South African Reserve Bank
Working Paper Series
WP/14/06

Straightforward approximate stochastic equilibria for nonlinear Rational Expectations models

Michael K. Johnston ,Robert G. King ,Denny Lie

Authorised for distribution by Chris Loewald

October 2014

South African Reserve Bank

South African Reserve Bank Working Papers are written by staff members of the South African Reserve Bank and on occasion by consultants under the auspices of the Bank. The papers deal with topical issues and describe preliminary research findings, and develop new analytical or empirical approaches in their analyses. They are solely intended to elicit comments and stimulate debate.

South African Reserve Bank Working Papers are externally refereed.

Information on South African Reserve Bank Working Papers can be found at
http://www.resbank.co.za/Research/ResearchPapers/WorkingPapers/Pages/WorkingPapers-Home.aspx

Enquiries

Head: Research Department
South African Reserve Bank
P O Box 427
Pretoria 0001

Tel. no.: +27 12 313-3911
0861 12 SARB (0861 12 7272)

Non-technical summary

Many theoretical economic models contain expectations of future outcomes as variables in the system. Rational expectations, a concept in which information from the entire model is used to determine expectations of future outcomes, has been the dominant method used in theoretical models for the past half century (since the introduction of the concept by John Muth (1961)).

The most commonly used methods for working with Dynamic Stochastic General Equilibrium (DSGE) models rely on a combination of linear approximation and rational expectations to achieve a solution. Linear approximation is attractive because it is tractable, produces solutions in state-space form which can be easily estimated, and allows questions of solution existence and uniqueness to be easily answered relative to a non-stochastic steady state.

Despite the advantages in terms of tractability and feasibility, linear approximation methods have a number of disadvantages. For example, they impose that agents' decision rules do not change in response to changes in the variance of exogenous shocks, that the economy responds the same to shocks regardless of whether the economy is in a recession or expansion, that the response of the economy is always linearly proportional to the size of the shocks, and that positive and negative shocks have symmetric effects. These assumptions preclude answers about stochastic volatility or changes in uncertainty, asymmetric behaviour over the business cycle, and many other interesting questions.

We present an analysis which shows that non-linear solutions can be almost as straightforward as linear approximation methods for a wide class of models. Just as linear approximations can be regarded as first-order Taylor series differential approximations, our method explains how to use second- and third-order Taylor series differential approximations to account for model non-linearities. The analysis uses nothing more than textbook differential matrix calculus (for example, Magnus and Neudecker (2002)) to achieve the result.

Questions of existence and uniqueness are the same as in the first-order, and the non-linear solution has a linear state space representation. We explain how to use existing linear solution methods multiple times to achieve a non-linear approximate solution. The linear state-space representation means that the non-linear solution can be used for simulation and impulse response analysis nearly as easily as a linear solution. Finally, we conclude with a numerical example in which two models — one with firm heterogeneity and one without — are identical in linear approximations, but which have important differences that show up when non-linear solution methods are used.

Straightforward approximate stochastic equilibria for nonlinear Rational Expectations models*

Michael K. Johnston† Robert G. King‡ Denny Lie§

First version: September 2009

This version: September 2014

Abstract

We present a new approach to the approximation of equilibrium solutions to nonlinear rational expectations models that applies to any order of approximation. The approach relies on a particular version of Taylor series approximations—the *differential* version—and on a scalar perturbation of the support of the entire history of shocks. The resulting solution for any order can always be directly cast in a linear state-space form, permitting the solution to be used for many practical applications such as forecasting, estimation, and computing impulse responses. Using the approach, we show that there cannot be multiple solutions in any order of approximation if the associated first-order approximate solution is determinate. Our approach can be used simply to verify key propositions of the earlier literature, to extend its range of applications, and to resolve puzzles left by it. While the paper only provides an explicit solution up to a third-order approximation, extensions to any higher order approximations are straightforward.

JEL Classification: C63, C68, E17, E37

Keywords: Solution methods; higher order approximations; perturbation, differential Taylor series approximation; nonlinear rational expectations models; pruning; DSGE

*The authors thank participants at the 2009 Dynare conference and seminar participants at Boston University, Bank of Canada, Federal Reserve Bank of Richmond, Federal Reserve Bank of Boston, University of Sydney, and University of Melbourne for helpful discussions, comments and contributions.

†South African Reserve Bank.

‡Boston University and the National Bureau of Economic Research

§The University of Sydney — corresponding author (email: denny.lie@sydney.edu.au).

1 Introduction

Since the emergence of the real business cycle literature it has become commonplace to obtain solutions to nonlinear rational expectations models using log-linear or linear (first-order) approximations. Various studies on first-order approximation methods have shown that first-order approximate solutions have many appealing properties, e.g.: (i) local uniqueness and stability can be easily characterized, (ii) the method is simple to apply and is computationally scalable, especially compared to fully non-linear methods, and (iii) the resulting solution can be cast in a state-space form, which permits practical applications such as estimation, forecasting, and impulse response analysis.[1] Furthermore, for many standard rational expectations models such as real business cycle (RBC) and standard New Keynesian models, first-order approximations have been shown to approximate the true nonlinear solutions quite well. In spite of this, there exists a large number of modern macroeconomic models where it is necessary to move beyond linear approximations. Models with state-dependent decisions, e.g. in pricing or investment, are one such example. First-order approximations are ill-suited for these models since the linearized solutions render the effect of a given exogenous shock independent of the state of the economy, partly nullifying the models' state-dependence feature. And since large shocks have proportionally the same effect as small shocks in a linearized solution, applying the method to models in which risk and uncertainty matter a great deal, e.g. asset pricing models and models with financial frictions, may yield inaccurate solutions. In addition, it is now well known that using first-order approximations for evaluating welfare and comparing alternative policies may yield incorrect conclusions.[2] For these reasons a number of recent studies have proposed solution algorithms for second- or higher-order approximations and championed their use to solve rational expectations models — among others, Schmitt-Grohe and Uribe (2004), Swanson, Anderson, and Levin (2005), Lombardo and Sutherland (2007) and Kim, Kim, Schaumburg, and Sims (2008).[3]

Despite this recent advance in higher-order approximations there remains unresolved issues regarding the properties of the solutions and their suitability for various analyses. First, unlike in the first order, the stability and the uniqueness of higher-order approximate solutions remain an open question. The presence of quadratic terms in second-order approximations, for example,

[1] A non-exhaustive list of studies on first-order approximations: Blanchard and Khan (1980), Anderson and Moore (1984), King and Watson (1998, 2002), Sims (2001), and Klein (2004).

[2] See, e.g., Kim and Kim (2003).

[3] With the exception of Swanson, Anderson, and Levin, all other studies only concern up to second-order approximations.

opens the possibility for multiple solutions, even when uniqueness is achieved under a first-order approximation. Another important issue regards the form of the solutions. As discussed in Kim, et.al. (2008), conventional second-order Taylor-series approximations like in Schmitt-Grohe and Uribe (2004) cannot be directly used to perform standard analyses such as impulse response due to the presence of "garbage" higher order terms — for this reason it is necessary to first "prune" these extraneous terms from the approximate solutions before performing various analyses. This pruning mechanism leads to a solution in a state-space form, which is useful for forecasting, estimation, and simulation exercise. Although it leads to a usable solution form, a fundamental question of interest is whether pruning is legitimate, in a sense that it reflects a true second-order approximate solution based on the Taylor series. Third, with the exception of Swanson, Anderson, and Levin (2008) and Lan and Meyer-Gohde (2013), all existing studies focus only on second-order approximations. Given the highly nonlinear and stochastic nature of modern macroeconomic models, it is of interest to develop a unified solution algorithm for third- and higher-order approximations and to understand their stability and uniqueness properties. There is also the practical issue of how to perform a pruning-like mechanism to these higher-order approximate solutions and obtain consistent state-space representations. Finally, since higher-order systems get large fast, even for a small model, it is important to have a solution algorithm that is both fast and efficient.

In this paper we describe a new approach to computing approximate stochastic equibrilia for nonlinear rational expectations that applies to any order of approximation. It is a straightforward approach that relies on a particular version of Taylor series approximations—the *differential* version—and on a scalar perturbation of the support of the history of shocks. The approach is based on the simple idea that both the approximate model equations and the approximate solution system are stochastic difference equations which are intimately linked. For any order of approximation, we deduce restrictions directly on the stochastic differentials of approximate equilibrium stochastic processes; and we do so in a relatively straightforward manner, using little more than differential calculus and some linear algebra.

Our approach has several advantages over existing approaches and provides answers to unresolved issues on higher-order approximations described above. First, our differential approach naturally leads to a solution in a state-space form for *any order* of approximation — the state-space form is linear, but the effect of nonlinearity is preserved through the interactions among variables and shocks. Moreover, our approach provides an *endogenous* pruning-like mechanism, which indicates that the *exogenous* pruning mechanism performed, for example, in Kim, et.al.

(2008) is, in fact, unnecessary. Second, we show that stability and uniqueness issues are settled in the first order. As long as the first-order approximate solution is stable and unique, so is the solution to any higher-order approximation. Third, as in Lombardo and Sutherland and Kim, et.al., our solution algorithm utilizes familiar linear rational expectations techniques to succesively solve for higher-order approximations. As there is already a wealth of computational resources on linear approximation techniques, this leads to faster and more efficient computation. Fourth, the approach is general, in a sense that we do not need to take a stand on whether to impose restrictions on variables or restrictions on equations. That is, the approach can be used either in an algorithm where a researcher has to specify which variables are predetermined and jump variables, e.g. as in Blanchard and Khan (1980), King and Watson (1998, 2002) and Klein (2000), or in Sims' (2001) algorithm where the only necessary information is on which equations are subject to endogenous forecast errors. Fifth, our approach can be used simply to verify key propositions of the earlier studies, to extend its range of application, and to resolve puzzles left by it. In terms of propositions, there is a direct implication of the core theorem (Theorem 1) in Schmitt-Grohe and Uribe (2004), which indicates that second order approximations originating from the simultaneous stretching of shocks have no effects on first or second order approximation coefficients, except for a constant term adjustment in the second order case which depends on the extent of the variance of the shocks. We extend this theorem and show that in third-order approximate solution the effect of shock uncertainty is time-varying and depends on the state of the economy. Finally, while this paper only provides explicit solutions up to a third-order approximation, extensions to any higher-order are direct and straightforward. And importantly, our approach is unified in a sense that for any order of approximation the general form of the restrictions is identical and thus, the same core algorithm can be used to obtain the solutions — what is different for different orders are the elements of the driving processes.

The rest of the paper is organized as follows. Section 2 describes the general forms of the model to be approximated and the approximate equilibrium solutions. Section 3 contains the details of our differential approach and how to compute the solutions to any order of approximation. We provide explicit state-space form solutions up to a third-order approximation. Properties of the approximate solutions and links to existing approaches in the literature are discussed in Section 4. Section 5 applies the method to two nonlinear rational expectations models. Section 6 concludes.

1.1 Comparison to existing studies

Table 1 provides a summary of comparison between our approach and several existing studies in the literature, namely Schmitt-Grohe and Uribe (SGU, 2004), Swanson, Anderson, and Levin (SAL, 2005), Lombardo and Sutherland (LS, 2007), Kim, et.al. (KKSS, 2008), and Lan and Meyer-Gohde (LM, 2013). We differentiate each study into several main characteristics, listed in the rows of Table 1. The last column describes the characteristics of our approach.

Looking first at the order of approximations, SGU, LS, and KKSS each describes a method that is valid up to a second-order approximation. The perturbation AIM method of SAL is valid up to a 7th-order of approximation. As mentioned above, we provide an approach that is valid for any order of approximation. In terms of the type of Taylor series approximation, all four studies utilize the conventional Taylor series approximation, in contrast to our differential version of the Taylor series. The differential approach is central to our method as it allows us to naturally and consistently obtain solutions in a state-space form without the use of "exogenous," somewhat ad-hoc, pruning mechanism like in LS and KKSS. In addition, it leads to clean and straightforward stability and uniqueness conditions. While LS also discuss similar conditions for second-order approximate solutions, we extend the result and show that the stability and uniqueness conditions are the same irrespective of the order of approximation. On the type of restrictions being imposed on rational expectations models in order to obtain solutions, unlike the existing studies where restrictions are imposed either on equations (SAL, KKSS, and LM) or variables (SGU and LS), our approach can be applied to both types of restrictions. In terms of the computation method, we follow the direct method akin to the canonical-variable method in Blanchard and Khan (1980), which is also used in LS and KKSS. This is in contrast to the undetermined coefficients-like approach used in SGU, SAL, and LM. The direct method permits faster computation since it does not rely on approximating the policy functions and hence, does away from the implicit function theorem. Finally, as in LS and KKSS, our solution algorithm utilizes familiar techniques from the linear-approximation literature.

Our approximation approach is closest to that in Lan and Meyer-Gohde (2013), which, like ours, is based on the idea that a solution to rational expectations model is a function of the entire infinite history of shocks. Their solution method is also valid for any order of approximation and similarly, the stability of the solution to any order of approximation is shown to only depend on the stability of the first-order approximate solution. However, the execution in terms of finding the approximate

solution is quite different. LM posit that the unknown policy function is a function of the entire history of shocks and derive restrictions directly based on the Taylor series in these infinite series of shocks (a "Volterra" series), expanded around the certainty point. These restrictions, for a given order of approximation, can then be used to solve for the unknown partial derivatives (the coefficients of the policy function) using the undetermined coefficients-like method used, for example, in SGU (2004). This approach of finding the solution leads to an approximate policy function that is a nonlinear function of the infinite history of shocks (for a second-order approximation or higher), i.e. a nonlinear moving average process. We show in Appendix C, however, that approximating the entire history of shocks is equivalent to taking a standard Taylor-series approximation via differential. This novel observation means that we do not need to deal with these infinite number of shock terms, making our approach highly tractable. In addition, it allows us to operate with standard state-space methods widely used in the literature.[4] While one can use the solution form in Lan and Meyer-Gohde to properly perform impulse response analysis, it is less useful for other purposes, e.g. for estimation, due to the infinite number of terms related to the shocks. This will not be an issue in our approach since our use of state-space methods allows the solution to be directly cast in a standard linear state-space form, permitting the potential use of standard Kalman filter for estimation purpose (see Kollman, 2013).

2 Model and approximate solution: The general forms

We first stipulate the general class of models we are interested in and the general form of an *n-th* order solution, including the general structures of the restriction and the solution for each of its elements.

The models of interest We seek an approximate solution to the following class of rational expectations models:

$$0 = g(z_{t+1}, z_t, \eta_{t+1}) \tag{1}$$

$$0 = E_t f(z_{t+1}, z_t, \eta_{t+1}) \ , \tag{2}$$

where the the first block contains n_g equations which are *stochastic* as of t and the second block contains n_f equations that are *exact* as of t (including expectational restrictions). Both vector-

[4] Except for Lan and Meyer-Gohde (2013), all the studies mentioned in Table 1 use state-space methods.

valued function g and f are possibly, and in general, nonlinear. There are n_z variables in the system, which are contained in the vector z_t, and there are as many equations as variables, i.e. $n_z = n_g + n_f$. $\eta_{t+1} = \sigma\varepsilon_{t+1}$ are the vector of i.i.d. exogenous shocks or disturbances with zero mean, stretched by the perturbation parameter, σ. Some of the variables in z_t are jump variables, while some are predetermined. The number of exogenous shocks in the system is n_ε. We leave the distribution of the exogenous shocks unspecified.

Taking the conditional expectations, the two block of equations in (1) and (2) can also be compactly written as

$$0 = E_t\Gamma(z_{t+1}, z_t, \eta_{t+1}) \ . \tag{3}$$

We will use the general forms (1)-(2) and (3) interchangeably in subsequent expositions of our approach. Note that we do not restrict the way exogenous shocks affect the variables and hence, they are allowed to enter nonlinearly. In subsequent discussions, we do not take a stand on whether restrictions are imposed on variables or equations for the existence of a solution — our approach can be used irrespective of the choice. Appendix A contains a computational application of our approach where the restrictions are imposed on equations as in Sims (2001).

The *n-th* order solution The approximate solution to the system (3) (or equivalently, (1) and (2)) to the *n-th* order is of the form

$$z_t \simeq \bar{z} + y_t^{(1)} + \frac{1}{2}y_t^{(2)} + \frac{1}{6}y_t^{(3)} + ... + \frac{1}{n!}y_t^{(n)} \ , \tag{4}$$

where $\bar{z}$ is the vector of deterministic steady-state values of z_t, satisfying

$$0 = \Gamma(\bar{z}, \bar{z}, 0) \ , \tag{5}$$

and $y_t^{(i)}$, $i = 1, ..., n$, denotes the i-th order element of the solution (based on the Taylor series). Obtaining the approximate solution in (4) is tantamount to obtaining the solutions to $y_t^{(i)}$, consistent with restrictions imposed by the nonlinear system of equations in (3). If the the solution to each $y_t^{(i)}$ is obtained in a state-space form, the *n-th* order solution is also in a state-space form. Our approach to finding the solution is to solve for $y_t^{(i)}$ successively, starting from $i = 1$.[5] As will be clear later, $y_t^{(i)}$ is in fact the solution to the restrictions imposed by the *i-th differential* of (3).

[5] This successive computation approach is first advocated by Lombardo and Sutherland (2007).

General forms of approximate-model restrictions and solutions For each $y_t^{(i)}$, $i = 1, ..., n$, in the *n-th* order approximate solution, we will show that the restrictions for the solution are linear in form and given by

$$A_1 y_{t+1}^{(i)} = B_1 y_t^{(i)} + C_1^{(i)} x_{t+1}^{(i)} \tag{6}$$

$$A_2 E_t y_{t+1}^{(i)} = B_2 y_t^{(i)} + C_2^{(i)} E_t x_{t+1}^{(i)} \ , \tag{7}$$

with the driving process evolves according to[6]

$$x_t^{(i)} = \gamma^{(i)} + \gamma_\varsigma^{(i)} \varsigma_{t-1}^{(i)} + \gamma_v^{(i)} v_t^{(i)} \tag{8}$$

$$\varsigma_t^{(i)} = \phi^{(i)} + \phi_\varsigma^{(i)} \varsigma_{t-1}^{(i)} + \phi_v^{(i)} v_t^{(i)} \tag{9}$$

The driving variables $x_t^{(i)}$ are treated as exogenous for each i, but are not necessarily exogenous for the whole system (3). Notice that the $n_g \times n_z$ matrices A_1 and B_1 and the $n_f \times n_z$ matrices A_2 and B_2 have no superscript i and hence, are identical for all i. The vector of innovations, $v_t^{(i)}$, has zero mean and variance $\Omega_v^{(i)}$.

The restrictions above are familiar restrictions from the literature on linear approximation methods. As such, the solution to $y_t^{(i)}$ can be obtained using any linear approximation method and can be written in a linear state-space form,

$$y_t^{(i)} = \theta_y^{(i)} + \theta_{yk}^{(i)} k_t^{(i)} + \theta_{y\varsigma}^{(i)} \varsigma_t^{(i)} \tag{10}$$

$$k_t^{(i)} = \theta_k^{(i)} + \theta_{kk}^{(i)} k_{t-1}^{(i)} + \theta_{k\varsigma}^{(i)} \varsigma_{t-1}^{(i)} + \theta_{kv}^{(i)} v_t^{(i)} \tag{11}$$

for some states $k_t^{(i)}$.[7] It follows then the *n-th* order approximate solution is also available in a state-space form — moreover, the resulting state-space form is linear. Despite the linear form, nonlinearity is preserved for a second-order approximation and higher in the definition of variables and shocks in $\varsigma_t^{(i)}$, $k_t^{(i)}$, and $v_t^{(i)}$. At this point we note that in order for the solutions in (10) and (11) to exist, we assume standard regularity conditions from the linear approximation literature, see. e.g. the conditions in King and Watson (1998). We also impose the "no unit root condition"

[6] Alternatively, taking the time-t conditional expectations, we can write the restrictions as $AE_t y_{t+1}^{(i)} = By_t^{(i)} + C^{(i)} E_t x_{t+1}^{(i)}$, with $A = \left[\begin{array}{cc} A_1^T & A_2^T \end{array} \right]^T$, $B = \left[\begin{array}{cc} B_1^T & B_2^T \end{array} \right]^T$, and $C^{(i)} = \left[\begin{array}{cc} C_1^{(i)T} & C_2^{(i)T} \end{array} \right]^T$.

[7] In this general state-space form solution, we allow for $y_t^{(i)}$ and $k_t^{(i)}$ to depend on constants, $\theta_y^{(i)}$ and $\theta_k^{(i)}$ — but these could be removed by appropriately redefining the definition of the endogenous and state variables.

for ease of exposition of our approach.

3 Our approach: Taylor Series approximation via differential

An equilibrium solution to a rational expectations model in (3) is a stochastic process which satisfies the model's equations, including those with expectations, at all dates and from specified initial conditions. This stochastic process can alternatively be represented as a function of the entire history of shocks,

$$z_t = Z(\{\eta_{t-j}\}_{j=0}^{\infty}) \ , \tag{12}$$

where $\eta_{t-j} = \sigma\varepsilon_{t-j}$. A central difficulty is that this true stochastic process—or the true solution function, Z—is not known and is only implicit in model restrictions. Our solution approach is to find an *approximate stochastic process* by taking a Taylor series approximation to the unknown true solution (12), around the deterministic stationary point implicit in (5). This approximation can be viewed as a Taylor series approximation in the perturbation parameter σ at every date, around the deterministic point, $\sigma = 0$. Hence, treating an underlying distribution function for $\eta_t = \sigma\varepsilon_t$ as fixed at every date, we can stretch this distribution and write the Taylor formula as

$$Z(\{\eta_{t-j}\}_{j=0}^{\infty}) = z + \sum_{i=1}^{\infty} (i!)^{-1} \frac{\partial^i Z(\{\sigma\varepsilon_{t-j}\}_{j=0}^{\infty})}{(\partial\sigma)^i}\sigma^i$$

As is apparent from the above formula, the coefficient terms rapidly become complicated functions of the history of shocks, even for the second-order approximation ($i = 2$). However, while these terms express the nature of the perturbation approximation that we are undertaking, there is no need to actually calculate the relevant components of, say, $\frac{\partial^2 Z(\{\sigma\varepsilon_{t-j}\}_{j=0}^{\infty})}{(\partial\sigma)^2}\sigma^2$. As shown in Appendix C, the Taylor formula above is equivalent to taking a Taylor series approximation via *differential*:

$$z_t = Z(\{\eta_{t-j}\}_{j=0}^{\infty}) \approx \bar{z} + dz_t + \frac{1}{2}d^2 z_t + \frac{1}{6}d^3 z_t + ... \tag{13}$$

Hence, to obtain the approximation to the *n-th* order all we need to do is to calculate the solution for the differentials, dz_t, ..., $d^n z_t$. These differential solutions have to be consistent with the restrictions imposed by the model's equations in (1) and (2).

We note that our approximation strategy differs from the standard approach in the literature in two respects. First, it is conventional to find an equilibrium solution by first positing unknown

policy functions,

$$\begin{aligned} z_t &= z(s_t;\sigma) \\ s_t &= m(s_{t-1},\eta_t;\sigma)\ , \end{aligned}$$

for some states s_t, before proceeding to find the approximation to such policy functions, e.g. as in Schmitt-Grohe and Uribe (2006). This is in contrast to our approach of finding the approximation to the stochastic process directly. Even if there is a one-to-one relationship between the approximate solutions arising from the two approaches, our approach offers a potentially faster solution as it does not need to rely on implicit function theorem, which is needed in the policy functions approximation approach. Moreover, our direct approach also makes the stability and the uniqueness conditions transparent. The second difference is that we use the differential version of the Taylor series approximation (13) instead of the conventional one. In Section 4, we contrast the two versions of the Taylor series approximation and show why the differential version is preferable.

3.1 A simple scalar example

To illustrate our approximation strategy let's first consider a simple scalar system

$$\begin{aligned} z_t &= f(s_t) \\ s_t &= m(s_{t-1},\eta_t), \end{aligned}$$

where the scalar functions f and m are fully known. The innovation η_t has zero mean and variance Ω_η. This system is descriptive, in a sense that there is no need to actually solve for the solution as the system itself is already a true solution. We are interested in the *approximate* solutions to the variable z_t, around the deterministic stationary point satisfying $\bar{z} = f(\bar{s})$ and $\bar{s} = m(\bar{s},0)$. Note that the system above belongs to the general class of models described in (3). For example, by defining $\tilde{z}_t = \begin{bmatrix} z_t & s_t \end{bmatrix}^T$ and shifting the dates one-period forward the system above can be written as

$$0 = \tilde{\Gamma}(\tilde{z}_{t+1},\tilde{z}_t,\eta_{t+1})\ ,$$

with $\tilde{\Gamma}$ now is a vector-valued function consisting of the functions f and m.

