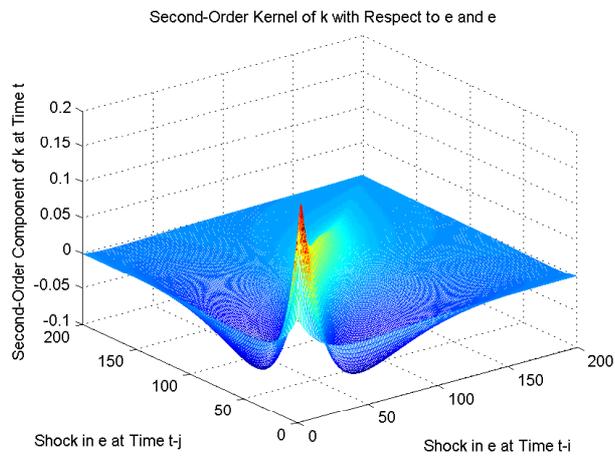


(b) Consumption

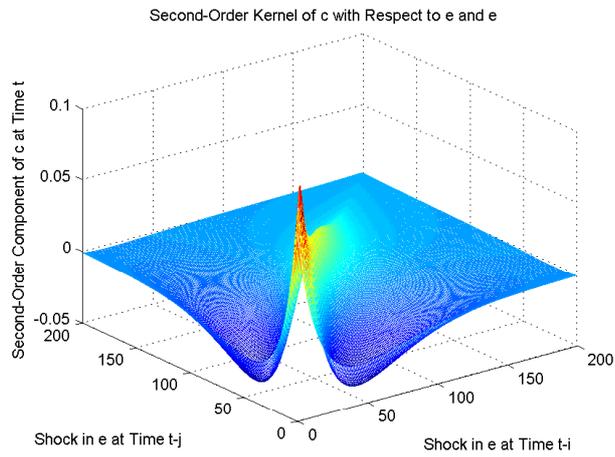


(c) Labor

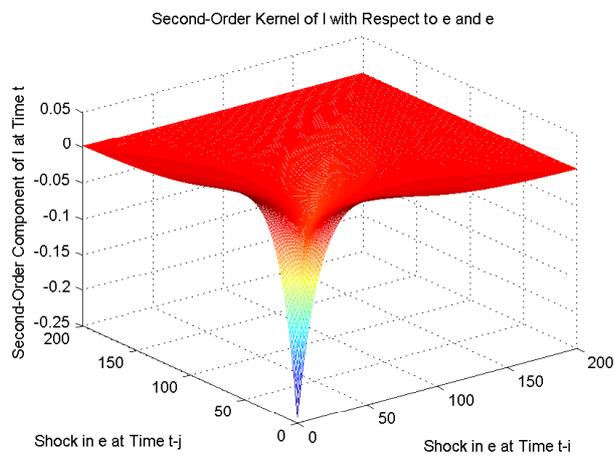
Figure 2: Impulse Responses to a Technology Shock, Model of Section 4.2  
Blue:  $\gamma = 2$ , Red  $\gamma = 5$ , Green  $\gamma = 10$