Consider the second-order Taylor series approximation to the above system — based on our

differential approach this is given by

$$z_t \simeq \bar{z} + dz_t + \frac{1}{2} d^2 z_t \ .$$

All we have to do is to find the solutions to the differentials, dz_t and $d^2 z_t$. Defining $f_s(.)$ and $f_{ss}(.)$ as the first and the second derivatives of the function f with respect to its argument (s_t), we have

$$\begin{aligned} dz_t &= [f_s(s_t)]\, ds_t \\ d^2 z_t &= [f_s(s_t)]\, d^2 s_t + [f_{ss}(s_t)]\, (ds_t ds_t) \ . \end{aligned}$$

The differentials ds_t and $d^2 s_t$ are in turn given by

$$\begin{aligned} ds_t &= [m_s(s_{t-1}, \eta_t)]\, ds_{t-1} + [m_\eta(s_{t-1}, \eta_t)]\, d\eta_t \\ d^2 s_t &= [m_s(s_{t-1}, \eta_t)]\, d^2 s_{t-1} + [m_{ss}(s_{t-1}, \eta_t)]\, (ds_{t-1} ds_{t-1}) \\ &\quad + [2m_{s\eta}(s_{t-1}, \eta_t)]\, (d\eta_t ds_{t-1}) + [m_{\eta\eta}(s_{t-1}, \eta_t)]\, (d\eta_t d\eta_t) \ , \end{aligned}$$

where $m_s(s_{t-1}, \eta_t)$ and $m_\eta(s_{t-1}, \eta_t)$ are the first partial derivatives of the function $m(.)$ with respect to its first argument and the second argument, respectively, and $m_{ss}(s_{t-1}, \eta_t)$, $m_{\eta\eta}(.)$, and $m_{s\eta}(.)$ are the second partial and cross derivatives. Notice that $d^2 s_t$ does not contain the term $d^2 \eta_t$ since $\eta_t = \sigma \varepsilon_t$, and hence, $d^j \eta_t = 0$ for all $j > 1$. The nature of perturbation we are undertaking also implies $d\eta_t = d(\sigma \varepsilon_{t+1}) = \varepsilon_{t+1}(\sigma - 0) = \eta_t$.

Evaluating all the coefficients at the stationary point, it follows then the second-order approximate solution is given by

$$\begin{aligned} z_t &\simeq \bar{z} + dz_t + \frac{1}{2} d^2 z_t \\ &= \bar{z} + [f_s]\, ds_t + \frac{1}{2} \begin{bmatrix} f_s & f_{ss} \end{bmatrix} \begin{bmatrix} d^2 s_t \\ ds_t ds_t \end{bmatrix} \ , \end{aligned} \tag{14}$$

with the full state vector evolves according too

$$\begin{bmatrix} ds_t \\ d^2 s_t \\ ds_t ds_t \end{bmatrix} = \begin{bmatrix} m_s & 0 & 0 \\ 0 & m_s & m_{ss} \\ 0 & 0 & m_s^2 \end{bmatrix} \begin{bmatrix} ds_{t-1} \\ d^2 s_{t-1} \\ ds_{t-1} ds_{t-1} \end{bmatrix} + \begin{bmatrix} m_\eta & 0 & 0 \\ 0 & 2m_{s\eta} & m_{\eta\eta} \\ 0 & 2m_s m_\eta & m_\eta^2 \end{bmatrix} \begin{bmatrix} \eta_t \\ \eta_t ds_{t-1} \\ \eta_t \eta_t \end{bmatrix} \ . \tag{15}$$

The above state-space form can then be used to perform various analyses such as forecasting and calculating impulse responses. In the first-order solution, $z_t \simeq \bar{z} + dz_t$, only the evolution of ds_t in the first line in (15) matters. For second order, the state vector consists of ds_t, d^2s_t, and $ds_t ds_t$. Note that here, while $ds_t = s_t - \bar{s}$ can be interpreted as the (level) deviation of s_t from its stationary point, the interpretation of the second-order differential, d^2s_t, is not as straightforward. For now, we will just refer $d^n s_t$, $n > 1$, as the the *n-th* differential of the state(s), s_t. We show later that these higher-order differential terms are directly related to the higher-order terms in the conventional Taylor series approximation. The evolution of the last element of the full state vector, $ds_t ds_t$, is derived by taking the square of the evolution of ds_t, i.e. the state evolution in the first-order differential. Unlike in the first-order, the second-order approximate solution makes the effect of exogenous shocks to be potentially state-dependent and conditionally heteroskedastic (the $\eta_t ds_{t-1}$ term) and size- and variance-dependent (the $\eta_t \eta_t$ term). There is also a nonzero mean to d^2s_t and $ds_t ds_t$, so that the second-order approximation no longer displays certainty equivalence in the mean. These properties will translate to the general class of models in (3) as well.

While simple, the above example illustrates many of the appealing properties of our approximation strategy mentioned in the introduction. For one, extensions to higher order approximations are direct. For example in the third-order approximation, all we need to do is to calculate the expression for d^3z_t, which depends on d^3s_t, which in turns depends on various third-order terms such as $ds_{t-1}ds_{t-1}ds_{t-1}$, $d^2s_t ds_t$, and $d\eta_t d\eta_t ds_{t-1}$ — and this third-order approximate solution can also be cast in a state-space form. It is also straightforward to determine the stability of the approximate solutions. Notice that the first-order solution is stable if $|m_s| < 1$. This condition also guarantees the stability of the second-order approximate solution as the stability of d^2s_t also depends on the same coefficient: m_s on d^2s_{t-1} and m_s^2 (which is < 1 if $|m_s| < 1$) on $ds_{t-1}ds_{t-1}$. For the third order, stability will depend on m_s, m_s^2, and m_s^3, which is smaller than one in absolute value if $|m_s| < 1$. Hence, the stability of the approximate solution for *any* order only depends on whether the first-order approximate solution is stable.

3.2 The general class of models

We now turn to applying the approximation strategy to the general class of models described in (1) and (2),

$$\begin{aligned} 0 &= g(z', z, \eta') \\ 0 &= E_t f(z', z, \eta') \ , \end{aligned}$$

where we have dropped time subscripts temporarily and use a prime to represent $t+1$ variables. Since our strategy involves computing the approximate solution successively starting from the lowest order, our presentation mirrors this.

3.2.1 The first-order approximate solution

Total differentiation of the model equations above provides a set of restrictions on the first-order differentials and their equilibrium dynamics:

$$0 = G_{z'} dz' + G_z dz + G_{\eta'} d\eta' \tag{16}$$

$$0 = E_t[F_{z'} dz' + F_z dz + F_{\eta'} d\eta'] \ , \tag{17}$$

where the matrices F_j and G_j, $j \in \{z', z, \eta'\}$, are the partial derivatives with respect to elements of j of the functions f and g, respectively, evaluated at the deterministic stationary point.[8] These differential restrictions can thus be written as

$$A_1 dz' = B_1 dz + C_2^{(1)} \eta' \tag{18}$$

$$A_2 E_t dz' = B_2 dz + C_2^{(1)} E_t \eta' \ , \tag{19}$$

which is of the general form of restrictions depicted in (6) and (7) with matrix coefficients $A_1 = -G_{z'}$, $B_1 = G_z$, $C_2^{(1)} = G_{\eta'}$, $A_2 = -F_{z'}$, $B_2 = F_z$, and $C_2^{(1)} = F_{\eta'}$. The vectors of variables and the driving process in these first-order differentials are $y^{(1)} = dz$ and $x'^{(1)} = d\eta' = \eta'$, respectively. Since η' is just the vector of the model's original i.i.d. exogenous shocks with $E_t \eta' = 0$ its evolution — in

[8] The matrix $F_{z'}$ is the n_f-by-n_z matrix of partial derivatives with respect to z', with the (i,j)th element being the partial derivative of the ith equation with respect to the jth element of z'. The matrix $G_{z'}$ is correspondingly a n_g-by-n_z matrix, with elements defined for equation i and the jth element of z'. The matrices G_z and F_z are defined similarly for elements of z, so are the n_g-by-n_ε matrix $G_{\eta'}$ and the n_f-by-n_ε matrix $F_{\eta'}$ for elements of the shock vector, η'.

terms of the general form in (8) and (8) — is very simple, i.e. there is no additional driving-process term, $\varsigma^{(1)}$, and we simply have $\eta' = I_\eta \eta'$, where I_η is an $n_\varepsilon \times n_\varepsilon$ identity matrix.

The determinacy (stability and uniqueness) condition and the solution The system of restrictions in (18) and (19) can be solved using any of the variety of linear rational expectations solution methods in the literature, e.g. King and Watson (1998) and Klein (2000), or in Sims' (2001). The solution can be written in a recursive, state-space form,

$$\begin{aligned} dz &= \Pi ds \\ ds' &= \Phi_s ds + \Phi_\eta \eta' \ , \end{aligned} \tag{20}$$

with $ds = s - \bar{s}$. The vector of state variables, s, could either consist of the original state variables in the model or linear combinations of the original states.[9] This solution is of the general form depicted in (10) and (11). Note that to obtain the above solution we implicitly assume that a stable solution exists and is unique.[10] For example, if we use the method of King and Watson (1998), the linear difference (differential) system

$$AE_t dz' = Bdz$$

where $A = \begin{bmatrix} A_1^T & A_2^T \end{bmatrix}^T$ and $B = \begin{bmatrix} B_1^T & B_2^T \end{bmatrix}^T$, must satisfy the following conditions.[11] First, there must be a number ϕ such that $|A\phi - B| \neq 0$. Second, the no-unit-root condition further requires that $|A - B| \neq 0$. Third, the relevant generalizations of the Blanchard and Kahn (1980) rank and order conditions must be satisfied, in that there must be as many stable eigenvalues as there are elements of s_t and it must be feasible to associate unstable canonical variables (which include components related to unstable and infinite eigenvalues) to remaining elements of z_t. Given the differential solution above, it follows then the first-order approximate solution is given by (adding back the time subscripts)

$$\begin{aligned} z_t &= \bar{z} + dz_t \\ &= \bar{z} + \theta_{z\xi}^{(1)} \xi_t^{(1)} \end{aligned} \tag{21}$$

[9] When restrictions are made on the variables like in King and Watson (1998) and Klein (2000), s_t is the vector of original state variables. Under Sims (2001) where restrictions are determined from the equations directly, s_t, in general, is a linear combination of the original states.

[10] It is generally possible to characterize the whole set of solutions under multiple equilibria (indeterminacy) , see e.g. Lubik and Schorfheide (2004). We leave this for future research.

[11] X^T represents the transpose of matrix X.

with

$$\xi_{t+1}^{(1)} = \theta_{\xi\xi}^{(1)}\xi_t^{(1)} + \theta_{\xi v}^{(1)}v_{t+1}^{(1)} \ , \tag{22}$$

where $\theta_{z\xi}^{(1)} = \Pi$, $\theta_{\xi\xi}^{(1)} = \Phi_s$, $\theta_{\xi v}^{(1)} = \Phi_\eta$, $\xi_t^{(1)} = ds_t$, and $v_t^{(1)} = \eta_t$.

3.2.2 The second-order approximate solution

Total differentiation of the first-order differential restrictions in (16) and (17) gives us the relevant restrictions on the second-order differentials.

$$0 = \Gamma_{z'}E_t d^2z' + \Gamma_z d^2z + E_t\delta'^{(2)} \ , \tag{23}$$

where

$$\Gamma_{z'} = \begin{bmatrix} G_{z'} \\ F_{z'} \end{bmatrix} \ , \Gamma_z = \begin{bmatrix} G_z \\ F_z \end{bmatrix}$$

and

$$\begin{aligned} \delta'^{(2)} \quad = \quad & \Gamma_{z'z'}(dz' \otimes dz') + 2\Gamma_{z'z}(dz' \otimes dz) + 2\Gamma_{z'\eta'}(dz' \otimes d\eta') \\ & + \Gamma_{zz}(dz \otimes dz) + 2\Gamma_{z\eta'}(dz \otimes d\eta') + \Gamma_{\eta'\eta'}(d\eta' \otimes d\eta') \ . \end{aligned} \tag{24}$$

In writing the restrictions above, we use the expectational form of the system of equations (3), instead of separating them into f-type and g-type equations, to save notations. The vector $\delta'^{(2)}$ can be interpreted as the driving process relevant to the second-order differentials, d^2z. The matrices Γ_{jk}, $j,k \in \{z', z, \eta'\}$, are Jacobian matrices consisting of partial derivatives of matrix Γ_j, with respect to each element of k. That is, $\Gamma_{jk}^h \equiv vec\left(\frac{\partial}{\partial k}\left[\Gamma_j^h\right]^T\right)^T$ for rows $h = 1, ..., n_z$.[12]

Several comments on the differential restrictions in (23) are in order. First, notice that the coefficient matrices $\Gamma_{z'}$ and Γ_z are identical to the matrices attached to $E_t dz'$ and dz' in the expectational form of the first-order differential restrictions in (16) and (17).[13] What is different between the first- and second-differential restrictions, aside from the differential order, is the element of the driving process. Instead of consisting of the stochastic shocks in the nonlinear model, the driving process in the second order, $\delta'^{(2)}$, is now a function of various cross (tensor) products of the elements of the first-order differentials and driving process. This second-order driving process

[12] $vec(.)$ is the matrix vectorization operator. If A is an $m \times n$ matrix and a_i is its i-th column, $vec(A) = \begin{bmatrix} a_1^T & a_2^T & ... & a_n^T \end{bmatrix}^T$. The dimension of $vec(A)$ is $mn \times 1$.

[13] In fully-expectational form, the first-order differential restrictions are $0 = \Gamma_{z'}E_t dz' + \Gamma_z dz + \Gamma_{\eta'}E_t\eta'$.

therefore, unlike the driving process in the first order (η'), is endogenous from the perspective of the model. However, since we have already known (and have computed) the first-order differential solutions, $\delta'^{(2)}$ can be treated as *exogenous* for the purpose of computing the solution to the second-order differentials, d^2z.[14] These observations have an important implication: The restrictions in (23), although they are nonlinear in nature, can now be cast in a linear form stipulated in (6) and (7). Despite this linear form, nonlinearity is still preserved, but now is present in the definition of the elements of the driving process, i.e. the cross products of the lower-order differential elements. These restrictions can therefore be solved using any standard linear rational-expectations solution method — an observation first pointed out by Lombardo and Sutherland (2007).

The discussion above also applies to any differential restrictions higher than second order. That is, the restrictions will have the same general linear form, but the elements of the driving process will be different for different orders. The driving process to any i-th order differential restrictions will consist of various cross products of the elements of all lower-order differential solutions and their driving processes, hence preserving the nonlinear nature of the model. And given that we have computed these lower-order solutions prior to computing the solution to the i-th order differentials of interest, the driving process can be treated as if they are exogenous. Any standard linear rational-expectations method can thus be applied to obtain the solution to any i-th differentials, provided that the lower-order differential solution(s) have been computed. Finally, since matrices $\Gamma_{z'}$ and Γ_z attached to $E_t d^i z'$ and $d^i z$ are identical for any i, it follows that we do not have worry about the determinacy condition. As long as the first-order differential solution is stable and unique, so is the solution to any order of approximation.

The evolution of the driving process To obtain the solution to d^2z we first need to obtain the evolution of the driving process, $\delta'^{(2)}$. As elaborated in Appendix B, we can write various cross products in (24) directly as a function of the first-order state-variables and driving process using the first-order differential solution in (20). This process yields (adding back the time subscripts)

$$\delta^{(2)}_{t+1} = C^{(2)} x^{(2)}_{t+1} \ ,$$

[14] In fully-expectational form, the first-order differential restrictions are $0 = \Gamma_{z'} E_t dz' + \Gamma_z dz + \Gamma_{\eta'} E_t \eta'$.

with $x_{t+1}^{(2)}$ defined by

$$x_{t+1}^{(2)} = \begin{bmatrix} vech(ds_t ds_t^T) \\ vech(\eta_{t+1}\eta_{t+1}^T) \\ vec(ds_t \eta_{t+1}^T) \end{bmatrix} . \tag{25}$$

We leave the detail of the matrix coefficient $C^{(2)}$ in Appendix B. Here, we just mention that $C^{(2)}$ consists of various cross products involving the coefficients in the first-order differential solution and the Jacobian matrices Γ_{jk}, $j,k \in \{z', z, \eta'\}$. The vector of the second-order driving process, $x_{t+1}^{(2)}$, consists of various cross products of the first-order state variables and driving process. In (25), $vec(.)$ is the standard matrix vectorization operator and $vech(.)$ is a vectorization operator that chooses only unique elements of a symmetric matrix.[15,16]

From (25), it follows then we can separate the elements of $x_t^{(2)}$ between those that are subject to innovations and those that are not — the latter ones become the additional ("pseudo") states relevant for the second-order differentials. Defining these additional states and the innovations as[17]

$$\begin{aligned} \varsigma_t^{(2)} &= vech(ds_t ds_t^T - \Omega_{ss}) \\ v_t^{(2)} &= \begin{bmatrix} vech(\eta_t\eta_t^T - \Omega_{\eta\eta}) \\ vec(ds_{t-1}\eta_t^T) \end{bmatrix} , \end{aligned}$$

we have

$$x_t^{(2)} = \begin{bmatrix} vech(ds_{t-1} ds_{t-1}^T) \\ vech(\eta_t\eta_t^T) \\ vec(ds_{t-1}\eta_t^T) \end{bmatrix} = \gamma^{(2)} + \gamma_\varsigma^{(2)}\varsigma_{t-1}^{(2)} + \gamma_v^{(2)} v_t^{(2)} \tag{26}$$

We leave the details of the coefficients $\gamma^{(2)}$, $\gamma_\varsigma^{(2)}$, and $\gamma_v^{(2)}$ in Appendix B. Ω_{ss} and $\Omega_{\eta\eta}$ are the unconditional expectations of the cross products of the first-order state variables and innovations, respectively — that is, $\Omega_{ss} = E[ds_t ds_t^T]$ and $\Omega_{\eta\eta} = E[\eta_t\eta_t^T]$.[18] Notice that we have chosen to define the additional states and the innovations, $\varsigma_t^{(2)}$ and $v_t^{(2)}$, in terms of deviations from their

[15] For any symmetric $n \times n$ matrix S, $vech(S)$ is a column vector of length $n(n+1)/2$, created through vertical concatenation of the unique elements of S — following Magnus and Neudecker (2002), the unique elements are chosen from the lower-triangular part of the symmetric matrix. The operator $vech(.)$ is related to the standard vectorization operator $vec(.)$ through a duplication matrix. That is, for any symmetric $n \times n$ matrix S there exists a unique $n^2 \times \frac{n(n+1)}{2}$ matrix D_s such that $D_s vech(S) = vec(S)$.

[16] The following properties of Kronecker product is used to obtain $C^{(2)}$ and $x_{t+1}^{(2)}$: (i) for any conformable matrices A, B, C, and D, $(AB \otimes CD) = (A \otimes C)(B \otimes D)$; and (ii) for any two column vectors a and b, $vec(ab^T) = b \otimes a$.

[17] We call these additional states as "pseudo states" as these state variables arise from lower-order differential solutions.

[18] Since η_t is i.i.d., $E[ds_{t-1}\eta_t^T] = 0$.

unconditional expectations, as this "demeaned" system would be most useful in terms of various practical applications such as impulse response analyses. This choice, however, is innocuous for the purpose of obtaining the second-order differential solution, as the "non-demeaned" system would be equally fine as well.[19] Finally, from the evolution of ds_t in (22) the pseudo states can be found to evolve as

$$\varsigma_t^{(2)} = \phi^{(2)} + \phi_\varsigma^{(2)}\varsigma_{t-1}^{(2)} + \phi_v^{(2)} v_t^{(2)} \ , \tag{27}$$

which is of the general form described in (9).[20]

The determinacy condition and the differential solution The differential restrictions in (23) can now be written in expectational linear-difference system,[21]

$$AE_t d^2 z_{t+1} = B d^2 z_t + C^{(2)} E_t x_{t+1}^{(2)} \ ,$$

with $A = -\Gamma_{z'}$ and $B = \Gamma_z$. Since the matrices A and B are identical to those in the first-order differential system in (18) and (19), $d^2 z_t$ is also determinate, i.e. the same determinacy conditions as in the first order hold. Given the evolution of the driving process in (26) and the pseudo states in (27), the solution is obtained as

$$d^2 z_t = \theta_y^{(2)} + \theta_{yk}^{(2)} d^2 s_t + \theta_{y\varsigma}^{(2)} \varsigma_t^{(2)} \tag{28}$$

$$d^2 s_t = \theta_k^{(2)} + \theta_{kk}^{(2)} d^2 s_{t-1} + \theta_{k\varsigma}^{(2)} \varsigma_{t-1}^{(2)} + \theta_{kv}^{(2)} v_t^{(2)} \tag{29}$$

Here, $d^2 s_t$ is the second differentials of the states s_t, identified in the first-order approximate solution. Again, we leave the full details of various coefficients above in Appendix B. But here, we just note that $\theta_{yk}^{(2)} = \theta_{yk}^{(1)} = \Pi$ and $\theta_{kk}^{(2)} = \theta_{kk}^{(1)} = \Phi_s$ due to identical matrices A and B attached to the differentials of interest in the first- and second-order restrictions. This result confirms our prior observation (and the simple example in Section 3.1) that the same determinacy conditions hold for both the first- and second-order differentials.

The state-space solution above can be further modified by expressing $d^2 s_t$ in terms of deviation from their unconditional expectations, $E[d^2 s]$, which can be obtained by from (29).[22] This results

[19] The vector of innovations still needs to be demeaned, as we require $E[v_t^{(i)}] = 0$ for all i.

[20] See Appendix B for the details of the coefficients.

[21] There should be no confusion in this system that some of the equations are non-expectational, i.e. they hold exactly. Also, here $A = \begin{bmatrix} A_1^T & A_2^T \end{bmatrix}^T$ and $B = \begin{bmatrix} B_1^T & B_2^T \end{bmatrix}^T$.

[22] From (29), $E[d^2 s_t] = \theta_k^{(2)} + \theta_{kk}^{(2)} E[d^2 s_{t-1}]$, implying $E[d^2 s] = \left(I - \theta_{kk}^{(2)}\right)^{-1} \theta_k^{(2)}$.

in a fully-demeaned system,

$$d^2 z_t = E[d^2 z] + \theta_{yk}^{(2)} \left(d^2 s_t - E[d^2 s] \right) + \theta_{y\varsigma}^{(2)} \varsigma_t^{(2)} \tag{30}$$

$$d^2 s_t - E[d^2 s] = \theta_{kk}^{(2)} \left(d^2 s_{t-1} - E[d^2 s] \right) + \theta_{k\varsigma}^{(2)} \varsigma_{t-1}^{(2)} + \theta_{kv}^{(2)} v_t^{(2)} \ . \tag{31}$$

$E[d^2 z] = \theta_y^{(2)} + \theta_{yk}^{(2)} E[d^2 s]$ is the unconditional expectation of the second-order differentials.

The second-order approximate solution Imposing our definition of the approximate solution in (4) the second-order approximate solution is

$$\begin{aligned} z_t &\simeq \bar{z} + dz_t + \frac{1}{2} d^2 z_t \\ &= \left\{ \bar{z} + \theta_{z\xi}^{(1)} \xi_t^{(1)} \right\} + \frac{1}{2} \left\{ E[d^2 z] + \theta_{z\xi}^{(2)} \xi_t^{(2)} \right\} \ , \end{aligned} \tag{32}$$

with the full state vector evolves according to

$$\begin{bmatrix} \xi_{t+1}^{(1)} \\ \xi_{t+1}^{(2)} \end{bmatrix} = \begin{bmatrix} \theta_{\xi\xi}^{(1)} & 0 \\ 0 & \theta_{\xi\xi}^{(2)} \end{bmatrix} \begin{bmatrix} \xi_t^{(1)} \\ \xi_t^{(2)} \end{bmatrix} + \begin{bmatrix} \theta_{\xi v}^{(1)} & 0 \\ 0 & \theta_{\xi v}^{(2)} \end{bmatrix} \begin{bmatrix} v_{t+1}^{(1)} \\ v_{t+1}^{(2)} \end{bmatrix} \tag{33}$$

In the above we have collected all the state variables and innovations in the second-order differentials into vectors $\xi_t^{(2)}$ and $v_t^{(2)}$, respectively:

$$\xi_t^{(2)} = \begin{bmatrix} d^2 s_t - E[d^2 s] \\ \varsigma_t^{(2)} \end{bmatrix} \equiv \begin{bmatrix} d^2 s_t - E[d^2 s] \\ vech(ds_t ds_t^T - \Omega_{ss}) \end{bmatrix} \ ,$$

$$v_t^{(2)} = \begin{bmatrix} vech(\eta_t \eta_t^T - \Omega_{\eta\eta}) \\ vec(ds_{t-1} \eta_t^T) \end{bmatrix} \ .$$

The details of the matrix coefficients $\theta_{z\xi}^{(2)}$, $\theta_{\xi\xi}^{(2)}$, and $\theta_{\xi v}^{(2)}$ can be directly inferred from (27), (30), and (31).[23] We discuss several important properties of the above state-space solution in Section 4.

3.2.3 The third-order approximate solution

For the third-order approximate solution, we focus on presenting the general steps to obtain the solution and its general form. Various details on the coefficients and the vectors of variables and

[23] See Appendix B for the details of the coefficients.

innovations are provided in Appendix B.

Total differentiation of the second-order differential restrictions in (23) leads to the third-order differential restrictions,

$$0 = \Gamma_{z'} E_t d^3 z' + \Gamma_z d^3 z + E_t \delta'^{(3)} \ . \tag{34}$$

The form of the restrictions is therefore similar to the first- and second-order differential restrictions ((16), (17), and (23), respectively), except for the elements of the driving process. The driving process, $\delta'^{(3)}$, is now a function of various cross products of the elements in the first- and second-order differential solutions, which can be treated as exogenous from the standpoint of computing the solution to $d^3 z$.

The evolution of the driving process Using the first-order differential solution in (21) and (22) and the second-order differential solution in (28) and (29), we can write the driving process as a function of the cross products of all the state variables and innovations in the first- and second-order solutions. Collecting these cross-products in a vector $x_{t+1}^{(3)}$, this process leads to

$$\delta_{t+1}^{(3)} = C^{(3)} x_{t+1}^{(3)} \ ,$$

The elements of $x_t^{(3)}$ can be further separated into those that are subject to innovations at time t, $v_t^{(3)}$, and the pseudo states, $\varsigma_{t-1}^{(3)}$, leading to

$$x_t^{(3)} = \gamma^{(3)} + \gamma_\varsigma^{(3)} \varsigma_{t-1}^{(3)} + \gamma_v^{(3)} v_t^{(3)} \ . \tag{35}$$

This process is the third-order counterpart to (26) in the second-order differentials. Here, for example,

$$\varsigma_t^{(3)} \equiv \begin{bmatrix} \xi_t^{(1)} \\ vec\left(\xi_t^{(1)} \tilde{\xi}_t^{(2)T} - \Omega_{\xi s} \right) \end{bmatrix} \ ,$$

with $\tilde{\xi}_t^{(2)} = \xi_t^{(2)} + E\left[\tilde{\xi}_t^{(2)}\right]$ is the non-demeaned version of the full state vector in the second-order differentials, $\xi_t^{(2)}$, and $\Omega_{\xi s} = E\left[\xi_t^{(1)} \tilde{\xi}_t^{(2)T}\right]$.[24] The third-order pseudo states therefore consists of the cross products of all state variables present in the first- and second-order differential solutions. As

[24] Specifically,

$$\tilde{\xi}_t^{(2)} \equiv \begin{bmatrix} d^2 s_t \\ vech(ds_t ds_t^T) \end{bmatrix} \ .$$

We use $\tilde{\xi}_t^{(2)}$, instead of the demeaned version $\xi_t^{(2)}$, simply for clarity of exposition.

in the second order, both $\varsigma_t^{(3)}$ and $v_t^{(3)}$ have been demeaned.

From the evolutions of the state variables in the first- and second-order solutions in (22) and (33), it follows then $\varsigma_t^{(3)}$ can be found to evolve according to

$$\varsigma_t^{(3)} = \phi^{(3)} + \phi_\varsigma^{(3)}\varsigma_{t-1}^{(3)} + \phi_v^{(3)}v_t^{(3)} \ . \tag{36}$$

The determinacy condition and the differential solution The differential restrictions in (34) can now be written in expectational linear-difference system,

$$AE_t d^3 z_{t+1} = Bd^3 z_t + C^{(3)}E_t x_{t+1}^{(3)} \ ,$$

in a manner similar to the first- and second-order differentials. Hence, the solution to $d^3 z_t$ is also determinate. Given (35) and (36), the solution is obtained as

$$\begin{aligned} d^3 z_t &= E[d^3 z] + \theta_{yk}^{(3)}\left(d^3 s_t - E[d^3 s]\right) + \theta_{y\varsigma}^{(3)}\varsigma_t^{(3)} & (37)\\ d^3 s_t - E[d^3 s] &= \theta_{kk}^{(3)}\left(d^3 s_{t-1} - E[d^3 s]\right) + \theta_{k\varsigma}^{(3)}\varsigma_{t-1}^{(3)} + \theta_{kv}^{(3)}v_t^{(3)}. & (38)\end{aligned}$$

This solution is already cast in a fully-demeaned form, as in (30) and (31) in the second order differentials. Here, $d^3 s_t$ is the third differentials of the states s_t and $E[d^3 z] = \theta_y^{(3)} + \theta_{yk}^{(3)}E[d^3 s]$ is the unconditional expectation of the second-order differentials.

The third-order approximate solution Imposing our definition of the approximate solution in (4) the third-order approximate solution is

$$\begin{aligned} z_t &\simeq \bar{z} + dz_t + \frac{1}{2}d^2 z_t + \frac{1}{6}d^3 z_t \\ &= \left\{\bar{z} + \theta_{z\xi}^{(1)}\xi_t^{(1)}\right\} + \frac{1}{2}\left\{E[d^2 z] + \theta_{z\xi}^{(2)}\xi_t^{(2)}\right\} + \frac{1}{6}\left\{E[d^3 z] + \theta_{z\xi}^{(3)}\xi_t^{(3)}\right\} \ , \end{aligned} \tag{39}$$

with the full state vector evolves according to

$$\begin{bmatrix} \xi_{t+1}^{(1)} \\ \xi_{t+1}^{(2)} \\ \xi_{t+1}^{(3)} \end{bmatrix} = \begin{bmatrix} \theta_{\xi\xi}^{(1)} & 0 & 0 \\ 0 & \theta_{\xi\xi}^{(2)} & 0 \\ 0 & 0 & \theta_{\xi\xi}^{(3)} \end{bmatrix} \begin{bmatrix} \xi_t^{(1)} \\ \xi_t^{(2)} \\ \xi_t^{(3)} \end{bmatrix} + \begin{bmatrix} \theta_{\xi v}^{(1)} & 0 & 0 \\ 0 & \theta_{\xi v}^{(2)} & 0 \\ 0 & 0 & \theta_{\xi v}^{(3)} \end{bmatrix} \begin{bmatrix} v_{t+1}^{(1)} \\ v_{t+1}^{(2)} \\ v_{t+1}^{(3)} \end{bmatrix} \tag{40}$$

We have collected all the state variables in the third-order differentials into vectors $\xi_t^{(3)}$:

$$\xi_t^{(3)} = \begin{bmatrix} d^3 s_t - E[d^3 s] \\ \varsigma_t^{(3)} \end{bmatrix}.$$

The details of matrix coefficients $\theta_{z\xi}^{(3)}$, $\theta_{\xi\xi}^{(3)}$, and $\theta_{\xi v}^{(3)}$ can be directly inferred from (36), (37), and (38).

3.2.4 Higher than third order

Extensions to the fourth order and higher are straightforward, albeit tedious. We can follow similar mechanisms and steps as the above. To obtain the i-th order differential restrictions, we just need to take the total differentiation to the restrictions in the $(i-1)$-th order. These restrictions have similar form to those in any lower order, except for the driving process. The resulting driving process is a function of cross-products of all lower-order differential solutions, which in turn can be written as a function of cross products of all lower-order state variables and innovations. After identifying the pseudo state and innovation vectors and obtaining the evolution of the states, the solution to the i-th order differentials can then be obtained using any standard linear-rational expectations methods. Finally, collecting all the relevant differential solutions, one can obtain the n-th order approximate solution of interest defined in (4). The approximate solution to any n-th order can always be cast in a state-space form as each differential solution is obtained in a state-space form.

4 Discussion

4.1 Properties of the approximate solutions

We discuss here some important properties of the first-, second-, and third-order approximate solutions depicted in (21)-(22), (32)-(33), and (39)-(40). First, it is clear that all three approximate solutions can be cast in a linear state-space form. Hence, no matter what the approximation order is, one can, for example, generate impulse responses or perform simulations using the state-space form. In terms of determinacy, as previously indicated, the solution to any approximation order is determinate as long as the first-order solution satisfies the uniqueness and stability conditions. These determinacy properties are reflected in various coefficients in the second-order and third-order solutions. As shown in Appendix B, the eigenvalues of $\theta_{\xi\xi}^{(2)}$ and $\theta_{\xi\xi}^{(3)}$ are guaranteed to be

inside unity as long as the eigenvalues of $\theta^{(1)}_{\xi\xi} = \Phi_s$ are inside unity. It follows then, much as in the simple example in Section 3.1, both the second- and third-order state vectors, $\xi^{(2)}_t$ and $\xi^{(3)}_t$, are non-explosive, implying the stationarity of solutions for all orders of approximation.

The effect of uncertainty, however, are captured differently across different orders. As is well known, there is certainty equivalence in the first-order approximate solution, both in terms of the policy and state evolution equations ((21) and (22)). That is, the coefficients $\theta^{(1)}_{z\xi}$ and $\theta^{(1)}_{\xi\xi}$ are identical to their non-stochastic counterparts. In the second-order solution, uncertainty affects only up to a constant term, reflected by the presence of $E[d^2 z]$ term in (32). The coefficients $\theta^{(2)}_{z\xi}$ and $\theta^{(2)}_{\xi\xi}$, however, are identical to those in the non-stochastic version of the model. These results are therefore consistent with Theorem 1 in Schmitt-Grohe and Uribe (2004).[25] On the elements of the innovation term, notice that since

$$v^{(2)}_t \equiv \begin{bmatrix} vech(\eta_t \eta_t^T - \Omega_{\eta\eta}) \\ vec(ds_{t-1}\eta_t^T) \end{bmatrix},$$

our second-order state-space solution also implies that there is a time-varying conditional variance-covariance matrix of the one-step-ahead forecast errors, with second moments of the η_t playing a role. Due to the nonlinearity of the underlying model, the size of the shocks, η_t, matters and has a second-order uncertainty effect (the $vech(\eta_t \eta_t^T - \Omega_{\eta\eta})$ term) and the effect of the shocks to the system is state dependent and conditionally heteroskedastic (the $vec(ds_{t-1}\eta_t^T)$ term).

The third-order state-space solution captures the effect of uncertainty one step further. In addition to the constant-term effect through the presence of $E[d^3 z]$, there is an additional time-varying, state-dependent effect of uncertainty. Although not directly visible from (39)-(40), this state dependence is captured through the presence of ds_t in $\xi^{(3)}_t$ — as the constant term, this term only appears when the model is stochastic. It is still the case, however, that the coefficients $\theta^{(3)}_{z\xi}$ and $\theta^{(3)}_{\xi\xi}$ are identical to those in the non-stochastic version of the model. We therefore extend Theorem 1 in Schmitt-Grohe and Uribe to include the third-order approximation as well.

[25] The constant correction terms in Schmitt-Grohe and Uribe (2004) — $h_{\sigma\sigma}$ and $g_{\sigma\sigma}$, using their notations — can be shown to be identical to those in our second-order solution in (32) and (33). Using the details of the coefficients of our solution in Appendix B, it is the case that $h_{\sigma\sigma} = \theta^{(2)}_k - \theta^{(2)}_{k\varsigma}\Omega_{ss} - \theta^{(2)}_{kv}\Omega_{\eta\eta}$ and $g_{\sigma\sigma}$ is equal to the first n_y elements of $E[d^2 z] - \theta^{(2)}_{yk} E[d^2 s] - \theta^{(2)}_{y\varsigma}\Omega_{ss}$. Here, we assume that endogenous (jump) variables are ordered first in the vector of all variables, z_t, and n_y is the number of endogenous variables.

4.2 Conventional vs. differential Taylor-series approximation and pruning

Our use of a differential Taylor-series (TS) approximation instead of the conventional one (e.g. as used in Schmitt-Grohe and Uribe, 2005) requires some discussion. To compare the two versions of TS approximations, we utilize the simple scalar, descriptive model analyzed in Section 3.1, simply for ease of exposition. The analysis based on this simple model and the subsequent discussion extend to the general class of model as well. We also provide an explanation of the "pruning" mechanism advocated in Kim, et al. (2008), especially in relation to our linear state-space solution. To permit direct comparison to the two aforementioned studies, we focus on the second-order approximate solution.[26]

The approximate solution to simple model based on the differential TS approximation is given in (14)-(15). This state-space form can be rewritten as and is equivalent to

$$z_t \simeq \bar{z} + f_s\left(ds_t + \frac{1}{2}d^2 s_t\right) + \frac{1}{2}f_{ss}\left(ds_t ds_t\right) \ , \tag{41}$$

$$\begin{bmatrix} ds_t + \frac{1}{2}d^2 s_t \\ ds_t ds_t \end{bmatrix} = \begin{bmatrix} m_s & \frac{1}{2}m_{ss} \\ 0 & m_s^2 \end{bmatrix} \begin{bmatrix} ds_{t-1} + \frac{1}{2}d^2 s_{t-1} \\ ds_{t-1} ds_{t-1} \end{bmatrix} + \begin{bmatrix} m_\eta & m_{s\eta} & \frac{1}{2}m_{\eta\eta} \\ 0 & 2m_s m_\eta & m_\eta^2 \end{bmatrix} \begin{bmatrix} \eta_t \\ \eta_t ds_{t-1} \\ \eta_t \eta_t \end{bmatrix} . \tag{42}$$

The second-order approximate solution based on the conventional TS approximation to the same model is on the other hand given by

$$z_t \simeq \bar{z} + f_s\left(\widetilde{ds}_t\right) + \frac{1}{2}f_{ss}\left(\widetilde{ds}_t \widetilde{ds}_t\right) \ , \tag{43}$$

with the state-evolution equation

$$\widetilde{ds}_t = m_s \widetilde{ds}_{t-1} + m_\eta \eta_t + \frac{1}{2}\left(m_{ss}\widetilde{ds}_t \widetilde{ds}_t + \boxed{2m_{s\eta}\widetilde{ds}_{t-1}\eta_t} + \boxed{m_{\eta\eta}\eta_t\eta_t}\right) \tag{44}$$

Here, $\widetilde{ds}_t \equiv s_t - \bar{s}$. The expressions above are obtained by directly employing the standard Taylor series formula on the functions $f(.)$ and $s(.)$. In Schmitt-Grohe and Uribe, the boxed terms are not present because the shocks in the underlying solution of the model are assumed to only directly affect predetermined variables and are additively separable (linear). These terms are included here

[26] Schmitt-Grohe and Uribe (2004) and Kim, et al. (2008) only provide solutions up to a second-order approximation.

for comparability and completeness in the discussion. Also, notice that since the simple model is completely descriptive, the constant correction (stochastic) terms are not present.

Pruning Observe that if one were to use (43) and (44) directly to generate impulse responses or perform simulations, z_t would accumulate terms that are of order three or higher. These higher-order terms do not necessarily improve the accuracy of the second-order solution and generally lead to explosive forecasts and simulated time paths. Kim, et al. (2008) advocate that these extraneous higher-order terms should be "pruned" out. This "pruning" mechanism entails making $\widetilde{ds}_t\widetilde{ds}_t$ in (43) and (44) to evolve according to the first-order approximate solution. Expressing the first-order solution as $ds_t = m_s ds_{t-1} + m_\eta \eta_t$, this means

$$\begin{aligned} \widetilde{ds}_t\widetilde{ds}_t &\approx ds_t ds_t \\ &= m_s^2 ds_{t-1}ds_{t-1} + 2m_s m_\eta \eta_t ds_{t-1} + m_\eta^2 \eta_t\eta_t \ . \end{aligned}$$

As the term $\widetilde{ds}_{t-1}\eta_t$ in (44) also leads to extraneous higher-order terms, it also needs to be pruned out, leading to $\widetilde{ds}_{t-1}\eta_t \approx ds_{t-1}\eta_t$. Hence, the pruned second-order solution based on the conventional TS approximation leads to the following state-space form:

$$z_t \simeq \bar{z} + f_s\left(\widetilde{ds}_t\right) + \frac{1}{2}f_{ss}\left(ds_t ds_t\right) \ , \tag{45}$$

$$\begin{bmatrix} \widetilde{ds}_t \\ ds_t ds_t \end{bmatrix} = \begin{bmatrix} m_s & \frac{1}{2}m_{ss} \\ 0 & m_s^2 \end{bmatrix} \begin{bmatrix} \widetilde{ds}_{t-1} \\ ds_{t-1}ds_{t-1} \end{bmatrix} + \begin{bmatrix} m_\eta & m_{s\eta} & \frac{1}{2}m_{\eta\eta} \\ 0 & 2m_s m_\eta & m_\eta^2 \end{bmatrix} \begin{bmatrix} \eta_t \\ \eta_t ds_{t-1} \\ \eta_t\eta_t \end{bmatrix} \ . \tag{46}$$

How does this pruned solution differ from our solution based on the differential TS approximation? Comparing (41)-(42) with (45)-(46) it is clear that two state-space forms are identical, provided that

$$\widetilde{ds}_t = ds_t + \frac{1}{2}d^2 s_t \ .$$

In fact, this is our interpretation of the second differentials of the states, $d^2 s_t$: it is the second-order expansions of the state vector based on the conventional TS approximation, net the first-order expansions. One can therefore interpret $d^2 s_t$ as an approximate second-order expansion term necessary to make the second-order approximate solution to be non-explosive.

Another way to look at the relationship between our differential TS approach and pruning

under conventional TS is by considering a descriptive nonlinear autoregressive scalar system, $z_t = m(z_{t-1}, \eta_t)$.[27] The second-order solution based on the conventional TS approximation to this system is

$$\tilde{z}_t = m_z \tilde{z}_{t-1} + m_\eta \eta_t + \frac{1}{2}\left[m_{zz}\tilde{z}_{t-1} + 2m_{z\eta}\tilde{z}_{t-1}\eta_t + m_{\eta\eta}\eta_t^2\right] \ ,$$

with $\tilde{z}_t \equiv z_t - \bar{z}$. The exogenous pruning mechanism involves replacing $\tilde{z}_{t-1}$ in the squared bracket above by its first-order approximation, dz_{t-1}, yielding

$$\tilde{z}_t \simeq m_z \tilde{z}_{t-1} + m_\eta \eta_t + \frac{1}{2}\left[m_{zz}(dz_{t-1})^2 + 2m_{z\eta}dz_{t-1}\eta_t + m_{\eta\eta}\eta_t^2\right] \ .$$

To show that $\tilde{z}_t \simeq dz_t + \frac{1}{2}d^2 z_t$, replace $\tilde{z}_{t-1}$ in the above expression with $\tilde{z}_{t-1} = dz_{t-1} + \frac{1}{2}d^2 z_{t-1}$,

$$\begin{aligned}
\tilde{z}_t &\simeq m_z\left(dz_{t-1} + \frac{1}{2}d^2 z_{t-1}\right) + m_\eta \eta_t + \frac{1}{2}\left[m_{zz}(dz_{t-1})^2 + 2m_{z\eta}dz_{t-1}\eta_t + m_{\eta\eta}\eta_t^2\right] \\
&= (m_z dz_{t-1} + m_\eta \eta_t) \\
&\quad + \frac{1}{2}\left[m_z d^2 z_{t-1} + m_{zz}(dz_{t-1})^2 + 2m_{z\eta}dz_{t-1}\eta_t + m_{\eta\eta}\eta_t^2\right]
\end{aligned}$$

The right-hand-side of the above expression is exactly the second-order approximation based on the differential TS, i.e. $dz_t + \frac{1}{2}d^2 z_t$.

We thus conclude that our approximation approach based on the differential TS leads to an identical solution to the pruned approximate solution based on the conventional TS approximation in Kim, et al. Our use of the differential version of the TS approximation therefore renders pruning unnecessary. Yet, on the flip side, our approach also justifies that the "exogenous" pruning mechanism advocated in Kim, et.al is in fact "approximation-consistent," in a sense of the Taylor-series expansions. Our approach, however, offers an advantage over the conventional TS approximation in that the solution is always obtained in a linear state-space for *any order* of approximation. And this linear state-space form always leads to non-explosive impulse responses, forecasts, and simulated time paths.

Comparing the pre-pruned solution based on the conventional TS approximation in (43)-(44) with our solution in (41)-(42), we can also deduce that our approximation strategy provides the researchers with instructions about what to place on the right-hand side of the equality sign: it is a complete description of how to construct the approximate stochastic process. And this approximate stochastic process is always non-explosive as long as the linear model is non-explosive. By contrast,

[27] Thanks to May Li for pointing out this example.

the conventional TS approximation leaves some latitude. If one were working econometrically so that actual lagged values were available, for example, one could use data on s_{t-1} and $(s_{t-1})^2$. However, to simulate the system, as in the construction of impulse responses, it is necessary to somehow construct or approximate these elements.

5 Applications

In this section, we apply our approximation approach to two particular models: the one-sector neoclassical growth model and the Q-theory model of lumpy investment in Miao and Wang (2013). For each model, we compute the second-order approximate solution. We choose the neoclassical growth model to facilitate direct comparison to the second-order approximation approach in Schmitt-Grohe and Uribe (2004) and to show how to employ our approach. The second model is chosen to show that the dynamics obtained from the second-order approximate solution could be very different compared to those coming from the first-order solution.