(a) Capital

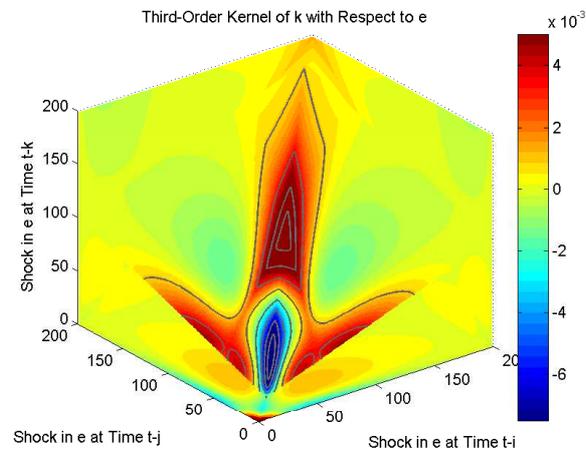


(b) Consumption

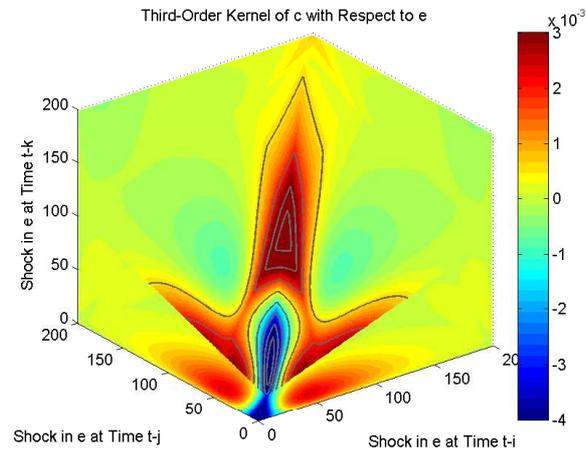


(c) Labor

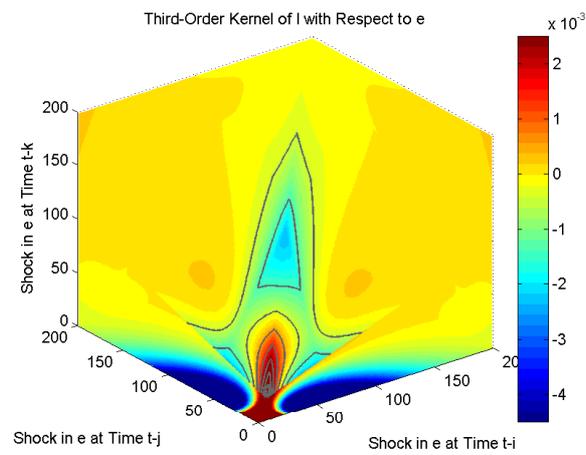
Figure 3: Second-Order Kernels, Model of Section 4.2



(a) Capital

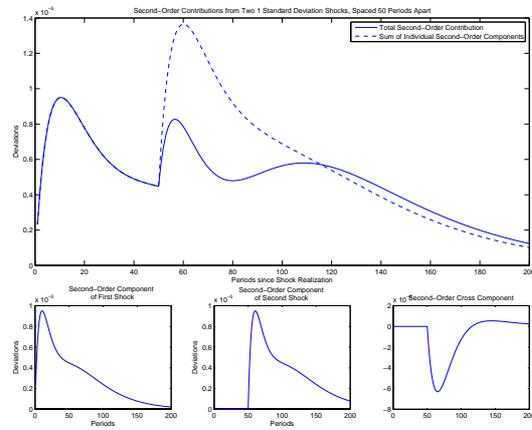


(b) Consumption

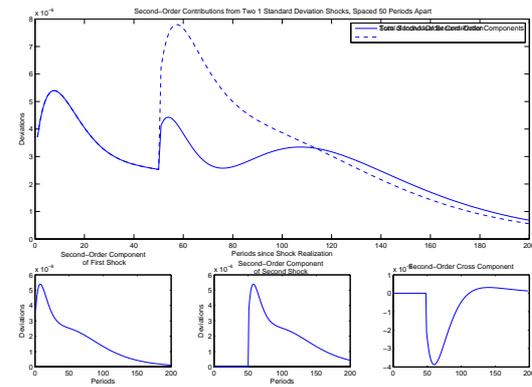


(c) Labor

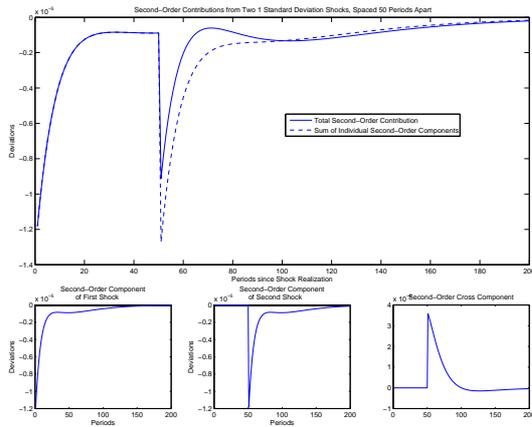
Figure 4: Third-Order Kernels, Model of Section 4.2