5.1 The neoclassical growth model

The simple neoclassical growth model consists of the following equations:

$$c_t^{-\gamma} = \beta E_t c_{t+1}^{-\gamma} \left[\alpha A_{t+1} k_{t+1}^{\alpha-1} + 1 - \delta\right] \ ,$$

$$c_t + k_{t+1} = A_t k_t^{\alpha} + (1-\delta) k_t \ ,$$

$$\ln A_{t+1} = \rho \ln A_t + \sigma \varepsilon_{t+1} \ .$$

For this application, we use a computation approach in which the researchers have to specify which variables are endogenous and exogenous, following King and Watson (1998, 2002), Klein (2004), and Schmitt-Grohe and Uribe (2004). Here, c_t is a non-predetermined endogenous variables, k_t is a predetermined endogenous variable (at time t), and A_t is an exogenous variable. k_0 and A_0 are given. ε_t is i.i.d. with zero mean and unit variance. Following Schmitt-Grohe and Uribe (2004) we set $\beta = 0.95$, $\delta = 1$, $\alpha = 0.3$, $\rho = 0$, $\gamma = 2$, and $\sigma = 1$.[28] The calibration implies $\ln \bar{c} = -0.8734$, $\ln \bar{k} = -1.7932$, and $\ln \bar{A} = 0$ in the deterministic steady state. To solve the model using our

[28] Note that the notations for the parameters and variables used in the model apply only for this subsection.

approach, we define

$$z_t = \begin{bmatrix} \widetilde{c}_t \\ \widetilde{k}_t \\ \widetilde{A}_t \end{bmatrix} \equiv \begin{bmatrix} \ln c_t \\ \ln k_t \\ \ln A_t \end{bmatrix}$$

and

$$\eta_{t+1} = [\varepsilon_{t+1}] \ ,$$

so that the nonlinear model can be represented by the system of nonlinear equations as in (3), $0 = E_t\Gamma(z_{t+1}, z_t, \eta_{t+1})$. Here, without any loss of generality, we order the state variables (non-predetermined endogenous and exogenous variables) last in z_t.[29]

Computing the approximate equilibrium solution to the second-order, we obtain a state-space form as in (32)-(33) with

$$\xi_t^{(1)} = \begin{bmatrix} d\widetilde{k}_t \\ d\widetilde{A}_t \end{bmatrix} \equiv \begin{bmatrix} \hat{k}_t \\ \hat{A}_t \end{bmatrix} \ ,$$

$$\xi_t^{(2)} = \begin{bmatrix} d^2\widetilde{k}_t - E\left[d^2\widetilde{k}\right] \\ d^2\widetilde{A}_t - E\left[d^2\widetilde{A}\right] \\ \left(d\widetilde{k}_t\right)^2 - E\left[\left(d\widetilde{k}\right)^2\right] \\ d\widetilde{A}_t d\widetilde{k}_t - E\left[d\widetilde{A}d\widetilde{k}\right] \\ \left(d\widetilde{A}_t\right)^2 - E\left[\left(d\widetilde{A}\right)^2\right] \end{bmatrix} \equiv \begin{bmatrix} d^2\widetilde{k}_t - E\left[d^2\widetilde{k}\right] \\ d^2\widetilde{A}_t - E\left[d^2\widetilde{A}\right] \\ \hat{k}_t^2 - E\left[\hat{k}^2\right] \\ \hat{A}_t\hat{k}_t - E\left[\hat{A}\hat{k}\right] \\ \hat{A}_t^2 - E\left[\hat{A}^2\right] \end{bmatrix} \ ,$$

where $\hat{x}_t \equiv d\widetilde{x}_t = \ln(x_t/\bar{x})$ for any variable x_t.[30] The coefficients of the first-order terms are:[31]

$$\theta_{z\xi}^{(1)} = \begin{bmatrix} 0.2525 & 0.8417 \\ 1 & 0 \\ 0 & 1 \end{bmatrix} , \theta_{\xi\xi}^{(1)} = \begin{bmatrix} 0.4191 & 1.3970 \\ 0 & 0 \end{bmatrix} .$$

[29] Instead of defining variables in terms of their logarithms, one can alternatively define $z_t = \begin{bmatrix} c_t & k_t & A_t \end{bmatrix}^T$. We use logarithms to permit direct comparison with the computation results in Schmitt-Grohe and Uribe (2004).

[30] The MATLAB codes are available upon request.

[31] Several of 0's 1's in various coefficients are the results of rounding — but they are practically 0 or 1.

The coefficients of the second-order terms are given by:

$$\theta^{(2)}_{z\xi} = \begin{bmatrix} 0.2525 & 0.8417 & -0.0051 & -0.0341 & -0.0569 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \end{bmatrix},$$

$$\theta^{(2)}_{\xi\xi} = \begin{bmatrix} 0.4191 & 1.3970 & -0.0070 & -0.0467 & -0.0778 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0.1757 & 1.1710 & 1.9517 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

The vectors of innovations are

$$v^{(1)}_{t+1} = \begin{bmatrix} \varepsilon_{t+1} \end{bmatrix}, \; v^{(2)}_{t+1} = \begin{bmatrix} (\varepsilon_{t+1})^2 \\ \hat{k}_t \varepsilon_{t+1} \\ \hat{A}_t \varepsilon_{t+1} \end{bmatrix},$$

with coefficients

$$\theta^{(1)}_{\xi v} = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \quad \text{and}$$

$$\theta^{(1)}_{\xi v} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0.4191 & 1.3970 \\ 1.000 & 0 & 0 \end{bmatrix}$$

Finally, various expected values of the second-differentials and the covariance matrix of the state vectors are given by:

$$E[d^2 z] = E\left\{ \begin{bmatrix} d^2\widetilde{c} \\ d^2\widetilde{k} \\ d^2\widetilde{A} \end{bmatrix} \right\} = \begin{bmatrix} -0.0926 \\ 0.6674 \\ 0 \end{bmatrix} \quad \text{and} \quad E\left\{ \begin{bmatrix} d^2\widetilde{k} \\ d^2\widetilde{A} \\ \hat{k}^2 \\ \hat{A}\hat{k} \\ \hat{A}^2 \end{bmatrix} \right\} = \begin{bmatrix} 0.6674 \\ 0 \\ 2.3676 \\ 0 \\ 1 \end{bmatrix}.$$

The analysis in Section 4.2 shows that it is possible to associate each of the coefficients above to those arising from the "pre-pruned" approximate solution, like the method used in Schmitt-Grohe and Uribe (2004). Applying the analysis to the neoclassical growth model above leads to the following laws of motion for the two endogenous variables:

$$\hat{c}_t = 0.2525\hat{k}_t + 0.8417\hat{A}_t + \frac{1}{2}\left[-0.1921 - 0.0051\hat{k}_t^2 - 0.0341\hat{A}_t\hat{k}_t - 0.0569\hat{A}_t^2\right]$$

and

$$\hat{k}_{t+1} = 0.4191\hat{k}_t + 1.3970\hat{A}_t + \frac{1}{2}\left[0.4820 - 0.0070\hat{k}_t^2 - 0.0467\hat{A}_t\hat{k}_t - 0.0778\hat{A}_t^2\right] \ .$$

These representations are identical to those in Schmitt-Grohe and Uribe (see Section 5.1 in their paper).[32]

5.2 The Q-theory model

Our second application is on the Q-theory model with lumpy investment of Miao and Wang (2013). Rather than presenting all the equations in the model, here we just describe the general features of the model. The model in Miao and Wang is a simple, analytically tractable investment model that incorporates both convex capital adjustment costs and lumpy investment at the micro level. Lumpy investment is achieved through the presence of fixed costs of investments, as in the generalized (S,s) model of Caballero and Engel (1999). The fixed costs are assumed to be random and drawn independently across firms in each period from a time-invariant distribution. Firms that draw a fixed cost that is lower than a cutoff fixed cost value will optimally choose to adjust their capital levels, while firms drawing a fixed cost higher than this cutoff value will choose to wait until at least next period to invest. Despite this lumpiness at the micro level, at the aggregate level, the model is shown to be isomorphic to a standard Q-theory model with only convex adjustment costs (and without fixed costs) as in Tobin (1969) and Hayashi (2003).[33,34]

This isomorphism result is an appealing feature since that means in the aggregate the model is essentially a real business cycle (RBC) model with a standard investment feature — and it is well known that RBC models are highly linear, i.e. a first-order approximate solution to the model is

[32] See footnote 23 on how to compute the second-order constant terms.

[33] Three conditions need to be satisfied for the isomorphism results: (i) the production technology is of constant return to scale, (ii) the convex adjustment cost function is homogenous of degree one in capital and investment, and (iii) the fixed costs do not become negligible (e.g. proportional to the level of capital).

[34] The fact that lumpy investment does not appear to matter for the aggregate dynamics means that the model is consistent with the findings in Thomas (2002) and Khan and Thomas (2009).

very accurate. Yet, the presence of lumpy investment means there is a potential gain in accuracy when the model is solve to a second-order approximation. That is, the fact that the model features fixed investment costs implies that firms' capital adjustment decision is state dependent and hence, there are potentially important nonlinearity effects that are not captured by a first-order (linear) approximation. Consider the following two equations in the model:

$$\begin{aligned} \frac{I_t}{K_t} &= i_t \left[\int_0^{\xi_t^*} \phi(\xi) d\xi \right] \\ K_{t+1} &= (1-\delta) K_t + g(i_t) \left[\int_0^{\xi_t^*} \phi(\xi) d\xi \right] K_t \end{aligned}$$

The first equation shows that the aggregate investment rate, I_t/K_t, positively depends on the firms' investment target i_t and the proportion of firms that choose to adjust in a given period (the adjustment hazard rate), $\left[\int_0^{\xi_t^*} \phi(\xi) d\xi\right]$. This hazard rate positively depends on the fixed cost cutoff, ξ_t^*. The former thus represents the intensive margin effect, while the latter is the extensive margin effect. It is the additional presence of the extensive margin effect that makes the model to be potentially highly nonlinear — the degree of nonlinearity, of course, depends on the shape of the fixed-cost distribution (PDF), $\phi(.)$. The nonlinearity effect of the extensive margin also affects the rest of the variables in the model, e.g. the capital stock K_{t+1}, as is apparent from the second equation above ($g(.)$ is the convex adjustment cost function, with $g'(i) > 0$).

The rest of the model is standard, e.g. the utility function is of constant-relative-risk-aversion type and the production function is Cobb-Douglas.[35] The convex adjustment cost function is set to $g(i) = (\psi/(1-\theta))i^{1-\theta}$. For the distribution of the fixed costs, we do not use a specific functional form. Instead, we calibrate its level and curvatures to have certain values at the steady state, as is common in the neoclassical investment literature.[36] The rest of the calibration is standard.[37] We solve this model to the second order using our approach and generate impulse responses to aggregate productivity shocks. Productivity is assumed to followed an AR(1) process $\ln A_{t+1} = \rho \ln A_t + \sigma\varepsilon_{t+1}$, with $\rho = 0.95$, $\sigma = 0.0072$, and $\sigma \sim i.i.d\ (0,1)$, which implies $\eta_{t+1} = \sigma\varepsilon_{t+1} \sim i.i.d\ (0, \sigma^2)$.[38]

[35] The utility function is $u(c_t, n_t) = (1-\sigma)^{-1} c_t^{1-\sigma} - \chi(1+\eta)^{-1} n_t^{1+\eta}$, with the production function given by $y_t = K_t^{\alpha}(A_t n_t)^{1-\alpha}$.

[36] Specifically, we calibrate the hazard rate (the CDF), $\Phi(\xi^*)$ to be 0.2 in the steady state. We also calibrate $\xi^* = 0.0015$ and $\xi^*\phi(\xi^*)/\Phi(\xi^*) = 1$ (unit elasticity of the hazard rate) to obtain the value for the PDF, $\phi(\xi^*)$. Lastly, to obtain the curvature value, $\phi'(\xi^*)$, we calibrate $\xi^*\phi'(\xi^*)/\phi(\xi^*) = 10$. Note that one could alternatively look for and use a specific distribution function that generates the calibrated level and curvature values.

[37] We set $\sigma = 1$, $\eta = 0.05$, $\psi = 0.5$, $\theta = 0.05$, $\alpha = 1/3$. The depreciation rate and the discount factor are $\delta = 0.04$ and $\beta = 0.99$, respectivelly. χ is chosen so that the steady-state labor, n, is 0.2.

[38] Equivalently, we can set $\sigma = 1$ and $\varepsilon_{t+1} \sim i.i.d\ (0, \sigma^2)$.

A one-time productivity shock We first look at the a one-time shock to productivity (at period 0), depicted in Figure 1. The size of the shock is three standard deviations and the economy was at the the steady state prior to the shock. For each variable in the figure, we compare the responses based on the first-order and the second-order approximate solutions in (21)-(22) and (32)-(33), respectively.

Looking first at the top four panels, a three-standard-deviation positive productivity shock translates to a roughly 2.1% increase in productivity on impact. Since A_t follows a log-linear AR(1) process, it follows then the second-order responses are identical to those in the first order. Higher productivity level leads to higher aggregate output, consumption, and investment rate. As described above, the rise in investment rate has both the intensive and extensive margin effects. That is, since the marginal product of capital and the capital price (marginal Q) increase in response to a positive shock to productivity, both the investment target for each capital-adjusting firm and the number of firms investing (the hazard rate) also rise. Comparing the first- and second-order responses across the top four panels, we see that there are some sizeable differences quantitatively. Interestingly, consumption increases by less on impact and exhibits a hump-shaped response when the model is solved to the second order — this hump-shaped feature is not present in the first-order response. The investment rate also increases by about 3% more on impact in the second-order approximation compared to that in the first order. This is a sizeable difference, considering that the size of the shock to productivity is quite small. The higher increase in the investment rate translates to higher output responses in the second order approximation. These quantitative differences persist for a number of periods after the initial impact of the shock.

The explanation behind the differences between the first- and second-order responses, particularly on the investment rate, is provided in the bottom four panels of Figure 1. The higher increase in the investment rate in the second-order approximation is fueled entirely by a much higher increase in the hazard rate — here, it increases by roughly 3.5% more in the second order on impact compared to that in the first order. Putting it differently, based on the second-order approximation there is now about 35% more firms (3.5%/10%) that choose to adjust on impact compared to when the model is solved to the first order. This increase in the extensive margin is more than enough to offset the lower increase in the investment target (the intensive margin effect) in the second-order approximation. The fixed cost cutoff seems to increase by less in the second-order approximation, but this is by no means contradicts the higher increase in the hazard rate. Finally, the higher investment rate in the second order approximation also translates to a higher increase in capital

stock — this is especially visible in the latter periods. We note that without the lumpy investment feature in the model, the two approximation orders produce virtually identical responses for all variables.

The results in Figure 1 thus show that although qualitatively largely similar, the responses based on the first- and second-order approximate solutions could be quite different quantitatively. These differences may matter for other practical applications such as forecasting and simulation, to name a few. This is also true for other models that feature state-dependent agents' decision, such as the pricing model of Dotsey, King, and Wolman (1999).[39] For these models, solving the model only up to a first order approximation may not appropriately capture the important nonlinear elements.

Repeated productivity shocks Notice that from the second-order state-space form solution in (32)-(33), if the economy is assumed to be at the steady state prior to a shock, there is no state-dependent effect of the shock. That is, the $vec(ds_{t-1}\eta_t^T)$ is always zero for all periods since $ds_{t-1} = 0$ prior to the shock. Although this by no means precludes the second-order solution capturing the nonlinearity effect of the model, this suggests that the differences between the responses based on the first- and second-order approximations could be more pronounced when the economy were not initially at the steady state. To capture this state-dependent effect of the shock, we conduct an experiment where the economy is hit by two consecutive three-standard-deviation shocks, each in the first and second periods (periods 0 and 1). This means that in the second period, $ds_{t-1} \neq 0$, so that the effect of the shock is state-dependent from the second period onwards.

Figure 2 displays the impulse responses for this repeated-shock experiment. It is clear that for all variables, the quantitative differences are now even larger compared to those in the one-time shock case in Figure 1. For example, the investment rate is only about 15% higher $((23\% - 20\%)/20\%)$ in the second order approximation on impact in the first period compared to that in the first-order approximation (identical to the one-time shock case). However, in the second period when the economy is again hit by a three-standard-deviation productivity shock, the investment rate is now about 26% higher $((48\% - 38\%)/38\%)$ when the model is solved to the second order. This result indicates again the potential gain in solving the model up to a second-order (or higher) approximation.

[39] The extent of the differences, of course, depends on the specification of other elements in the models, e.g. the shape of the fixed cost distribution in our application.

6 Conclusion

This paper presents a novel and straightforward approach in computing the approximate solutions to nonlinear rational expectations models, based on a differential version of Taylor series approximations. The approach can be applied to any order of approximation and the resulting solution for any order can always be directly cast in a linear state-space form, permitting the solution to be used for many practical applications such as forecasting, estimation, and computing impulse responses. Using the approach, we have shown that there cannot be multiple solutions in any order of approximation if the associated first-order approximate solution is determinate. Since our approach permits the use of linear rational-expectations solution methods, the solutions can be obtained much faster and more efficiently. While the paper only provides an explicit solution up to a third-order approximation, extensions to any higher order approximations are straightforward, albeit tedious. In terms of applications, our approach seems particularly suited for analysis of models in which there is a high degree of nonlinearity and uncertainty such as models with state-dependent agents' decisions and asset pricing models.

In terms of practical applications, the fact that the state-space form is linear for any approximation order means that the standard Kalman filter can be used to obtain the likelihood functions and estimate the parameters and latent states of DSGE models when these models are solved to higher than a first-order (linear) approximation. A recent paper by Kollman (2013) has taken a step on this front for a second-order approximation, but in principal, we can apply the same strategy for any higher order. In terms of future research extensions, we believe that our approximation strategy can be applied to solve Markov-switching models to any higher order as well, given its straightforward nature — perhaps using an algorithm similar to that in Farmer, Waggoner, and Zha (2011) for linear models. Finally, while we presently only focus on computing the unique solution in each approximation order, it is also possible to extend the approach to encompass sunspot solutions, in a manner similar to Lubik and Schorfheide (2004) for a first-order approximation.

Table 1 : Comparison of Existing Studies

	SGU (2004)*	SAL (2005)*	LS (2007)*	KKSS (2008)*	LM (2013)*	JKL (2014)*
Order of approximation	Up to 2nd-order	Up to 7-th order	Up to 2nd-order	Up to 2nd-order	Valid for any order	Valid for any order
Approximation method	Conventional Taylor series (TS)	Conventional TS	Conventional TS	Conventional TS	Volterra/moving-average series (non-state-space method)	Differential TS
Computation method	Undetermined coefficients (UC)	UC	Direct method†	Direct method†	UC	Direct method†
Use of LRE method?^	No	No	Yes	Yes	No	Yes
Restrictions type	On variables	On equations	On variables	On equations	On equations	Indifferent
Solution in state-space form?	No	No	Yes (with "exogenous" pruning)	Yes (with "exogenous" pruning)	No	Yes (approximation-consistent)
Stability	Not addressed	Not addressed	Partly addressed (only up to 2nd-order)	Not addressed	Fully addressed**	Fully addressed**

Notes:

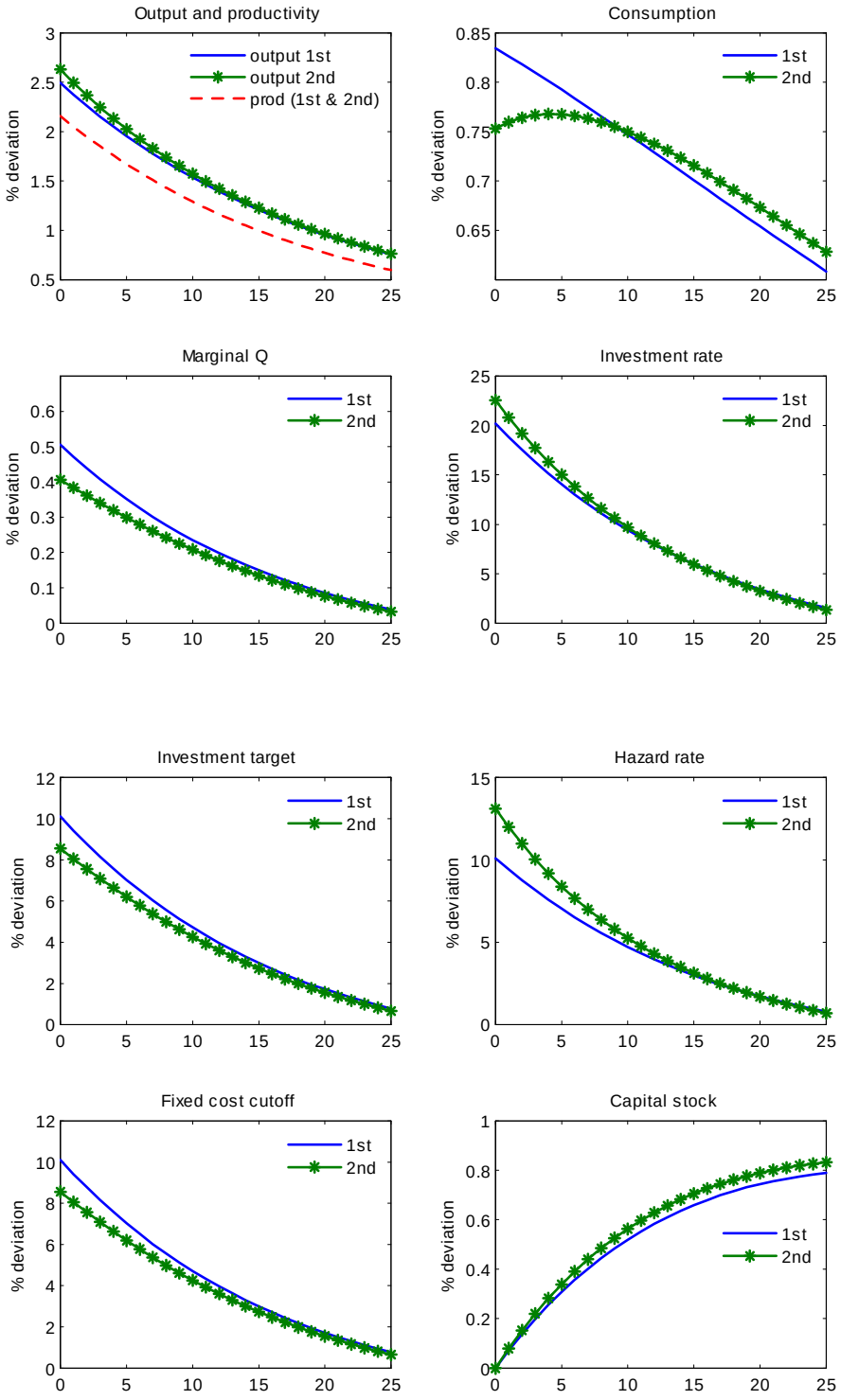
* SGU = Schmitt-Grohe and Uribe, SAL = Swanson, Anderson, and Levin, LS = Lombardo and Sutherland, KKSS = Kim, Kim, Schaumburg, and Sims, LM = Lan and Meyer-Gohde, JKL = Johnston, King, and Lie.

† By direct method, we mean the standard canonical-variable approach, first used in Blanchard and Kahn (1980).

^ LRE = Linear Rational Expectations

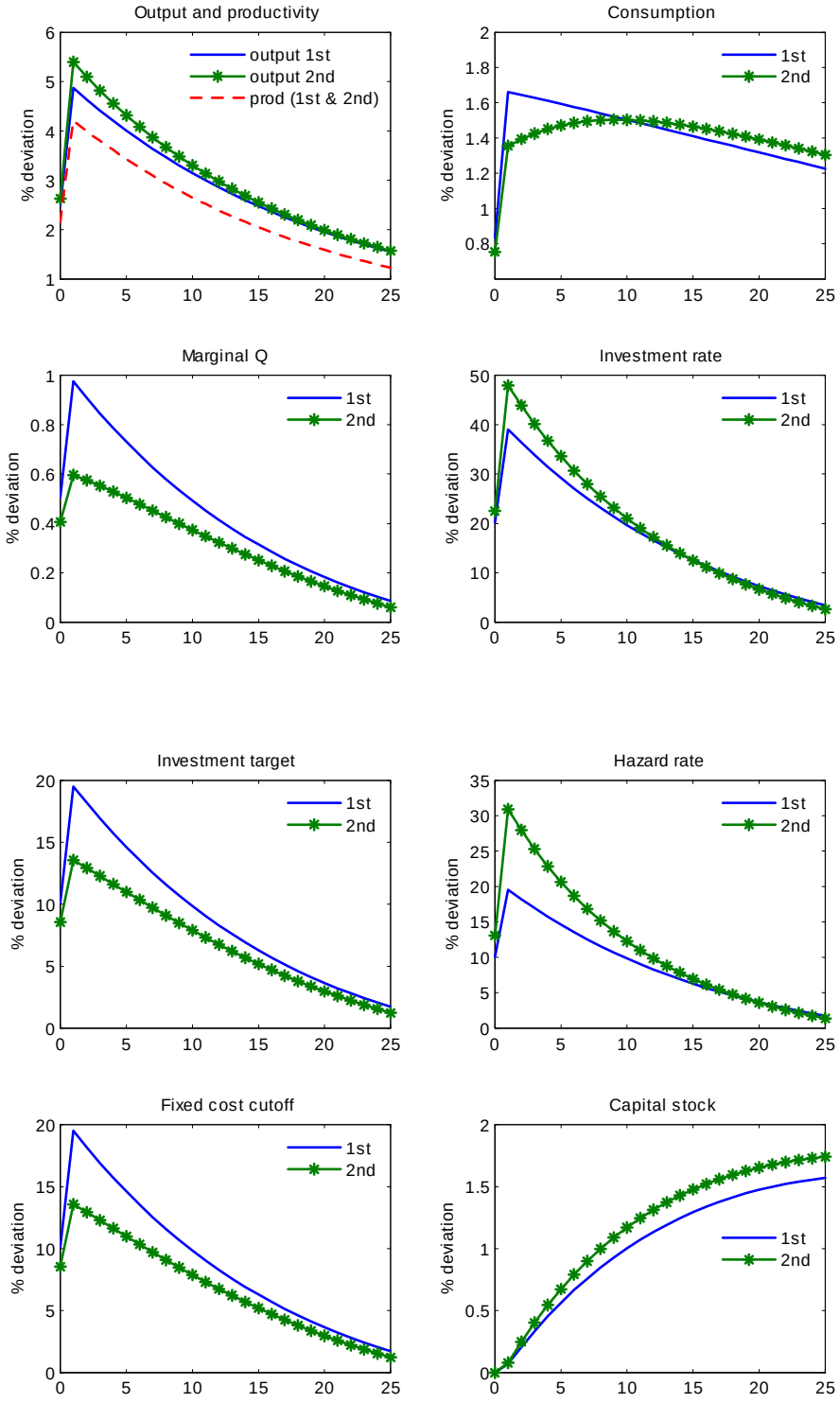
** Any *n*-th order approximate solution is always stable as long as the 1st-order solution is stable.

Figure 1: Q-theory model — Impulse response to productivity shock
One-time, 3 standard-devation shock



Note: "1st" and "2nd" refer to the responses based on the first- and second-order approximate solutions, respectively.

Figure 2: Q-theory model — Impulse response to productivity shock
Repeated (two-period), 3 standard-devation shocks



Note: "1st" and "2nd" refer to the responses based on the first- and second-order approximate solutions, respectively.

References

[1] Anderson, Gary S., and George Moore, 1985. "A Linear Algebraic Procedure for Solving Linear Perfect Foresight Models," *Economics Letters* 17, 247-252.

[2] Blanchard, Olivier J. and Charles M. Kahn, 1980. "The Solution of Linear Difference Models under Rational Expectations," *Econometrica* 48,1305-1313.

[3] Caballero, Ricardo and Eduardo Engel, 1999. "Explaining Investment Dynamics in U.S. Manufacturing: A Generalized (S,s) Approach," *Econometrica* 67(4), 783-826.

[4] Dotsey, Michael, Robert G. King, and Alexander L. Wolman, 1999. "State-Dependent Pricing and the General Equilibrium Dynamics of Money and Output," *The Quarterly Journal of Economics* 114(2), 655-690.

[5] Gutknecht, M.H. and T. Schmelzer, 2009. "The Block Grade of a Block Krylov Space," *Linear Algebra and its Applications* 430(1), 174-185.

[6] Farmer, Roger E.A., Daniel F. Waggoner, and Tao Zha, 2011. "Minimal State Variable Solutions to Markov-Switching Rational Expectations Models," *Journal of Economic Dynamics and Control*, Elsevier, vol. 35(12), pages 2150-2166.

[7] Hayashi, Fumio, 1982. "Tobin's Marginal q and Average q: A Neoclassical Interpretation," *Econometrica* 50, 213-224.

[8] Jonsson, I. and B. Kagstrom, 2002. "Recursive Blocked Algorithms for Solving Triangular Systems—Part I: One-Sided and Coupled Sylvester-Type Matrix Equations," *ACM Transactions on Mathematical Software* (TOMS) 28(4), 392-415.

[9] Khan, Aubhik and Julia Thomas, 2008. "Idiosyncratic Shocks and the Role of Nonconvexities in Plant and Aggregate Investment Dynamics," *Econometrica* 76,395-436.

[10] Kim, Jinill and Sunghyun Kim, 2003. "Spurious Welfare Reversals in International Business Cycle Models," *Journal of International Economics* 60(2), 471-500, August.

[11] Kim, Jinill, Sunghyun Kim, Ernst Schaumburg, and Christopher A. Sims, 2008. "Calculating and Using Second Order Accurate Solutions of Discrete Time Dynamic Equilibrium Models", *Journal of Economic Dynamics and Control* 32, 3397-3414.

[12] King, Robert G., and Mark W. Watson, 1998. "The Solution of Singular Linear Difference Systems Under Rational Expectations," *International Economic Review* 39, 1015-26.

[13] King, Robert G., and Mark W. Watson, 2002. "System Reduction and Solution Algorithms for Singular Linear Difference Systems under Rational Expectations," *Computational Economics* 20, 57–86.

[14] Klein, Paul, 2004. "Using the Generalized Schur Form to Solve a Multivariate Linear Rational Expectations Model," *Journal of Economic Dynamics and Control* 24, 1405–1423.

[15] Kollman, Robert, 2013. "Tractable Latent State Filtering for Non-Linear DSGE Models Using a Second-Order Approximation," CAMA Working Paper 29, Centre for Applied Macroeconomic Analysis, Crawford School of Public Policy, The Australian National University.

[16] Lan, Hong and Alexander Meyer-Gohde, 2013. "Solving DSGE Models with a Nonlinear moving average," *Journal of Economic Dynamics and Control* 37, 2543-67.

[17] Lombardo, Giovanni and Alan Sutherland, 2007. "Computing Second-order-accurate Solutions for Rational Expectation Models: Using Linear Solution Methods," *Journal of Economic Dynamics and Control* 31, 515-30.

[18] Lubik, Thomas A. and Frank Schorfheide, 2004. "Testing for Indeterminacy: An Application to U.S. Monetary Policy," *American Economic Review* 94(1), 190-217.

[19] Magnus, Jan R., and Heinz Neudecker, 2002. *Matrix Differential Calculus with Applications in Statistics and Econometrics*, John Wiley & Sons, (Reprinted, July).

[20] Miao, Jianjun and Pengfei Wang, 2013. "A Q-Theory Model of Lumpy Investment," *Economic Theory*, forthcoming.

[21] Moler, C. B., and G. W. Stewart, 1973. "An Algorithm for Generalized Matrix Eigenvalue Problems," *SIAM Journal of Numerical Analysis*, volume 10, 241-256, April.

[22] Schmitt-Grohe, Stephanie and Martin Uribe, 2004. "Solving Dynamic General Equilibrium Models Using a Second-Order Approximation to the Policy Function," *Journal of Economic Dynamics and Control* 28, 755 – 75.

[23] Sims, Christopher A., 2001. "Solving Linear Rational Expectations Models," *Journal of Computational Economics* 20(1-2), 1-20.

[24] Stewart, G.W., 2001. *Matrix Algorithms Volume II: Eigensystems,* SIAM.

[25] Swanson, Eric T., Gary S. Anderson, and Andrew T. Levin, 2005. "Higher-Order Perturbation Solutions to Dynamic, Discrete-Time Rational Expectations Models," manuscript.

[26] Thomas, Julia, 2002. "Is Lumpy Investment Relevant for the Business Cycle?," *Journal of Political Economy* 110, 508-534.

[27] Tobin, James, 1969. "A General Equilibrium Approach to Monetary Policy," *Journal of Money, Credit, and Banking* 1, 15-29.

[28] Watson, Mark W., 1989. "Recursive Solution Methods for Dynamic Linear Rational Expectations Models," *Journal of Econometrics* 41(1), 65-89, May.

APPENDICES

"Straightforward Approximate Stochastic Equilibria for Nonlinear Rational Expectations Models" (Johnston, King, and Lie, 2014)

[last modified on 25 August 2014]

A Computation using LRE method and the existence and uniqueness of the solution

1 Overview

In this appendix, we describe one of the linear rational expectations solution methods that can be employed for our approximation strategy (to any order). The method is based on the QZ decomposition as in prior work of Sims (2002), Klein (2000), and Kim, Kim, Schaumburg and Sims (2008). As in Sims, the algorithm described here does not rely on the user specifying which variables are predetermined and non-predetermined. The notations in this appendix are self-contained and may overlap with those in the main text.

The differentials of our non-linear rational expectations model suggests the importance of studying the linear rational expectations model of the form

$$\begin{aligned} A_1 y_{t+1} &= B_1 y_t + C_1 x_{t+1} && (1) \\ A_2 E_t y_{t+1} &= B_2 y_t + C_2 E_t x_{t+1} && (2) \end{aligned}$$

with the driving process being

$$\begin{aligned} x_t &= \gamma + \gamma_w w_{t-1} + \gamma_v v_t && (3) \\ w_t &= \phi + \phi_w w_{t-1} + \phi_v v_t && (4) \end{aligned}$$

with $v_t \sim N(0, \Omega)$. We are interested in obtaining the following recursive solution:

$$\begin{aligned} y_t &= \theta_y + \theta_{yk} k_t + \theta_{yw} w_t && (5) \\ k_t &= \theta_k + \theta_{kk} k_{t-1} + \theta_{kw} w_{t-1} + \theta_{kv} v_t && (6) \end{aligned}$$

We will refer to Equation (5) as the observation equation, and Equation (6) as the state equation, as is common in the state space econometrics literature. All of the coefficient matrices and the variables k_t are to be determined. However, the derivations make clear how one could also compute perfect foresight solutions for specified paths of x. In (5) we make y_t to depend contemporaneously on w_t and k_t (and a constant, θ_y) — yet it's clear that by employing the state equation (6), one can also make y_t to depend directly on w_{t-1}, k_{t-1}, and v_t along the lines of Watson (1989).

We will proceed in two steps. First, consider the expectational system implied by the law of iterated expectations,

$$A E_t y_{t+1} = B y_t + C E_t x_{t+1} \quad (7)$$

and find a solution of the form

$$\begin{aligned} k_t &= \theta_k + \theta_{kk} k_{t-1} + \theta_{kw} w_{t-1} \\ y_t &= \theta_y + \theta_{yk} k_t + \theta_{yw} w_t \end{aligned}$$

Second, solve for the innovations $\theta_{kv} v_t$ in (6) using

$$A_1 (y_{t+1} - E_t y_{t+1}) = C_1 (x_{t+1} - E_t x_{t+1})$$

a direct implication of Equation (1).

For dimension and sizing information, let $n(A)$ denote the number of rows of A, and $m(A)$ denote the number of columns of A. Let $r(A)$ denote the rank of the matrix A, and let $\lambda_i(A)$ denote the ith eigenvalue of the matrix A. Finally, let A^* denote the complex conjugate transpose of A.

2 Expectational System

A conditional expectation of (1) and (2) leads to a standard rational expectations model that has been much studied in the literature, Equation (7), which we repeat for convenience

$$A E_t y_{t+1} = B y_t + C E_t x_{t+1}$$

A solution to this system allows us to (i) determine a vector of state variables k in the recursive solution (6); (ii) the matrix θ_{kk} in (6); and (iii) the matrix θ_{yk} in (5). These state variables and matrices are not unique, but are determined up to a nonsingular transformation.[1]

We assume there exists a scalar α such that $\det |A - B\alpha| \neq 0$ so that the matrix pencil (A, B) is regular (see Definition 4.2 of Stewart (2001)). Under this condition, Theorem 4.5 of Stewart (2001) establishes that A and B can be jointly decomposed into upper triangular matrices N and J and orthonormal matrices Q and Z such that $Q^*Q = Z^*Z = I$ and[2]

$$\begin{aligned} QNZ &= A \\ QJZ &= B \end{aligned}$$

We also assume that there are no unit roots $|A - B| \neq 0$, although we discuss ways of relaxing this assumption below.

Multiplying the system (6) by Q^*, we see that there is an equivalent block diagonal system

$$\begin{bmatrix} N_{ss} & N_{su} \\ 0 & N_{uu} \end{bmatrix} E_t \begin{bmatrix} s_{t+1} \\ u_{t+1} \end{bmatrix} = \begin{bmatrix} J_{ss} & J_{su} \\ 0 & J_{uu} \end{bmatrix} \begin{bmatrix} s_t \\ u_t \end{bmatrix} + \begin{bmatrix} \Psi_s \\ \Psi_u \end{bmatrix} E_t x_{t+1}$$

with

$$\Psi = \begin{bmatrix} \Psi_s \\ \Psi_u \end{bmatrix} = Q^* C$$

As in some prior literature, we refer to s_t as the stable canonical variable vector and to u_t as the unstable canonical variable vector. These canonical variable vectors are linked to the original vector of endogenous variables y_t through the transformation

$$\begin{bmatrix} s_t \\ u_t \end{bmatrix} = Z y_t \tag{8}$$

[1] For example, if k is the vector of state variables which we determine and G is an invertible square matrix of the same dimension, then we can always replace k with $\widehat{k} = Gk$ and employ the solution matrices $\widehat{\theta}_{kk} = \theta_k G^{-1}$ and $\widehat{\theta}_{yk} = \theta_{yk} G^{-1}$.

[2] If we need the decomposition to be unique, we can choose the Weierstrass form of the decomposition,

$$A - \alpha B = Q (J_A - \alpha J_B) Z^{-1}$$

such that the matrix pencil $A - \alpha B$ is transformed to a block diagonal matrix pencil $J_A - \alpha J_B$, with upper triangular blocks for each eigenvalue. The Weierstrass form is to QZ as Jordan form is to Schur. It similarly shares the numerical instability properties of the Jordan canonical form, so we will use a non-unique upper triangular QZ in practice.

and the partitioning of the block diagonal system is such that there are n_s variables s_t associated generalized eigenvalues $\left|\frac{J_{ii}}{N_{ii}}\right| < 1$.[3] Let $R \equiv Z^*$ be the reverse transformation from y_t to canonical variables.

The separated equation for u_t can be solved forward as

$$\begin{aligned} u_t &= J_{uu}^{-1} N_{uu} E_t u_{t+1} - J_{uu}^{-1} \Psi_u E_t x_{t+1} \\ &= -\sum_{j=1}^{\infty} (J_{uu}^{-1} N_{uu})^{j-1} J_{uu}^{-1} \Psi_u E_t x_{t+j} \end{aligned} \tag{9}$$

provided $\lim_{j\to\infty}(J_{uu}^{-1}N_{uu})^j = 0$, a condition which is satisfied given our ordering of roots and assumption against unit roots. Given the forcing dynamics (5) and (6), this implies that

$$u_t = g_u + g_{uw} w_t \tag{10}$$

where the details of calculating g_u and g_{uw} are described further below, after specification of a driving process.

The separated equation for s_t is

$$E_t s_{t+1} = (N_{ss}^{-1} J_{ss}) s_t - (N_{ss}^{-1} N_{su}) E_t u_{t+1} + (N_{ss}^{-1} J_{su}) u_t + (N_{ss}^{-1} \Psi_s) E_t x_{t+1}$$

and the forcing dynamics (5) and (6) similarly imply

$$E_t s_{t+1} = \left(N_{ss}^{-1} J_{ss}\right) s_t + g_s + g_{sw} w_t \tag{11}$$

where the details of calculating g_s and g_{sw} are described further below.

We have now identified a state vector, $k_t = s_t$ (although this may not be of minimum dimension) and we have determined that the coefficients in the state evolution equation are

$$\begin{aligned} \theta_k &= g_s \\ \theta_{kk} &= N_{ss}^{-1} J_{ss} \\ \theta_{kw} &= g_{sw} \end{aligned}$$

which highlights the fact that state dynamics are governed by stable eigenvalues.

Next, from (8) and (10),

$$\begin{aligned} Z y_t &= \begin{bmatrix} s_t \\ g_u + g_{uw} w_t \end{bmatrix} = \begin{bmatrix} 0 \\ g_u \end{bmatrix} + \begin{bmatrix} I_s \\ 0 \end{bmatrix} s_t + \begin{bmatrix} 0 \\ g_{uw} \end{bmatrix} w_t \\ &\Downarrow \\ y_t &= Z^* \begin{bmatrix} 0 \\ g_u \end{bmatrix} + Z^* \begin{bmatrix} I_s \\ 0 \end{bmatrix} s_t + Z^* \begin{bmatrix} 0 \\ g_{uw} \end{bmatrix} w_t \end{aligned}$$

Using the fact that $k_t = s_t$, this formula provides the coefficients θ_{yk} and θ_{yw},

$$\begin{aligned} \theta_y &= Z^* \begin{bmatrix} 0 \\ g_u \end{bmatrix} \\ \theta_{yk} &= Z^* \begin{bmatrix} I_s \\ 0 \end{bmatrix} \\ \theta_{yw} &= Z^* \begin{bmatrix} 0 \\ g_{uw} \end{bmatrix} \end{aligned}$$

although we have yet to provide the detailed formulas for calculating g_u and g_{uw}, but will do so in Section (1.4) below.

[3] These generalized eigenvalues are the roots of $|Nr - J| = 0$

3 Complete System with Innovations

To this point, we have determined how the stability criterion influences the recursive solution via the selection of the state variable (k_t) and described the determination of the feedback coefficients θ_{yk} and θ_{kk}. Our next task is to determine the nature of the innovations to the system.

3.1 Existence and Uniqueness

We now proceed to solve for coefficient $\{\theta_{kv}\}$ in (6) using Equation (1), which has the implication that

$$A_1 (y_{t+1} - E_t y_{t+1}) = C_1 (x_{t+1} - E_t x_{t+1}) \tag{12}$$

The innovations implied by (9) are

$$u_{t+1} - E_t u_{t+1} = -\sum_{j=1}^{\infty} (J_{uu}^{-1} N_{uu})^{j-1} J_{uu}^{-1} \Psi_u (E_{t+1} x_{t+j+1} - E_t x_{t+j+1}) \tag{13}$$

Equation (12) and the variable transformation (8) imply that

$$A_1 \begin{bmatrix} R_s & R_u \end{bmatrix} \begin{bmatrix} s_{t+1} - E_t s_{t+1} \\ u_{t+1} - E_t u_{t+1} \end{bmatrix} = C_1 (x_{t+1} - E_t x_{t+1})$$

where $R = \begin{bmatrix} R_s & R_u \end{bmatrix}$ is the transformation defined earlier which maps y_t to canonical variables. Let $\widetilde{A} = A_1 R_s$. Using (13) and some rearrangement, we have that

$$\widetilde{A} (s_{t+1} - E_t s_{t+1}) = \sum_{j=0}^{\infty} \widetilde{B_j} (E_{t+1} x_{t+j+1} - E_t x_{t+j+1})$$

where $\widetilde{B_0} = C_1$ and $\widetilde{B_j} = A_1 R_u \left(J_{uu}^{-1} N_{uu}\right)^{j-1} J_{uu}^{-1} \Psi_u$ for $j > 0$. We can see that it will be possible to solve this system for $(s_{t+1} - E_t s_{t+1})$ provided that

$$span\left(\left\{\widetilde{B_j}\right\}_{j=0}^{n(y_t)-n(k_t)}\right) \subset span\left(\widetilde{A}\right) \tag{14}$$

an **existence** condition which is analogous to Equation (41) on page 11 of Sims (2002).[4]

Using the standard SVD,

$$\widetilde{A} = U \Sigma V^*$$

so that

$$\Sigma V^* \eta_{t+1} = U^* \sum_{j=0}^{\infty} \widetilde{B_j} (E_{t+1} x_{t+j+1} - E_t x_{t+j+1})$$

[4] Define $K_j(M,S) \equiv span\begin{pmatrix} S, & MS, & \cdots, & M^{j-1}S \end{pmatrix}$ for square M and conformable S.Then $K_j(M,S)$ is a block Krylov space. Furthermore, observe that $K_j(M,S) \subseteq K_{j+1}(M,S)$. By Lemma 5 of Gutknecht and Schmelzer (2009) and the general properties of Krylov subspaces outlined in Proposition 1 of their paper, $K_j(M,S) = K_{j+1}(M,S)$ for $j \geq n(M)$ and $K_{n(M)}(M,S)$ is the smallest M-invariant subspace that contains S. This is why it is sufficient to check that $K_{n(M)}(M,S) \subseteq span(A_1 R_s)$ instead of $K_\infty(M,S) \subseteq span(A_1 R_s)$ as it might seem.

where $r = r\left(\widetilde{A}\right)$ is the number of non-zero singular values in Σ. It is possible to write this as a partitioned system in a way that makes the location of zeros clear,

$$\begin{bmatrix} \Sigma_{11} & 0_{r,n(k)-r} \\ 0_{n(A_1)-r,r} & 0_{n(A_1)-r,n(k)-r} \end{bmatrix} V^* \eta_{t+1} = \begin{bmatrix} T_1 \\ T_2 \end{bmatrix} \sum_{j=0}^{\infty} \widetilde{B_j} \left(E_{t+1} x_{t+j+1} - E_t x_{t+j+1} \right)$$

for $U^* = \begin{bmatrix} T_1 \\ T_2 \end{bmatrix}$. Obviously if $r\left(\widetilde{A}\right) < n\left(A_1\right)$ so that the T_2 partition exists, then $T_2 = 0$ is a necessary condition for existence. If the existence condition (14) is satisfied, it should be possible to meet this requirement as well. Observe that

$$\begin{aligned} & \begin{bmatrix} U_{11} & U_{12} \\ U_{21} & U_{22} \end{bmatrix} \begin{bmatrix} \Sigma_{11} & 0_{r,n(k)-r} \\ 0_{n(A_1)-r,r} & 0_{n(A_1)-r,n(k)-r} \end{bmatrix} \begin{bmatrix} V_{11} & V_{12} \\ V_{21} & V_{22} \end{bmatrix} \\ = & \begin{bmatrix} U_{11}\Sigma_{11}V_{11} & U_{11}\Sigma_{11}V_{12} \\ U_{21}\Sigma_{11}V_{11} & U_{21}\Sigma_{11}V_{12} \end{bmatrix} \end{aligned}$$

so that only the first $r\left(\widetilde{A}\right)$ columns of U really matter, and we can chose $U_{12} = U_{22} = 0$ so that $T_2 = 0$.

If there are columns of zeros in Σ, then the solution is not unique. We require

$$n\left(k_t\right) = r\left(\widetilde{A}\right) \tag{15}$$

for **uniqueness**.

3.2 Structure of General Solution

Taking the existence condition (14) as satisfied, the fundamental solution for $\eta_{t+1} = (s_{t+1} - E_t s_{t+1})$ is a vector η such that

$$\widetilde{A}\eta_{t+1} = \sum_{j=0}^{\infty} \widetilde{B_j} \left(E_{t+1} x_{t+j+1} - E_t x_{t+j+1} \right)$$

holds as an identity and the $\widetilde{B_j}$ defined previously. The general solution for the innovations is thus

$$\eta_{t+1} = \widetilde{A}^{+} \sum_{j=0}^{\infty} \widetilde{B_j} \left(E_{t+1} x_{t+j+1} - E_t x_{t+j+1} \right) + \left(I - \widetilde{A}^{+}\widetilde{A} \right) \zeta_{t+1}$$

where ζ is a free vector and $\widetilde{A}^{+}$ is the Moore-Penrose generalized inverse of $\widetilde{A}$, which can be constructed using the singular value decomposition. The first component of this expression, which we will refer to as the fundamental component, arises if the solution is unique (in which case $\left(I - \widetilde{A}^{+}\widetilde{A} \right) = 0$) or if there are multiple solutions, so that we concentrate on determining it.

The fundamental component of the solution is then

$$\eta_{t+1}^{f} = V\Sigma^{+}U^* \sum_{j=0}^{\infty} \widetilde{B_j} \left(E_{t+1} x_{t+j+1} - E_t x_{t+j+1} \right)$$

which is the general solution, above, when the non-fundamental component is set to zero. Recall from the specification of the state equation in equation (6) above that

$$k_{t+1} - E_t k_{t+1} = \theta_{kv} v_{t+1}$$

and that $k_t = s_t$ so that

$$\begin{aligned}\theta_{kv} v_{t+1} &= V\Sigma^{+}U^{*}C_1(E_{t+1}x_{t+1} - E_t x_{t+1}) \\ &\quad + V\Sigma^{+}U^{*}\sum_{j=1}^{\infty} A_1 R_u \left(J_{uu}^{-1}N_{uu}\right)^{j-1} J_{uu}^{-1}\Psi_u \left(E_{t+1}x_{t+j+1} - E_t x_{t+j+1}\right)\end{aligned} \tag{16}$$

4 Implications of the driving process

The expression (16) holds for a wide range of stochastic processes for x. We now derive the implications of the driving process (3) and (4). First, it is direct that

$$\begin{aligned}E_{t+1}x_{t+1} - E_t x_{t+1} &= \gamma_v v_{t+1} \\ E_{t+1}x_{t+j+1} - E_t x_{t+j+1} &= \gamma_w \left(\phi_w\right)^{j-1}\phi_v v_{t+1} \text{ for } j > 0\end{aligned} \tag{17}$$

so that it is important to determine the matrix sum that arises when (16) is combined with (17), i.e.,

$$H \equiv \sum_{j=1}^{\infty}\left(J_{uu}^{-1}N_{uu}\right)^{j-1} J_{uu}^{-1}\Psi_u\gamma_w\left(\phi_w\right)^{j-1} \ , \tag{18}$$

which may be shown to be equivalent to the matrix equation

$$H - J_{uu}^{-1}N_{uu}H\phi_w = J_{uu}^{-1}\Psi_u\gamma_w$$

by a small amount of algebra.[5] Because this matrix equation is a contraction mapping in H it can be solved using iterative methods, or as $vec\left(H\right) = \left(I - \left(\phi_w^{*}\otimes J_{uu}^{-1}N_{uu}\right)\right)^{-1} vec\left(J_{uu}^{-1}\Psi_u\gamma_w\right)$. There is a unique solution if and only if $\left(J_{uu}^{-1}N_{uu}\right)$ and $\left(\phi_w^{*}\right)^{-1}$ have disjoint spectra, i.e., $\lambda_i\left(J_{uu}^{-1}N_{uu}\right)\lambda_j\left(\phi_w^{*}\right) \neq 1, \ \forall i,j$ (Jonsson and Kagstrom (2002)), a condition which is clearly satisfied by construction of $J_{uu}^{-1}N_{uu}$ and assumption of stationarity in the driving process.