(a) Capital



(b) Consumption



(c) Labor

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

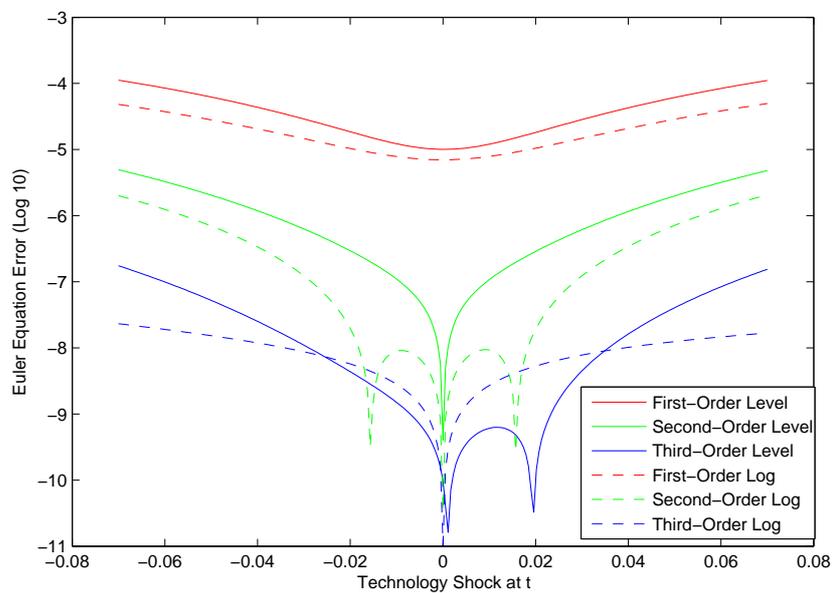
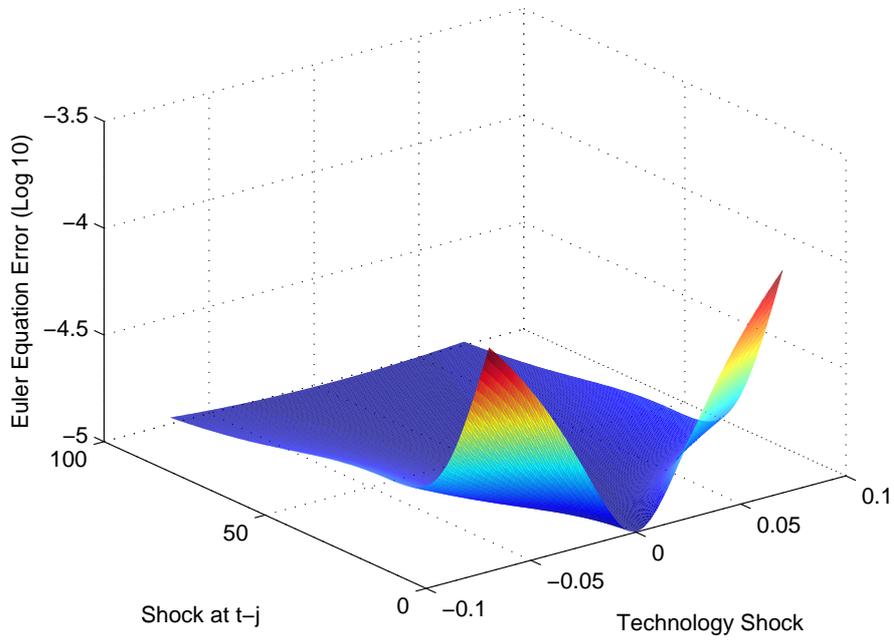
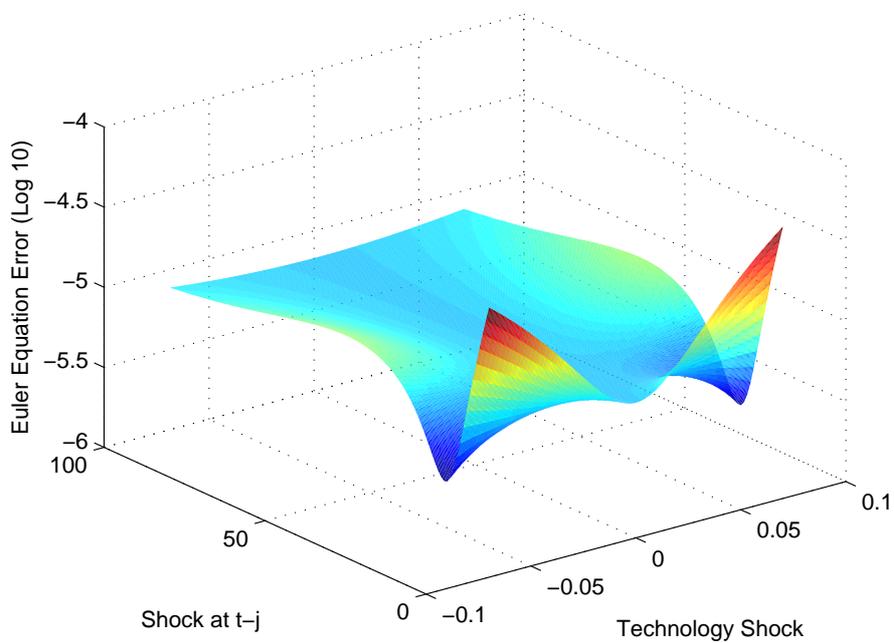