Proposition 1 *Define $H\left(x\right) = J_{uu}^{-1}\left(N_{uu}x\phi_w - \Psi_u\gamma_w\right)$. Then $H\left(x\right)$ is a contraction mapping.*

Proof. An m by n matrix A is a linear map from x in $\mathbb{R}^n$ to y in $\mathbb{R}^m$, so let $\left\|A\right\|_2$ denote the operator norm induced on A when the L2 norm is applied to both x and y. Using sub-multiplicativity of

[5] For the purpose of this footnote write the sum as

$$\begin{aligned}H &= \sum_{j=1}^{\infty} A^{j-1}BC^{j-1} \\ &= B + ABC + A^2BC^2 + ...\end{aligned}$$

where A, B, and C have distinct meanings from the main text. Then, an implication is that

$$AHC = ABC + A^2BC^2 + ...$$

so that

$$H - AHC = B$$

Hence, we may solve for the elements of H via

$$vec\left(H\right) = \left(I - \left(C^T\otimes A\right)\right)^{-1} vec\left(B\right)$$

operator norms,

$$\begin{aligned}\|H(x)-H(y)\|_2 &= \left\|J_{uu}^{-1}N_{uu}(x-y)\phi_w\right\|_2 \\ &\leq \left\|J_{uu}^{-1}N_{uu}\right\|_2\|x-y\|_2\|\phi_w\|_2 \\ &= \tau\|x-y\|_2\end{aligned}$$

for $\tau = \left\|J_{uu}^{-1}N_{uu}\right\|_2\|\phi_w\|_2 < 1$ because all of the eigenvalues of $J_{uu}^{-1}N_{uu}$ and ϕ_w are stable, and the norm $\|A\|_2$ is the square root of the largest eigenvalue of A^*A (equivalently, the largest singular value of A). ■

Hence, we may write the shocks in the recursive representation as

$$\theta_{kv}v_{t+1} = V\Sigma^+U^*\left(C_1\gamma_v + A_1R_uH\phi_v\right)v_{t+1}$$

where we take H to be defined by (18).

Similarly we can use the solution for H to also determine an important coefficient matrix left unspecified above when we wrote $u_t = g_u + g_{uw}w_t$ in (10), i.e.[6]

$$\begin{aligned}u_t &= -\sum_{j=1}^{\infty}(J_{uu}^{-1}N_{uu})^{j-1}J_{uu}^{-1}\Psi_uE_tx_{t+j} \\ &= -J_{uu}^{-1}\Psi_uE_tx_{t+1} - \sum_{j=2}^{\infty}(J_{uu}^{-1}N_{uu})^{j-1}J_{uu}^{-1}\Psi_uE_tx_{t+j} \\ &= -Hw_t + \left(-\left(I-J_{uu}^{-1}N_{uu}\right)^{-1}J_{uu}^{-1}\Psi_u\gamma - \tilde{N}\phi\right)\end{aligned}$$

where

$$\begin{aligned}\tilde{N} &= \left(I-J_{uu}^{-1}N_{uu}\right)^{-1}J_{uu}^{-1}N_{uu}J_{uu}^{-1}\Psi_u\gamma_w(I-\phi_w)^{-1} - \tilde{N}_2 \\ \text{and } \tilde{N}_2 &= \sum_{j=2}^{\infty}(J_{uu}^{-1}N_{uu})^{j-1}J_{uu}^{-1}\Psi_u\gamma_w(I-\phi_w)^{-1}(\phi_w)^{j-1}\end{aligned}$$

We thus have

$$\begin{aligned}g_u &= -\left(I-J_{uu}^{-1}N_{uu}\right)^{-1}J_{uu}^{-1}\Psi_u\gamma - \tilde{N}\phi \\ g_{uw} &= -H\end{aligned}$$

To compute $\tilde{N}_2$, by the same argument as in Proposition 1, $\tilde{N}_2$ is a contraction mapping. Hence, it can be computed using an iterative method using (see footnote 5)

$$\tilde{N}_2 - (J_{uu}^{-1}N_{uu})\tilde{N}_2\phi_w = J_{uu}^{-1}N_{uu}J_{uu}^{-1}\Psi_u\gamma_w(I-\phi_w)^{-1}\phi_w$$

Alternatively, it can be computed directly via

$$vec\left(\tilde{N}_2\right) = \left(I-\left(\phi_w^*\otimes J_{uu}^{-1}N_{uu}\right)\right)^{-1}vec\left(J_{uu}^{-1}N_{uu}J_{uu}^{-1}\Psi_u\gamma_w(I-\phi_w)^{-1}\phi_w\right)$$

[6] Based on the driving process evolution in (3) and (4), $E_tx_{t+1} = \gamma + \gamma_w w_t$ and $E_tx_{t+j} = \gamma + \gamma_w\sum_{h=1}^{j-1}\phi_w^{h-1}\phi + \gamma_w\phi_w^{j-1}w_t$, for $j > 1$. The summation can be computed as $\sum_{h=1}^{j-1}\phi_w^{h-1} = (I-\phi_w)^{-1}(I-\phi_w^{j-1})$.

Further, we also left unspecified g_s and g_{sw} in (11), but we can now also manipulate

$$\begin{aligned} E_t s_{t+1} &= (N_{ss}^{-1} J_{ss}) s_t + (N_{ss}^{-1} J_{su}) u_t + (N_{ss}^{-1} \Psi_s) E_t x_{t+1} - (N_{ss}^{-1} N_{su}) E_t u_{t+1} \\ &= (N_{ss}^{-1} J_{ss}) s_t + \left(N_{ss}^{-1} J_{su} g_{uw} + N_{ss}^{-1} \Psi_s \gamma_w - N_{ss}^{-1} N_{su} g_{uw} \phi_w \right) w_t \\ &\quad + \left(N_{ss}^{-1} J_{su} g_u + N_{ss}^{-1} \Psi_s \gamma - N_{ss}^{-1} N_{su} (g_u + g_{uw} \phi) \right) \end{aligned}$$

to deduce that

$$\begin{aligned} g_{sw} &= N_{ss}^{-1} \left(J_{su} g_{uw} + \Psi_s \gamma_w - N_{su} g_{uw} \phi_w \right) \\ g_s &= N_{ss}^{-1} \left(J_{su} g_u + \Psi_s \gamma - N_{su} (g_u + g_{uw} \phi) \right) \end{aligned}$$

4.1 The Solution Algorithm and Its Steps

Compute QZ Take the expectational model

$$A E_t y_{t+1} = B y_t + C E_t x_{t+1}$$

and find the decoupled representation

$$\begin{aligned} & \begin{bmatrix} N_{ss} & N_{su} \\ 0 & N_{uu} \end{bmatrix} E_t \begin{bmatrix} s_{t+1} \\ u_{t+1} \end{bmatrix} \\ = & \begin{bmatrix} J_{ss} & J_{su} \\ 0 & J_{uu} \end{bmatrix} \begin{bmatrix} s_t \\ u_t \end{bmatrix} + \begin{bmatrix} \Psi_s \\ \Psi_u \end{bmatrix} E_t x_{t+1} \end{aligned}$$

via

$$\begin{aligned} QNZ &= A \\ QJZ &= B \\ Q\Psi &= C \end{aligned}$$

where the canonical variables are

$$\begin{bmatrix} s_t \\ u_t \end{bmatrix} = Z y_t$$

Compute $\{g_u, g_{uw}, g_s, g_{sw}\}$ Find g_{uw} such that

$$J_{uu}^{-1} N_{uu} g_{uw} \phi_w - g_{uw} = J_{uu}^{-1} \Psi_u \gamma_w$$

Thus one way of computing the matrix $H = -g_{uw}$ is to guess x and iterate until $x = H(x)$. The sum H could also be computed directly by discarding terms with $j > p$ for some large integer p. Alternatively, one may also use tricks involving the vec operator and the Kronecker product to obtain $vec(H) = \left(I - \left(\phi_w^* \otimes J_{uu}^{-1} N_{uu} \right) \right)^{-1} vec\left(J_{uu}^{-1} \Psi_u \gamma_w \right)$, although this requires computing the inverse to a potentially large matrix. Compute

$$g_u = -\left(I - J_{uu}^{-1} N_{uu} \right)^{-1} J_{uu}^{-1} \Psi_u \gamma - \tilde{N} \phi$$

Next, compute

$$\begin{aligned} g_{sw} &= N_{ss}^{-1} \left(J_{su} g_{uw} + \Psi_s \gamma_w - N_{su} g_{uw} \phi_w \right) \\ g_s &= N_{ss}^{-1} \left(J_{su} g_u + \Psi_s \gamma - N_{su} (g_u + g_{uw} \phi) \right) \end{aligned}$$

Compute $\{\theta_k, \theta_{kw}, \theta_{kk}, \theta_y, \theta_{yk}, \theta_{yw}\}$

$$\begin{aligned}
\theta_k &= g_s \\
\theta_{kk} &= N_{ss}^{-1} J_{ss} \\
\theta_{kw} &= g_{sw}
\end{aligned}$$

$$\begin{aligned}
\theta_y &= R \begin{bmatrix} 0 \\ g_u \end{bmatrix} \\
\theta_{yk} &= R \begin{bmatrix} I_s \\ 0 \end{bmatrix} \\
\theta_{yw} &= R \begin{bmatrix} 0 \\ g_{uw} \end{bmatrix}
\end{aligned}$$

Check spanning condition A necessary and sufficient condition for the existence of a solution is

$$span\left(\left\{C_1, \left\{A_{1u} Z_u \left(J_{uu}^{-1} N_{uu}\right)^{j-1} \left(J_{uu}^{-1}\right) \Psi_u\right\}_{j=1}^{n(y)-n(k)}\right\}\right) \subset span\left(A_{1s} Z_s\right)$$

This section closely follows Sims (2002), Section (5). We want to check the spanning condition

$$span\left(B\right) \subset span\left(A\right)$$

for two matrices A and B of dimensions $m \times n_a$ and $m \times n_b$.

If A is full column rank, then the projection matrix $P = A\left(A^* A\right)^{-1} A^*$ exists, and we can project B onto A and check to see that we get the original matrix B back in return

$$\left\| \left(I - A\left(A^* A\right)^{-1} A^* \right) B \right\| < \varepsilon$$

If A is less than full column rank, then the projection matrix as defined does not exist. We need an alternative way of summarizing the column space. Using the SVD once again,

$$A = U \Sigma V^*$$

the columns of U form an orthonormal basis for the subset of $\mathbb{R}^m$ spanned by A. Because U is orthonormal, its projection matrix is just $P = UU^*$. Thus, we should be able to check that

$$\left\| \left(I - UU^* \right) B \right\| < \varepsilon$$

However, if B is not scaled well, then this may not work well in practice. Another option would be to use the SVD on B as well,

$$B = TDW^*$$

so that the columns of T likewise form an orthonormal basis for the subset of $\mathbb{R}^m$ spanned by B. Thus we could check

$$\left\| \left(I - UU^* \right) T \right\| < \varepsilon$$

and this would likely be more numerically stable. Note that this only works with the outer product version of the SVD.

In summary, create the matrices A and B, above, as:

$$\begin{aligned} A &= A_1 R_s \\ B_0 &= C_1 \\ B_j &= A_1 R_u \left(J_{uu}^{-1} N_{uu}\right)^{j-1} \left(J_{uu}^{-1}\right) \Psi_u,\ j > 0 \end{aligned}$$

and then use the SVD to find orthonormal bases for some subsets of $\mathbb{R}^m$. Then use the basis for A to project B_j onto the subspace spanned by A and verify that the matrix B_j is recovered within some tolerance, for all $j \leq n\left(y_t\right) - n\left(k_t\right)$.

Compute errors $\{\theta_{kv}\}$ Use SVD

$$A_1 R_s = U \Sigma V^*$$

to compute

$$\theta_{kv} = V \Sigma^{+} U^* \left(C_1 \gamma_v - A_1 R_u g_{uw} \phi_v\right)$$

B The (pseudo) driving processes and the differential solutions

This appendix shows how to obtain the driving process present in the differential restrictions, along with its evolution. Here we only derive the process up to the 3rd order — however, the general steps outlined below can be followed for any i-th order of interest. The derivation — for the particular processes in the 2nd- and 3rd-order differentials — focuses on presenting the elements of the driving processes as transparently as possible. For computation purpose, there are ways in which we can make the construction of the driving processes more efficient — we describe this alternative representation in a separate technical appendix. For each differential order, we also provide the solution, as this solution will be needed in order to derive the driving processes in the next orders.

Recall that for any i-th differential, the general form of restrictions is

$$\begin{aligned} A_1 d^i z_{t+1} &= B_1 d^i z_t + C_1^{(i)} x_{t+1}^{(i)} && (19) \\ A_2 E_t d^i z_{t+1} &= B_2 d^i z_t + C_2^{(i)} E_t x_{t+1}^{(i)} \ , && (20) \end{aligned}$$

with the driving process evolves according to

$$\begin{aligned} x_t^{(i)} &= \gamma^{(i)} + \gamma_\varsigma^{(i)} \varsigma_{t-1}^{(i)} + \gamma_v^{(i)} v_t^{(i)} && (21) \\ \varsigma_t^{(i)} &= \phi^{(i)} + \phi_\varsigma^{(i)} \varsigma_{t-1}^{(i)} + \phi_v^{(i)} v_t^{(i)} && (22) \end{aligned}$$

The coefficient matrices $A_1 = -G_{z'}$, $A_2 = -F_{z'}$, $B_1 = G_z$, and $B_2 = F_z$ are identical for all i-th differentials. Given these restrictions, the solution for the i-th differential is

$$\begin{aligned} d^i z_t &= \theta_y^{(i)} + \theta_{yk}^{(i)} d^i s_t + \theta_{y\varsigma}^{(i)} \varsigma_t^{(i)} && (23) \\ d^i s_t &= \theta_k^{(i)} + \theta_{kk}^{(i)} d^i s_{t-1} + \theta_{k\varsigma}^{(i)} \varsigma_{t-1}^{(i)} + \theta_{kv}^{(i)} v_t^{(i)} && (24) \end{aligned}$$

In the above, we have directly replaced $k_t^{(i)}$ with $d^i s_t$, the i-th differential of the states. For each i, we need to derive the matrices $C_1^{(i)}$ and $C_2^{(i)}$, the elements of $x_t^{(i)}$, the pseudo or latent states, $\varsigma_t^{(i)}$, along with their evolutions, and the shock process, $v_t^{(i)}$. The coefficient matrices of the state-space solution in (23) and (24) can then be found using the method described in Appendix A.

Since types of equations ("f" or "g") do not matter for the derivation of the driving process, we simply use the fully-expectational system of nonlinear equation

$$0 = E_t \Gamma(z_{t+1}, z_t, \eta_{t+1}) \ ,$$

and define the restrictions in (19) and (20) as

$$0 = \Gamma_{z'} E_t d^i z_{t+1} + \Gamma_z d^i z_t + C^{(i)} E_t x_{t+1}^{(i)}, \tag{25}$$

where

$$\begin{aligned} \Gamma_{z'} &= -A = -\begin{bmatrix} A_1 \\ A_2 \end{bmatrix} \\ \Gamma_z &= B = \begin{bmatrix} B_1 \\ B_2 \end{bmatrix} \\ C^{(i)} &\equiv \begin{bmatrix} C_1^{(i)} \\ C_2^{(i)} \end{bmatrix} \end{aligned}$$

We will use the restrictions in (25) when taking the total differentiation for each i below. Note that: (i) the matrices Γ_j, $j \in \{z', z, \eta'\}$, are Jacobian matrices consisting of the derivatives of Γ with respect to each element of j; (ii) for the first-order differential restrictions, $C^{(1)} = \Gamma_{\eta'}$.

1 The general steps

We can follow these steps to compute the driving process for any ith order differential restrictions (for $i > 1$, since the first order is a special case where there is no driving process beyond η_{t+1}).

STEP 1: **Take the total differentiation of the previous ith-1 differential restriction, yielding the general form in (19) and (20).**

In this step, $x_{t+1}^{(i)}$ contains various cross (tensor) products of the elements of the lower-order differential solutions. That is, if we collect the vector of variables z_{t+1}, z_t and the vector of shocks $\eta_{t+1} = \sigma\varepsilon_{t+1}$ into a stacked vector

$$Y_{t+1} = \begin{bmatrix} z_{t+1} \\ z_t \\ \eta_{t+1} \end{bmatrix},$$

$x_{t+1}^{(i)}$ consists of every (unique) permutation of such cross-products — for example, for $i = 2$, 3, and 4:

$$\begin{aligned} x_{t+1}^{(2)} &= [dY_{t+1} \otimes dY_{t+1}] \ , \\ x_{t+1}^{(2)} &= \begin{bmatrix} d^2Y_{t+1} \otimes dY_{t+1} \\ dY_{t+1} \otimes dY_{t+1} \otimes dY_{t+1} \end{bmatrix} \\ x_{t+1}^{(3)} &= \begin{bmatrix} d^3Y_{t+1} \otimes dY_{t+1} \\ d^2Y_{t+1} \otimes d^2Y_{t+1} \\ d^2Y_{t+1} \otimes dY_{t+1} \otimes dY_{t+1} \\ dY_{t+1} \otimes dY_{t+1} \otimes dY_{t+1} \otimes dY_{t+1} \end{bmatrix} \end{aligned}$$

Alternatively, one can also write $x_{t+1}^{(i)}$ directly in terms of the cross-product of lower-order states and shocks (see Step 2 below). This alternative representation reduces the dimension of $x_{t+1}^{(i)}$ and thus, would be computationally more efficient.

STEP 2: **Given the lower-order differential solutions, determine $x_{t+1}^{(i)}$ as a function of cross-products of the lower-order states and shocks. In the process, determine the new pseudo states, $\varsigma_{t-1}^{(i)}$, and shocks (a function of the original shocks, η_{t+1}), $v_t^{(i)}$.**

This step thus amounts to deriving (21) in the general form above.

STEP 3: **Compute the evolution of these new states in a recursive manner, yielding the form (22)**.

STEP 4: **Collect terms so we have the full system of restrictions (19)-(22), which now can be solved to yield (23) and (24).**

Below, we perform these steps for the 2nd- and 3rd-order differentials.

2 The second order

STEP 1: Elements of the driving process Total differentiating the 1st-order differential restrictions,

$$0 = \Gamma_{z'} E_t dz_{t+1} + \Gamma_z dz_t + \Gamma_{\eta'} E_t d\eta_{t+1}$$

leads to

$$0 = \Gamma_{z'} E_t d^2 z_{t+1} + \Gamma_z d^2 z_t + E_t \delta^{(2)}_{t+1} \ , \tag{26}$$

with

$$\begin{aligned} \delta^{(2)}_{t+1} \equiv\ & \Gamma_{z'z'}(dz_{t+1} \otimes dz_{t+1}) + 2\Gamma_{z'z}(dz_{t+1} \otimes dz_t) + 2\Gamma_{z'\eta'}(dz_{t+1} \otimes d\eta_{t+1}) \\ & + \Gamma_{zz}(dz_t \otimes dz_t) + 2\Gamma_{z\eta'}(dz_t \otimes d\eta_{t+1}) + \Gamma_{\eta'\eta'}(d\eta_{t+1} \otimes d\eta_{t+1}) \end{aligned}$$

The matrices Γ_{jk}, $j,k \in \{z', z, \eta'\}$, are Jacobian matrices consisting of the derivatives of Γ_j matrices with respect to each element of k. That is, $\Gamma^h_{jk} \equiv vec\left(\frac{\partial}{\partial k}\left[\Gamma^h_j\right]^T\right)^T$ for rows $h = 1, ..., n_j$. Hence, one way we can write the restrictions, in terms of the fully-expectational general form in (25), is as follows

$$AE_t d^2 z_{t+1} = Bd^2 z_t + \tilde{C}^{(2)} E_t \tilde{x}^{(2)}_{t+1} \ ,$$

with

$$\tilde{C}^{(2)} \equiv \left[\ \Gamma_{z'z'} \quad 2\Gamma_{z'z} \quad 2\Gamma_{z'\eta'} \quad \Gamma_{zz} \quad 2\Gamma_{z\eta'} \quad \Gamma_{\eta'\eta'}\ \right] \ ,$$

$$\tilde{x}^{(2)}_{t+1} \equiv \begin{bmatrix} dz_{t+1} \otimes dz_{t+1} \\ dz_{t+1} \otimes dz_t \\ dz_{t+1} \otimes d\eta_{t+1} \\ dz_t \otimes dz_t \\ dz_t \otimes d\eta_{t+1} \\ d\eta_{t+1} \otimes d\eta_{t+1} \end{bmatrix}$$

The driving process for the 2nd-order differentials thus consists of the cross (tensor) products of the 1st-order differential solution and its driving process. Though transparent, this way of writing the driving process is, however, not the most efficient way, since $C^{(i)}$ and $x^{(i)}_{t+1}$ can get very large fast. A more efficient way is to write the driving process $x^{(2)}_{t+1}$ directly in terms of the cross products of the *state* variables and the driving process in the 1st-order, which we undertake below.