Figure 6: Euler Equation Errors, Shock at Time  $t$ , Aruoba, Fernández-Villaverde, and Rubio-Ramírez's (2006) Baseline Case

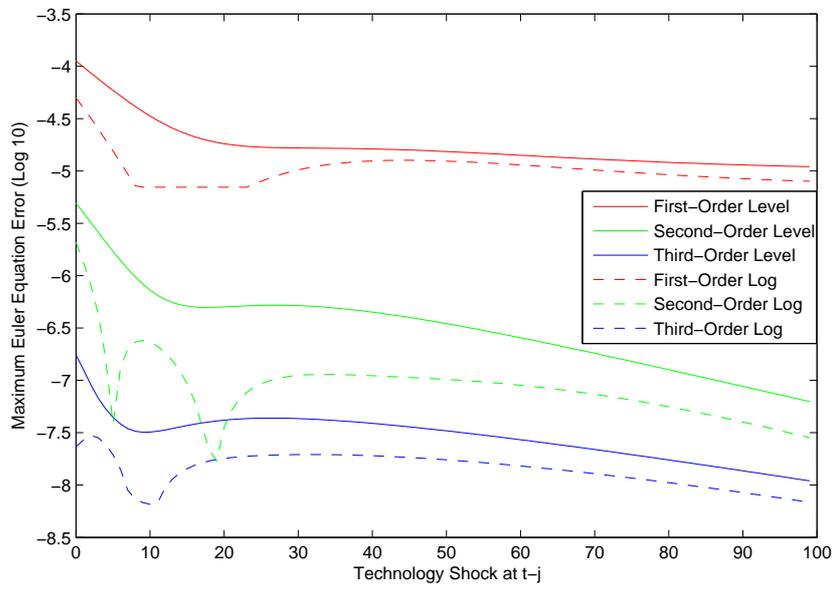


(a) In Levels

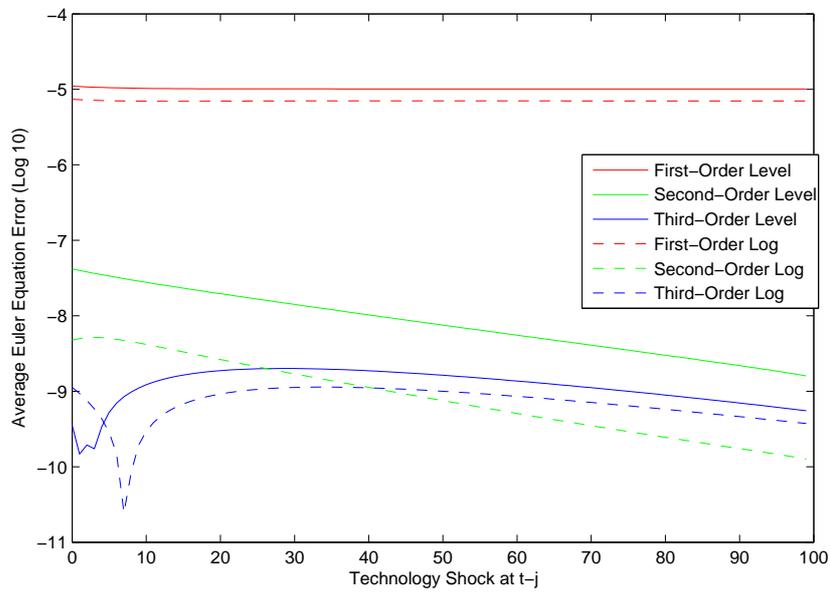


(b) In Logs

Figure 7: 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 8: Maximum and Average Euler Equation Errors, Aruoba, Fernández-Villaverde, and Rubio-Ramírez's (2006) Baseline Case