Given the 1st-order differential solution

$$dz_{t+1} = \Pi ds_{t+1} \tag{27}$$

$$ds_{t+1} = \Phi_s ds_t + \Phi_\eta \eta_{t+1} \ , \tag{28}$$

we have

$$\begin{aligned} dz_{t+1} \otimes dz_{t+1} &= \Pi ds_{t+1} \otimes \Pi ds_{t+1} \\ &= (\Pi\Phi_s ds_t + \Pi\Phi_\eta \eta_{t+1}) \otimes (\Pi\Phi_s ds_t + \Pi\Phi_\eta \eta_{t+1}) \\ &= (\Pi\Phi_s ds_t \otimes \Pi\Phi_s ds_t) + (\Pi\Phi_s ds_t \otimes \Pi\Phi_\eta \eta_{t+1}) + (\Pi\Phi_\eta \eta_{t+1} \otimes \Pi\Phi_s ds_t) \\ &\quad + (\Pi\Phi_\eta \eta_{t+1} \otimes \Pi\Phi_\eta \eta_{t+1}) \\ &= (\Pi\Phi_s \otimes \Pi\Phi_s) vec(ds_t ds_t^T) + (\Pi\Phi_\eta \otimes \Pi\Phi_\eta) vec(\eta_{t+1}\eta_{t+1}^T) \\ &\quad + (\Pi\Phi_\eta \otimes \Pi\Phi_s) vec(ds_t d\eta_{t+1}^T) + (\Pi\Phi_s \otimes \Pi\Phi_\eta) vec(\eta_{t+1} ds_t^T) \end{aligned}$$

In the above $vec(.)$ is the matrix vectorization operator. To get to the fourth line from the third line, we use the following two properties of Kronecker products: (i) for any conformable matrices A, B, C, and D, $(AB \otimes CD) = (A \otimes C)(B \otimes D)$; and (ii) for any two column vectors a and b, $vec(ab^T) = b \otimes a$. We have also utilized the fact that $d\eta_{t+1} = \eta_{t+1}$ and $d^j \eta_{t+1} = 0$ for all $j > 1$. The expression above can be further simplified using commutation and duplication matrices (see Magnus and Neudecker, 2002). Let D_s and D_η be the duplication matrix such as $D_s vech(ds_t ds_t^T) = vec(ds_t ds_t^T)$ and $D_\eta vech(\eta_{t+1}\eta_{t+1}^T) = vec(\eta_{t+1}\eta_{t+1}^T)$, respectively — here, $vech(S)$ is a vectorization operator that chooses only specific elements of the symmetric matrix S (the upper triangular part). Let $K_{s\eta}$ be the commutation matrix such that $vec(\eta_{t+1} ds_t^T) = K_{s\eta} vec(ds_t d\eta_{t+1}^T)$. Hence,

$$\begin{aligned} dz_{t+1} \otimes dz_{t+1} &= (\Pi\Phi_s \otimes \Pi\Phi_s) D_s vech(ds_t ds_t^T) + (\Pi\Phi_\eta \otimes \Pi\Phi_\eta) D_\eta vech(\eta_{t+1}\eta_{t+1}^T) \\ &\quad + \left\{(\Pi\Phi_\eta \otimes \Pi\Phi_s) + (\Pi\Phi_s \otimes \Pi\Phi_\eta) K_{s\eta}\right\} vec(ds_t d\eta_{t+1}^T) \end{aligned}$$

Using similar mechanisms, we also obtain

$$\begin{aligned} dz_{t+1} \otimes dz_t &= \Pi ds_{t+1} \otimes \Pi ds_t \\ &= (\Pi\Phi_s \otimes \Pi) D_s vech(ds_t ds_t^T) + (\Pi\Phi_\eta \otimes \Pi) vec(ds_t d\eta_{t+1}^T) \end{aligned}$$

$$\begin{aligned} dz_{t+1} \otimes \eta_{t+1} &= \Pi ds_{t+1} \otimes \eta_{t+1} \\ &= (\Pi\Phi_s \otimes I) K_{s\eta} vec(ds_t \eta_{t+1}^T) + (\Pi\Phi_\eta \otimes I) D_\eta vech(\eta_{t+1}\eta_{t+1}^T) \end{aligned}$$

$$\begin{aligned} dz_t \otimes dz_t &= \Pi ds_t \otimes \Pi ds_t \\ &= (\Pi \otimes \Pi) D_s vech(ds_t ds_t^T) \end{aligned}$$

$$\begin{aligned} dz_t \otimes \eta_{t+1} &= \Pi ds_t \otimes \eta_{t+1} \\ &= (\Pi \otimes I) vec(\eta_{t+1} ds_t^T) = (\Pi \otimes I) K_{s\eta} vec(ds_t \eta_{t+1}^T) \end{aligned}$$

Based on the above expressions, the driving process is now given by

$$\delta_{t+1}^{(2)} = C^{(2)} x_{t+1}^{(2)} \ ,$$

with

$$x_{t+1}^{(2)} \equiv \begin{bmatrix} vech(ds_t ds_t^T) \\ vech(\eta_{t+1}\eta_{t+1}^T) \\ vec(ds_t \eta_{t+1}^T) \end{bmatrix} \tag{29}$$

and

$$C^{(2)} \equiv \begin{bmatrix} \hat{C}_1^{(2)} & \hat{C}_2^{(2)} & \hat{C}_3^{(2)} \end{bmatrix} \ ,$$

$$\begin{aligned} \hat{C}_1^{(2)} &\equiv \Gamma_{z'z'}(\Pi\Phi_s \otimes \Pi\Phi_s) D_s + 2\Gamma_{z'z}(\Pi\Phi_s \otimes \Pi) D_s \\ &\quad + \Gamma_{zz}(\Pi \otimes \Pi) D_s \ , \end{aligned}$$

$$\begin{aligned} \hat{C}_2^{(2)} &\equiv \Gamma_{z'z'}(\Pi\Phi_\eta \otimes \Pi\Phi_\eta) D_\eta + 2\Gamma_{z'\eta'}(\Pi\Phi_\eta \otimes I) D_\eta \\ &\quad + \Gamma_{\eta'\eta'} D_\eta \ , \end{aligned}$$

$$\begin{aligned} \hat{C}_3^{(2)} &\equiv \Gamma_{z'z'} \left\{(\Pi\Phi_\eta \otimes \Pi\Phi_s) + (\Pi\Phi_s \otimes \Pi\Phi_\eta) K_{s\eta}\right\} \\ &\quad + 2\Gamma_{z'z}(\Pi\Phi_\eta \otimes \Pi) + 2\Gamma_{z'\eta'}(\Pi\Phi_s \otimes I) K_{s\eta} \\ &\quad + 2\Gamma_{z\eta'}(\Pi \otimes I) K_{s\eta} \end{aligned}$$

The matrices $C_1^{(2)}$ and $C_2^{(2)}$ are the first n_g rows of matrix $C^{(2)}$ and the last n_f rows of matrix $C^{(2)}$, respectively.

STEP 2: Driving process as a function of (pseudo) states and shocks Given $x_t^{(2)}$ in (29) the natural candidates for the pseudo state and shock vectors — $\varsigma_{t-1}^{(2)}$ and $v_t^{(2)}$ in (21) — are

$$\begin{aligned}\tilde{\varsigma}_{t-1}^{(2)} &= vech(ds_{t-1}ds_{t-1}^T)\\ \tilde{v}_t^{(2)} &= \begin{bmatrix} vech(\eta_t\eta_t^T) \\ vec(ds_{t-1}\eta_t^T) \end{bmatrix} .\end{aligned}$$

However, for practical purpose, it's better to define them in terms of deviations from their unconditional expectations, resulting in a *demeaned* system. That is, we instead define

$$\varsigma_{t-1}^{(2)} = vech(ds_{t-1}ds_{t-1}^T - \Omega_{ss}) \tag{30}$$

$$v_t^{(2)} = \begin{bmatrix} vech(\eta_t\eta_t^T - \Omega_{\eta\eta}) \\ vec(ds_{t-1}\eta_t^T) \end{bmatrix} , \tag{31}$$

where $\Omega_{ss} = E[ds_t ds_t^T]$ and $\Omega_{\eta\eta} = E[\eta_t\eta_t^T]$.[7] Hence, we can write

$$x_t^{(2)} = \begin{bmatrix} vech(ds_{t-1}ds_{t-1}^T) \\ vech(\eta_t\eta_t^T) \\ vec(ds_{t-1}\eta_t^T) \end{bmatrix} = \gamma^{(2)} + \gamma_\varsigma^{(2)}\varsigma_{t-1}^{(2)} + \gamma_v^{(2)}v_t^{(2)} , \tag{32}$$

with the coefficients given by

$$\begin{aligned}\gamma^{(2)} &\equiv \begin{bmatrix} vech(\Omega_{ss}) \\ vech(\Omega_{\eta\eta}) \\ 0 \end{bmatrix}\\ \gamma_\varsigma^{(2)} &\equiv \begin{bmatrix} I \\ 0 \\ 0 \end{bmatrix}\\ \gamma_v^{(2)} &= \begin{bmatrix} 0 & 0 \\ I & 0 \\ 0 & I \end{bmatrix}\end{aligned}$$

Each I above is an identity matrix of conformable dimension.

STEP 3: Evolution of the pseudo states The evolution of the pseudo states, $\varsigma_t^{(2)}$, can be found using the state-evolution equation in the 1st-order differential solution. Given $ds_t = \Phi_s ds_{t-1} + \Phi_\eta\eta_t$,

$$\begin{aligned}vech(ds_t ds_t) &= D_s^{-1}(\Phi_s \otimes \Phi_s)D_s vech(ds_{t-1}ds_{t-1}) + D_s^{-1}(\Phi_\eta \otimes \Phi_\eta)D_\eta vech(\eta_t\eta_t^T)\\ &\quad + D_s^{-1}\left[(\Phi_\eta \otimes \Phi_s) + (\Phi_s \otimes \Phi_\eta)K_{s\eta}\right] vec(ds_{t-1}\eta_t^T)\end{aligned}$$

In deriving the above, we have used the property that for conformable matrices A, B, and C, $vec(ABC) = (C^T \otimes A)vec(B)$.[8] Taking the unconditional expectation of the above expression yields

$$vech(\Omega_{ss}) = D_s^{-1}(\Phi_s \otimes \Phi_s)D_s vech(\Omega_{ss}) + D_s^{-1}(\Phi_\eta \otimes \Phi_\eta)D_\eta vech(\Omega_{\eta\eta}) ,$$

[7] Since η_t is i.i.d., $E[ds_{t-1}\eta_t^T] = 0$.

[8] See Magnus and Neudecker (2002).

which can be used to compute Ω_{ss} given $\Omega_{\eta\eta}$.[9] Combining the two expressions above, the evolution of $\varsigma_t^{(2)}$ is

$$\varsigma_t^{(2)} = \phi^{(2)} + \phi_\varsigma^{(2)}\varsigma_{t-1}^{(2)} + \phi_v^{(2)} v_t^{(2)} \ , \tag{33}$$

with the coefficients given by

$$\begin{aligned}
\phi^{(2)} &= 0 \\
\phi_\varsigma^{(2)} &= D_s^{-1}(\Phi_s \otimes \Phi_s) D_s \\
\phi_v^{(2)} &= \left[\begin{array}{cc} D_s^{-1}(\Phi_\eta \otimes \Phi_\eta) D_\eta & D_s^{-1}\left[(\Phi_\eta \otimes \Phi_s) + (\Phi_s \otimes \Phi_\eta) K_{s\eta}\right] \end{array} \right]
\end{aligned}$$

The elements of the pseudo state and the shock vectors, $\varsigma_t^{(2)}$ and $v_t^{(2)}$, are given in (30) and (31).

STEP 4: Compute the differential solution Given the 2nd-order differential restrictions, the driving process (29) and its evolution in (32), the pseudo state and shock vectors (30) and (31), and the evolution of the pseudo states in (33), the 2nd-order differential solution is

$$d^2 z_t = \theta_y^{(2)} + \left[\begin{array}{cc} \theta_{yk}^{(2)} & \theta_{y\varsigma}^{(2)} \end{array} \right] \left[\begin{array}{c} d^2 s_t \\ vech(ds_t ds_t^T - \Omega_{ss}) \end{array} \right]$$

with the states evolve according to

$$\begin{aligned}
\left[\begin{array}{c} d^2 s_t \\ vech(ds_t ds_t^T - \Omega_{ss}) \end{array} \right] &= \left[\begin{array}{c} \theta_k^{(2)} \\ 0 \end{array} \right] + \left[\begin{array}{cc} \theta_{kk}^{(2)} & \theta_{k\varsigma}^{(2)} \\ 0 & \phi_\varsigma^{(2)} \end{array} \right] \left[\begin{array}{c} d^2 s_{t-1} \\ vech(ds_{t-1} ds_{t-1}^T - \Omega_{ss}) \end{array} \right] \\
&\quad + \left[\begin{array}{c} \theta_{kv}^{(2)} \\ \phi_v^{(2)} \end{array} \right] \left[\begin{array}{c} vech(\eta_t \eta_t^T - \Omega_{\eta\eta}) \\ vec(ds_{t-1} \eta_t^T) \end{array} \right] .
\end{aligned}$$

The coefficients $\phi_\varsigma^{(2)}$ and $\phi_v^{(2)}$ have been computed above. Since the matrices A_1, A_2, B_1, and A_2 in the general form of differential restrictions in (19) and (20) are identical for all i-th orders, $\theta_{yk}^{(2)}$ and $\theta_{kk}^{(2)}$ are identical to their 1st-order counterparts and hence, are known: $\theta_{yk}^{(2)} = \Pi$ and $\theta_{kk}^{(2)} = \Phi_s$. The rest of the coefficients — $\theta_y^{(2)}$, $\theta_{y\varsigma}^{(2)}$, $\theta_k^{(2)}$, $\theta_{k\varsigma}^{(2)}$, and $\theta_{kv}^{(2)}$ — can be computed using the linear solution method described in Appendix A.

We can modify the state-space solution one step further by writing $d^2 s_t$ in terms of deviation from its unconditional expectation. Notice that

$$\begin{aligned}
E[d^2 s_t] &= \theta_k^{(2)} + \theta_{kk}^{(2)} E[d^2 s_{t-1}] \\
&\Downarrow \\
E[d^2 s] &= \left(I - \theta_{kk}^{(2)}\right)^{-1} \theta_k^{(2)} \\
&= \left(I - \Phi_s\right)^{-1} \theta_k^{(2)}
\end{aligned}$$

Hence, the *fully-demeaned* state-space solution is

$$d^2 z_t = E[d^2 z] + \left[\begin{array}{cc} \theta_{yk}^{(2)} & \theta_{y\varsigma}^{(2)} \end{array} \right] \left[\begin{array}{c} d^2 s_t - E[d^2 s] \\ vech(ds_t ds_t^T - \Omega_{ss}) \end{array} \right] \tag{34}$$

[9] An equivalent representation is

$$vec(\Omega_{ss}) = (I - \Phi_s \otimes \Phi_s)^{-1}(\Phi_\eta \otimes \Phi_\eta) vec(\Omega_{\eta\eta})$$

$$\begin{bmatrix} d^2 s_t - E[d^2 s] \\ vech(ds_t ds_t^T - \Omega_{ss}) \end{bmatrix} = \begin{bmatrix} \theta_{kk}^{(2)} & \theta_{k\varsigma}^{(2)} \\ 0 & \phi_\varsigma^{(2)} \end{bmatrix} \begin{bmatrix} d^2 s_{t-1} - E[d^2 s] \\ vech(ds_{t-1} ds_{t-1}^T - \Omega_{ss}) \end{bmatrix} + \begin{bmatrix} \theta_{kv}^{(2)} \\ \phi_v^{(2)} \end{bmatrix} \begin{bmatrix} vech(\eta_t \eta_t^T - \Omega_{\eta\eta}) \\ vec(ds_{t-1} \eta_t^T) \end{bmatrix} \tag{35}$$

where $E[d^2 z] \equiv \theta_y^{(2)} + \theta_{yk}^{(2)} E[d^2 s]$ is the unconditional expectation of the second differential, $d^2 z_t$.

3 The third order

For the 3rd order derivation (relevant for STEPS 2 and 3), it's more convenient to use the *non-demeaned* version of the 2nd-order differential solution. That is, from (34) and (35),

$$d^2 z_t = \theta_z^{(2)} + \theta_{z\xi}^{(2)} \tilde{\xi}_t^{(2)} \tag{36}$$

$$\tilde{\xi}_t^{(2)} = \theta_\xi^{(2)} + \theta_{\xi\xi}^{(2)} \tilde{\xi}_{t-1}^{(2)} + \theta_{\xi v}^{(2)} \tilde{v}_t^{(2)} \ , \tag{37}$$

with

$$\tilde{\xi}_t^{(2)} \equiv \begin{bmatrix} d^2 s_t \\ vech(ds_t ds_t^T) \end{bmatrix}$$

$$\tilde{v}_t^{(2)} \equiv \begin{bmatrix} vech(\eta_t \eta_t^T) \\ vec(ds_{t-1} \eta_t^T) \end{bmatrix}$$

$$\theta_{z\xi}^{(2)} \equiv \begin{bmatrix} \theta_{yk}^{(2)} & \theta_{y\varsigma}^{(2)} \end{bmatrix}$$

$$\theta_z^{(2)} \equiv E[d^2 z] - \theta_{z\xi}^{(2)} \begin{bmatrix} E[d^2 s] \\ vech(\Omega_{ss}) \end{bmatrix}$$

$$\theta_{\xi\xi}^{(2)} \equiv \begin{bmatrix} \theta_{kk}^{(2)} & \theta_{k\varsigma}^{(2)} \\ 0 & \phi_\varsigma^{(2)} \end{bmatrix}$$

$$\theta_{\xi v}^{(2)} \equiv \begin{bmatrix} \theta_{kv}^{(2)} \\ \phi_v^{(2)} \end{bmatrix}$$

$$\theta_\xi^{(2)} \equiv \left(I - \theta_{\xi\xi}^{(2)} \right) \begin{bmatrix} E[d^2 s] \\ vech(\Omega_{ss}) \end{bmatrix} - \theta_{\xi v}^{(2)} \begin{bmatrix} vech(\Omega_{\eta\eta}) \\ 0 \end{bmatrix}$$

STEP 1: Elements of the driving process Total differentiating the 2nd-order differential restrictions in (26) leads to:[10]

$$0 = \Gamma_{z'} E_t d^3 z_{t+1} + \Gamma_z d^3 z_t + E_t \delta_{t+1}^{(3)} \ , \tag{38}$$

[10] We utilize the following fact: $d\eta_{t+1} = \eta_{t+1}$ and $d^j \eta_{t+1} = 0$ for all $j > 1$.

with

$$\begin{aligned}\delta^{(3)}_{t+1} \equiv\ & 3\Gamma_{z'z'}\left(d^2z_{t+1}\otimes dz_{t+1}\right)+3\Gamma_{z'z}\left(d^2z_{t+1}\otimes dz_t\right)+3\Gamma_{z'\eta'}\left(d^2z_{t+1}\otimes d\eta_{t+1}\right)\\ &+3\Gamma_{zz'}\left(d^2z_t\otimes dz_{t+1}\right)+3\Gamma_{zz}\left(d^2z_t\otimes dz_t\right)+3\Gamma_{z\eta'}\left(d^2z_t\otimes \eta_{t+1}\right)\\ &+\Gamma_{z'z'z'}\left(dz_{t+1}\otimes dz_{t+1}\otimes dz_{t+1}\right)+3\Gamma_{z'z'z}\left(dz_{t+1}\otimes dz_{t+1}\otimes dz_t\right)\\ &+3\Gamma_{z'z'\eta'}\left(dz_{t+1}\otimes dz_{t+1}\otimes d\eta_{t+1}\right)+3\Gamma_{z'zz}\left(dz_{t+1}\otimes dz_t\otimes dz_t\right)\\ &+6\Gamma_{z'z\eta'}\left(dz_{t+1}\otimes dz_t\otimes d\eta_{t+1}\right)+3\Gamma_{z'\eta'\eta'}\left(dz_{t+1}\otimes d\eta_{t+1}\otimes d\eta_{t+1}\right)\\ &+\Gamma_{zzz}\left(dz_t\otimes dz_t\otimes dz_t\right)+3\Gamma_{zz\eta'}\left(dz_t\otimes dz_t\otimes d\eta_{t+1}\right)\\ &+3\Gamma_{z\eta'\eta'}\left(dz_t\otimes d\eta_{t+1}\otimes d\eta_{t+1}\right)+\Gamma_{\eta'\eta'\eta'}\left(d\eta_{t+1}\otimes d\eta_{t+1}\otimes d\eta_{t+1}\right)\end{aligned} \tag{39}$$

The matrices $\Gamma_{jkl},\ j,k,l\in\{z',z,\eta'\}$, are Jacobian matrices consisting of the derivatives of Γ_{jk} matrices with respect to each element of l. That is, $\Gamma^h_{jkl}\equiv vec\left(\frac{\partial}{\partial l}\left[\Gamma^h_{jk}\right]^T\right)^T$ for rows $h=1,...,n_j$. As in the 2nd order, the driving process is function of the cross (tensor) products of all lower-order solutions and driving processes — for the 3rd order, the relevant lower-order solutions are the 1st- and 2nd-order differential solutions.

Following the derivation of $x^{(2)}_{t+1}$ in Section 2, we also represent $x^{(3)}_{t+1}$ directly in terms of the cross products of lower-order (1st and 2nd) state variables and shocks. Given the 1st- and 2nd-order differential solutions in (27), (36), and (37), all the cross-product terms in (39) can be written as functions of various cross products involving ds_t, η_{t+1}, $\tilde{\xi}^{(2)}_t$, and $\tilde{v}^{(2)}_{t+1}$. That is, one can obtain the driving process as

$$\delta^{(3)}_{t+1}=C^{(3)}x^{(3)}_{t+1}\ ,$$

with

$$x^{(3)}_{t+1}\equiv\begin{bmatrix} ds_t\\ vec\left(ds_t\tilde{\xi}^{(2)T}_t\right)\\ \eta_{t+1}\\ vec\left(\eta_{t+1}\tilde{\xi}^{(2)T}_t\right)\\ vec\left(ds_t\tilde{v}^{(2)T}_{t+1}\right)\\ vec\left(\eta_{t+1}\tilde{v}^{(2)T}_{t+1}\right)\end{bmatrix} \tag{40}$$

and

$$C^{(3)}\equiv\begin{bmatrix}\hat{C}^{(3)}_1 & \hat{C}^{(3)}_2 & \hat{C}^{(3)}_3 & \hat{C}^{(3)}_4 & \hat{C}^{(3)}_5 & \hat{C}^{(3)}_6\end{bmatrix}\text{, with}$$

$$\begin{aligned}\hat{C}^{(3)}_1 \equiv\ & \left(\theta^{(2)}_z\otimes\Pi\right)\Phi_s+\left(\theta^{(2)}_{z\xi}\theta^{(2)}_{\xi}\otimes\Pi\Phi_s\right)+\left(\theta^{(2)}_z\otimes\Pi\right)+\left(\theta^{(2)}_{z\xi}\theta^{(2)}_{\xi}\otimes\Pi\right)\\ &+\left(\theta^{(2)}_z\otimes\Pi\Phi_s\right)+\left(\theta^{(2)}_z\otimes\Pi\right)\end{aligned}$$

$$\begin{aligned}\hat{C}^{(3)}_2 \equiv\ & \left(\theta^{(2)}_{z\xi}\theta^{(2)}_{\xi\xi}\otimes\Pi\Phi_s\right)+\left(\theta^{(2)}_{z\xi}\theta^{(2)}_{\xi\xi}\otimes\Pi\right)+\left(\theta^{(2)}_{z\xi}\otimes\Pi\Phi_s\right)+\left(\theta^{(2)}_{z\xi}\otimes\Pi\right)\\ &+(\gamma_{z'z'}\otimes\Pi\Phi_s)+(\gamma_{z'z'}\otimes\Pi)+(\gamma_{z'z}\otimes\Pi)+(\gamma_{zz}\otimes\Pi)\end{aligned}$$

$$\begin{aligned}\hat{C}^{(3)}_3 \equiv\ & \left(\theta^{(2)}_z\otimes\Pi\right)\Phi_\eta+\left(\theta^{(2)}_{z\xi}\theta^{(2)}_{\xi}\otimes\Pi\Phi_\eta\right)+\left(\theta^{(2)}_z\otimes I_\eta\right)+\left(\theta^{(2)}_{z\xi}\theta^{(2)}_{\xi}\otimes I_\eta\right)\\ &+\left(\theta^{(2)}_z\otimes\Pi\Phi_\eta\right)+\left(\theta^{(2)}_z\otimes I_\eta\right)\end{aligned}$$

$$\begin{aligned}\hat{C}_4^{(3)} &\equiv \left(\theta_{z\xi}^{(2)}\theta_{\xi\xi}^{(2)} \otimes \Pi\Phi_\eta\right) + \left(\theta_{z\xi}^{(2)}\theta_{\xi\xi}^{(2)} \otimes I_\eta\right) + \left(\theta_{z\xi}^{(2)} \otimes \Pi\Phi_\eta\right) + \left(\theta_{z\xi}^{(2)} \otimes I_\eta\right) \\ &\quad + (\gamma_{z'z'} \otimes \Pi\Phi_\eta) + (\gamma_{z'z'} \otimes I_\eta) + (\gamma_{z'z} \otimes I_\eta) + (\gamma_{zz} \otimes I_\eta)\end{aligned}$$

$$\begin{aligned}\hat{C}_5^{(3)} &\equiv \left(\theta_{z\xi}^{(2)}\theta_{\xi v}^{(2)} \otimes \Pi\Phi_s\right) + \left(\theta_{z\xi}^{(2)}\theta_{\xi v}^{(2)} \otimes \Pi\right) + (\alpha_{z'z'} \otimes \Pi\Phi_s) + (\alpha_{z'z'} \otimes \Pi) \\ &\quad + (\alpha_{z'z} \otimes \Pi)\end{aligned}$$

$$\begin{aligned}\hat{C}_6^{(3)} &\equiv \left(\theta_{z\xi}^{(2)}\theta_{\xi v}^{(2)} \otimes \Pi\Phi_\eta\right) + \left(\theta_{z\xi}^{(2)}\theta_{\xi v}^{(2)} \otimes I_\eta\right) + (\alpha_{z'z'} \otimes \Pi\Phi_\eta) + (\alpha_{z'z'} \otimes I_\eta) \\ &\quad + (\alpha_{z'z} \otimes I_\eta) + \left(\alpha_{z'\eta'} \otimes I_\eta\right) + \left(\alpha_{z\eta'} \otimes I_\eta\right) + \left(\alpha_{\eta'\eta'} \otimes I_\eta\right)\end{aligned}$$

In the above, we use I_η and I_s to represent $n_\eta \times n_\eta$ and $n_s \times n_s$ identity matrices, respectively. The superscript ^{T} denotes matrix transpose. Also, we have used the following compacted representations of various cross products present in the 2nd-order driving process:

$$\begin{aligned}dz_{t+1} \otimes dz_{t+1} &= \gamma_{z'z'}\tilde{\xi}_t^{(2)} + \alpha_{z'z'}\tilde{v}_{t+1}^{(2)}\ , \\ \gamma_{z'z'} &\equiv \left[\begin{array}{cc} 0 & (\Pi\Phi_s \otimes \Pi\Phi_s)D_s \end{array}\right] \\ \alpha_{z'z'} &\equiv \left[\begin{array}{cc} (\Pi\Phi_\eta \otimes \Pi\Phi_\eta)D_\eta & (\Pi\Phi_\eta \otimes \Pi\Phi_s) + (\Pi\Phi_s \otimes \Pi\Phi_\eta)K_{s\eta} \end{array}\right]\end{aligned}$$

$$\begin{aligned}dz_{t+1} \otimes dz_t &= \gamma_{z'z}\tilde{\xi}_t^{(2)} + \alpha_{z'z}\tilde{v}_{t+1}^{(2)}\ , \\ \gamma_{z'z} &\equiv \left[\begin{array}{cc} 0 & (\Pi\Phi_s \otimes \Pi)D_s \end{array}\right] \\ \alpha_{z'z} &\equiv \left[\begin{array}{cc} 0 & (\Pi\Phi_\eta \otimes \Pi) \end{array}\right]\end{aligned}$$

$$\begin{aligned}dz_{t+1} \otimes \eta_{t+1} &= \alpha_{z'\eta'}\tilde{v}_{t+1}^{(2)}\ , \\ \alpha_{z'\eta'} &\equiv \left[\begin{array}{cc} (\Pi\Phi_\eta \otimes I)D_\eta & (\Pi\Phi_s \otimes I)K_{s\eta} \end{array}\right]\end{aligned}$$

$$\begin{aligned}dz_t \otimes dz_t &= \gamma_{zz}\tilde{\xi}_t^{(2)}\ , \\ \gamma_{zz} &\equiv \left[\begin{array}{cc} 0 & (\Pi \otimes \Pi)D_s \end{array}\right]\end{aligned}$$

$$\begin{aligned}dz_t \otimes \eta_{t+1} &= \alpha_{z\eta'}\tilde{v}_{t+1}^{(2)}\ , \\ \alpha_{z\eta'} &\equiv \left[\begin{array}{cc} 0 & (\Pi \otimes I)K_{s\eta} \end{array}\right]\end{aligned}$$

$$\begin{aligned}\eta_{t+1} \otimes \eta_{t+1} &= \alpha_{\eta'\eta'}\tilde{v}_{t+1}^{(2)}\ , \\ \alpha_{\eta'\eta'} &\equiv \left[\begin{array}{cc} D_\eta & 0 \end{array}\right]\end{aligned}$$

STEP 2: Driving process as a function of (pseudo) states and shocks Given $x_t^{(3)}$ in (40) the natural candidates for the pseudo state and shock vectors — $\varsigma_{t-1}^{(3)}$ and $v_t^{(3)}$ in (21) — are

$$\begin{aligned}\tilde{\varsigma}_{t-1}^{(3)} &= \begin{bmatrix} ds_{t-1} \\ vec\left(ds_{t-1}\tilde{\xi}_{t-1}^{(2)T}\right) \end{bmatrix} \\ \tilde{v}_t^{(3)} &= \begin{bmatrix} \eta_t \\ vec\left(\eta_t\tilde{\xi}_{t-1}^{(2)T}\right) \\ vec\left(ds_{t-1}\tilde{v}_t^{(2)T}\right) \\ vec\left(\eta_t\tilde{v}_t^{(2)T}\right) \end{bmatrix}.\end{aligned}$$

As in 2nd order, however, we instead define them in terms of deviations from unconditional expectations, resulting in a *demeaned* system. That is, we instead define

$$\varsigma_{t-1}^{(3)} = \begin{bmatrix} ds_{t-1} \\ vec\left(ds_{t-1}\tilde{\xi}_{t-1}^{(2)T} - \Omega_{\xi s}\right) \end{bmatrix} \tag{41}$$

$$v_t^{(3)} = \tilde{v}_t^{(3)} = \begin{bmatrix} \eta_t \\ vec\left(\eta_t\tilde{\xi}_{t-1}^{(2)T}\right) \\ vec\left(ds_{t-1}\tilde{v}_t^{(2)T}\right) \\ vec\left(\eta_t\tilde{v}_t^{(2)T}\right) \end{bmatrix}, \tag{42}$$

where the matrix $\Omega_{\xi s} = E\left[ds_{t-1}\tilde{\xi}_{t-1}^{(2)T}\right]$ is still to be determined.[11] Hence, we can write the driving process as

$$x_t^{(3)} \equiv \begin{bmatrix} ds_{t-1} \\ vec\left(ds_{t-1}\tilde{\xi}_{t-1}^{(2)T}\right) \\ \eta_t \\ vec\left(\eta_t\tilde{\xi}_{t-1}^{(2)T}\right) \\ vec\left(ds_{t-1}\tilde{v}_t^{(2)T}\right) \\ vec\left(\eta_t\tilde{v}_t^{(2)T}\right) \end{bmatrix} = \gamma^{(3)} + \gamma_\varsigma^{(3)}\varsigma_{t-1}^{(3)} + \gamma_v^{(3)}v_t^{(3)}, \tag{43}$$

with the coefficients given by

$$\gamma^{(3)} \equiv \begin{bmatrix} 0 \\ vec\left(\Omega_{\xi s}\right) \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

$$\gamma_\varsigma^{(3)} \equiv \begin{bmatrix} I_s & 0 \\ 0 & I_{s\xi^{(2)}} \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}$$

[11] All elements of $\tilde{v}_t^{(3)}$ have zero unconditional expectations.

$$\gamma_v^{(3)} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ I_\eta & 0 & 0 & 0 \\ 0 & I_{\eta\xi^{(2)}} & 0 & 0 \\ 0 & 0 & I_{sv^{(2)}} & 0 \\ 0 & 0 & 0 & I_{\eta v^{(2)}} \end{bmatrix} .$$

We have used I_x to denote a $n_x \times n_x$ identity matrix (n_x is the size of the column vector x) and I_{xy} to denote a $(n_x * n_y) \times (n_x * n_y)$ identity matrix. Various zero vectors or sub-matrices above are of conformable dimensions.

STEP 3: Evolution of the pseudo states The evolution of the first element of $\varsigma_t^{(3)}$, ds_t, is already known from the 1st-order solution. The evolution of the second element can be found using the state-evolution equations in the 1st- and 2nd-order differential solutions in (28) and (37):

$$\begin{aligned} vec\left(ds_t\tilde{\xi}_t^{(2)T}\right) &= (\theta_\xi^{(2)} \otimes \Phi_s)ds_{t-1} + (\theta_{\xi\xi}^{(2)} \otimes \Phi_s)vec\left(ds_{t-1}\tilde{\xi}_{t-1}^{(2)T}\right) \\ &+(\theta_\xi^{(2)} \otimes \Phi_\eta)\eta_t + (\theta_{\xi\xi}^{(2)} \otimes \Phi_\eta)vec\left(\eta_t\tilde{\xi}_{t-1}^{(2)T}\right) \\ &+(\theta_{\xi v}^{(2)} \otimes \Phi_s)vec\left(ds_{t-1}\tilde{v}_t^{(2)T}\right) + (\theta_{\xi v}^{(2)} \otimes \Phi_\eta)vec\left(\eta_t\tilde{v}_t^{(2)T}\right) \end{aligned}$$

Taking the unconditional expectation of the above expression yields

$$vec(\Omega_{\xi s}) = (\theta_{\xi\xi}^{(2)} \otimes \Phi_s)vec(\Omega_{\xi s}) ,$$

which implies $\Omega_{\xi s} = 0$. The evolution of $\varsigma_t^{(3)}$ is then

$$\varsigma_t^{(3)} = \phi^{(3)} + \phi_\varsigma^{(3)}\varsigma_{t-1}^{(3)} + \phi_v^{(3)}v_t^{(3)} , \tag{44}$$

with the coefficients given by

$$\phi^{(3)} = 0$$

$$\phi_\varsigma^{(3)} = \begin{bmatrix} \Phi_s & 0 \\ (\theta_\xi^{(2)} \otimes \Phi_s) & (\theta_{\xi\xi}^{(2)} \otimes \Phi_s) \end{bmatrix}$$

$$\phi_v^{(3)} = \begin{bmatrix} \Phi_\eta & 0 & 0 & 0 \\ (\theta_\xi^{(2)} \otimes \Phi_\eta) & (\theta_{\xi\xi}^{(2)} \otimes \Phi_\eta) & (\theta_{\xi v}^{(2)} \otimes \Phi_s) & (\theta_{\xi v}^{(2)} \otimes \Phi_\eta) \end{bmatrix}$$

The elements of the pseudo state and the shock vectors, $\varsigma_t^{(3)}$ and $v_t^{(3)}$, are given in (41) and (42).

STEP 4: Compute the differential solution Given the 3rd-order differential restrictions, the driving process (40) and its evolution in (43), the pseudo state and shock vectors (41) and (42), and the evolution of the pseudo states in (44), the 3rd-order differential solution is

$$d^3z_t = \theta_y^{(3)} + \begin{bmatrix} \theta_{yk}^{(3)} & \theta_{y\varsigma}^{(3)} \end{bmatrix} \begin{bmatrix} d^3s_t \\ \varsigma_t^{(3)} \end{bmatrix}$$

with the states evolve according to

$$\begin{bmatrix} d^3s_t \\ \varsigma_t^{(3)} \end{bmatrix} = \begin{bmatrix} \theta_k^{(3)} \\ \phi^{(3)} \end{bmatrix} + \begin{bmatrix} \theta_{kk}^{(3)} & \theta_{k\varsigma}^{(3)} \\ 0 & \phi_\varsigma^{(3)} \end{bmatrix} \begin{bmatrix} d^3s_{t-1} \\ \varsigma_{t-1}^{(3)} \end{bmatrix} + \begin{bmatrix} \theta_{kv}^{(3)} \\ \phi_v^{(3)} \end{bmatrix} v_t^{(3)}$$

The coefficients $\phi^{(3)} = 0$, $\phi_{\varsigma}^{(3)}$ and $\phi_{v}^{(3)}$ have been computed above. Since the matrices A_1, A_2, B_1, and A_2 in the general form of differential restrictions in (19) and (20) are identical for all i-th orders, $\theta_{yk}^{(3)}$ and $\theta_{kk}^{(3)}$ are identical to their 1st-order counterparts and hence, are known: $\theta_{yk}^{(3)} = \Pi$ and $\theta_{kk}^{(3)} = \Phi_s$. The rest of the coefficients — $\theta_{y}^{(3)}$, $\theta_{y\varsigma}^{(3)}$, $\theta_{k}^{(3)}$, $\theta_{k\varsigma}^{(3)}$, and $\theta_{kv}^{(3)}$ — can be computed using the linear solution method described in Appendix A.

We can modify the state-space solution further by writing $d^3 s_t$ in terms of deviation from its unconditional expectation (as in the 2nd order). Notice that[12]

$$\begin{aligned} E[d^3 s_t] &= \theta_{k}^{(3)} + \theta_{kk}^{(3)} E[d^3 s_{t-1}] \\ &\Downarrow \\ E[d^3 s] &= \left(I - \theta_{kk}^{(3)}\right)^{-1} \theta_{k}^{(3)} \end{aligned}$$

Hence, the *fully-demeaned* state-space solution is

$$d^3 z_t = E[d^3 z] + \left[\begin{array}{cc} \theta_{yk}^{(3)} & \theta_{y\varsigma}^{(3)} \end{array} \right] \left[\begin{array}{c} d^3 s_t - E[d^3 s] \\ \varsigma_t^{(3)} \end{array} \right] \tag{45}$$

$$\left[\begin{array}{c} d^3 s_t - E[d^3 s] \\ \varsigma_t^{(3)} \end{array} \right] = \left[\begin{array}{cc} \theta_{kk}^{(3)} & \theta_{k\varsigma}^{(3)} \\ 0 & \phi_{\varsigma}^{(3)} \end{array} \right] \left[\begin{array}{c} d^3 s_{t-1} - E[d^3 s] \\ \varsigma_{t-1}^{(3)} \end{array} \right] + \left[\begin{array}{c} \theta_{kv}^{(3)} \\ \phi_{v}^{(3)} \end{array} \right] v_t^{(3)} \tag{46}$$

where $E[d^3 z] \equiv \theta_{y}^{(3)} + \theta_{yk}^{(3)} E[d^3 s]$ is the unconditional expectation of the third differential, $d^3 z_t$.

Notice that since $\Omega_{\xi s} = 0$, $\gamma^{(3)} = 0$. And since $\phi^{(3)} = 0$, we have $\theta_{k}^{(3)} = 0$ and $\theta_{y}^{(3)} = 0$ (see Appendix A). Hence, $E[d^3 s] = 0$ and $E[d^3 z] = 0$, though we have included them in (45) and (46) for completeness sake. This result does not mean that uncertainties (the stochastic shocks) do not have any effect on the third-order differential or Taylor-series approximation. In fact, the effect of uncertainties is indeed presence and is time-varying and state-dependent: it is embedded in vector $\varsigma_t^{(3)}$ through the presence of ds_t.

[12] $E[\varsigma_t^{(3)}] = \phi^{(3)} = 0$.

C The equivalence between differential Taylor series approximations and approximating the entire stochastic process

In this appendix, we show that finding an approximation to the entire stochastic process is equivalent to employing a differential version of the Taylor series approximation — this result thus justifies our approximation strategy in the main text.

For this purpose, we utilize a nonlinear, autoregressive stochastic difference equation,

$$s_t = m(s_{t-1}, \eta_t) \tag{47}$$

For simplicity and without any loss of generality, we assume s_t to be a scalar and the nonstochastic stationary point of this difference equation is 0, i.e., $0 = m(0,0)$. It is also assumed that $\{\eta_t\} = \{\sigma\varepsilon_t\}$ are a series of serially independent random variables with mean zero and constant variance $\omega = E\left[\eta_t^2\right] < \infty$. Finally, we assume that $m_s = \frac{\partial m(s,\eta)}{\partial s} < 1$ for all s, η so that the time series generated by the stochastic difference equation is a stationary stochastic process. This is the sort of stationarity assumption necessary in any autoregressive model. We focus below on a second-order approximation — but the result extends to any higher order.

1 Solution and approximation solutions

An exact solution to the stochastic difference equation is a function $S_t \equiv S(\{\eta_{t-j}\}_{t=0}^{\infty})$, which has the property that

$$S(\{\eta_{t-j}\}_{j=0}^{\infty}) = m(S(\{\eta_{t-1-j}\}_{j=0}^{\infty}), \eta_t)$$

Notice that the solution is implicit in this equation, as S appears on both sides.

An approximate solution is another function of $\{\eta_{t-j}\}_{j=0}^{\infty}$ that has specified properties. Using the perturbation parameter σ to increase the range of uncertainty in all random variables simultaneously, we take Taylor series approximations around the nonstochastic stationary point, $\sigma = 0 \Longrightarrow \eta_t = 0$.

1.1 Approximation via partial derivatives (approximating the entire stochastic process)

The direct approach via partial derivatives would be to calculate

$$s_t \approx s + \frac{\partial S_t}{\partial \sigma}\sigma + \frac{1}{2}\frac{\partial^2 S_t}{(\partial \sigma)^2}\sigma^2 \ .$$

We pursue this route initially, although it requires considerable work and use of the implicit function theorem to deduce the partial derivatives $\frac{\partial S_t}{\partial \sigma}$ and $\frac{\partial^2 S_t}{(\partial \sigma)^2}$.

First order approximation Consider the solution

$$s_t = S(\{\eta_{t-j}\}_{j=0}^{\infty}) \ .$$

A first-order Taylor series approximation of this function around the point $\sigma = 0$ is

$$s_t \simeq S(\{0\}) + \sum_{j=0}^{\infty} S_j \varepsilon_{t-j}\sigma = S(\{0\}) + \sum_{j=0}^{\infty} S_j \eta_{t-j} \ ,$$

where $S_j = \frac{\partial S}{\partial \eta_{t-j}}$ and these partial derivatives are evaluated at $\sigma = 0$ which implies that $\eta_{t-j} = 0$ for all j.

Application of the implicit function theorem supplies the necessary information on the level and derivatives above. In terms of the level of the function,

$$S(\{0\}) = m(S(\{0\}), 0) \Longrightarrow S(\{0\}) = 0 \ .$$

In terms of the derivatives of the function, $S(\{\eta_{t-j}\}_{t=0}^{\infty}) = m(S(\{\eta_{t-1-j}\}_{t=0}^{\infty}), \eta_t)$ implies that

$$\begin{aligned} S_0 &= m_\eta \\ S_j &= m_s S_{j-1} \ , \end{aligned}$$

where $m_\eta = \frac{\partial m(s,\eta)}{\partial \eta}$ and $m_s = \frac{\partial m(s,\eta)}{\partial s}$, with these partial derivatives being evaluated at $s = \eta = 0$. Accordingly, the first-order approximate solution takes the form

$$s_t \simeq l_t = \sum_{j=0}^{\infty} m_\eta (m_s)^j \eta_{t-j} = m_s l_{t-1} + m_\eta \eta_t \ ,$$

which is a linear stochastic difference equation.

Second order approximation We now follow the same approach to determine a second-order approximation to the stochastic difference equation, leading to

$$\begin{aligned} s_t &\simeq \sum_{j=0}^{\infty} S_j \varepsilon_{t-j} \sigma + \frac{1}{2} \sum_{j=0}^{\infty} \sum_{h=0}^{\infty} S_{jh} \varepsilon_{t-j} \varepsilon_{t-h} \sigma^2 \\ &= \sum_{j=0}^{\infty} S_j \eta_{t-j} + \frac{1}{2} \sum_{j=0}^{\infty} \sum_{h=0}^{\infty} S_{jh} \eta_{t-j} \eta_{t-h} \ , \end{aligned}$$

where $S_j = \frac{\partial S}{\partial \eta_{t-j}}$ and $S_{jh} = \frac{\partial^2 S}{(\partial \eta_{t-j})(\partial \eta_{t-h})}$. The implicit function theorem applied to $S(\{\eta_{t-j}\}_{t=0}^{\infty}) = m(S(\{\eta_{t-1-j}\}_{t=0}^{\infty}), \eta_t)$ indicates that

$$\begin{aligned} S_{00} &= m_{\eta\eta} \ , \\ S_{0j} &= m_{s\eta} S_{j-1} \text{ for } j \geqslant 1 \ , \\ S_{jh} &= m_{ss} S_{j-1} S_{h-1} + m_s S_{j-1,h-1} \text{ for } j \geqslant 1 \text{ and for } h \geqslant 1 \ . \end{aligned}$$

Working with the new quadratic term in this expression, we note that

$$\begin{aligned} q_t &= \sum_{j=0}^{\infty} \sum_{h=0}^{\infty} S_{jh} \eta_{t-j} \eta_{t-h} \\ &= S_{00} \eta_t^2 + 2\eta_t \sum_{j=1}^{\infty} S_{0j} \eta_{t-j} + \sum_{j=1}^{\infty} \sum_{h=1}^{\infty} S_{jh} \eta_{t-j} \eta_{t-h} \\ &= m_{\eta\eta} \eta_t^2 + 2 m_{s\eta} \eta_t l_{t-1} + \sum_{j=1}^{\infty} \sum_{h=1}^{\infty} S_{jh} \eta_{t-j} \eta_{t-h} \ . \end{aligned}$$

Further, changing the dating and using the implicit function theorem restrictions, we find that

$$\begin{aligned} q_t &= m_{\eta\eta} \eta_t^2 + 2 m_{s\eta} \eta_t l_{t-1} + \sum_{j=0}^{\infty} \sum_{h=0}^{\infty} S_{j+1,h+1} \eta_{t-1-j} \eta_{t-1-h} \\ &= m_{\eta\eta} \eta_t^2 + 2 m_{s\eta} \eta_t l_{t-1} + \sum_{j=0}^{\infty} \sum_{h=0}^{\infty} [m_{ss} S_j S_h + m_s S_{j,h}] (\eta_{t-1-j} \eta_{t-1-h}) \ . \end{aligned}$$

The expression above has a simple recursive form, so that the quadratic (second-order) term in the approximation is given by

$$q_t = m_{\eta\eta}\eta_t^2 + 2m_{s\eta}\eta_t l_{t-1} + m_{ss}(l_{t-1})^2 + m_s q_{t-1} \ ,$$

in which q_{t-1}, l_{t-1} and l_{t-1}^2 figure as relevant historical determinants of q_t.

1.2 Approximation via differentials

It would not be very convenient if we had to go through all of this detail every time that we wanted to take a second-order approximation. Fortunately, quite standard differentials provide a short-cut. We write the Taylor series approximation as

$$s_t \simeq ds_t + \frac{1}{2}(d^2 s_t)$$

and we calculate

$$\begin{aligned} ds_t &= m_s(ds_{t-1}) + m_\eta\,(d\eta_t) \\ d^2 s_t &= m_{ss}(ds_{t-1})^2 + 2m_{s\eta}(ds_{t-1})(d\eta_t) + m_{\eta\eta}(d\eta_t)^2 \\ &\quad + m_s(d^2 s_{t-1}) + \boxed{m_\eta(d^2\eta_t)} \ . \end{aligned}$$

With ds_t, $d\eta_t$, and $d^2 s_t$ now taking the role of l_t, η_t, and q_t, respectively, taking differentials thus produces exactly the same result as when approximating the entire stochastic process (approximation via partial derivatives). Note that the boxed term above is zero since we are stretching the distribution according to $\eta_t = \sigma\varepsilon_t$, around $\sigma = 0$ — that is, $d\eta_t = d(\sigma\varepsilon_t) = \varepsilon_t(\sigma - 0) = \eta_t$ and $d^j\eta_t = 0$ for all $j > 1$. Accordingly, we can use the differential approach to rapidly generate results from simultaneously stretching all shocks, so long as we are careful in our understanding about how the nature of the perturbation in σ affects the differentials of endogenous variables and shocks.

The state space representation We now cast the second-order approximate solution in state space form. First, we have defined the approximation to be the sum of two components, $s_t \simeq ds_t + \frac{1}{2}d^2 s_t$. Second, the state vector must contain the three elements $ds_{t-1}, (ds_{t-1})^2$ and $d^2 s_{t-1}$ since these are the relevant history. Further, we know that

$$(ds_t)^2 = (m_s ds_{t-1} + m_\eta\eta_t)^2 = m_s^2(ds_{t-1})^2 + m_\eta^2\eta_t^2 + 2m_s m_\eta\eta_t ds_{t-1}$$

as an identity. Hence, a state equation for the second order approximation $s_t \simeq ds_t + \frac{1}{2}d^2 s_t$ evolves according to

$$\begin{aligned} \begin{bmatrix} ds_t \\ d^2 s_t \\ (ds_t)^2 \end{bmatrix} &= \begin{bmatrix} m_s & 0 & 0 \\ 0 & m_s & m_{ss} \\ 0 & 0 & m_s^2 \end{bmatrix} \begin{bmatrix} ds_{t-1} \\ d^2 s_{t-1} \\ (ds_{t-1})^2 \end{bmatrix} \\ &\quad + \begin{bmatrix} 0 \\ m_{\eta\eta}E\left[\eta_t^2\right] \\ m_\eta^2 E\left[\eta_t^2\right] \end{bmatrix} + \begin{bmatrix} m_\eta\eta_t \\ m_{\eta\eta}(\eta_t^2 - E\left[\eta_t^2\right]) + 2m_{s\eta}\eta_t ds_{t-1} \\ m_\eta^2(\eta_t^2 - E\left[\eta_t^2\right]) + 2m_s m_\eta\eta_t ds_{t-1} \end{bmatrix} \ . \end{aligned}$$

Note that the forecast errors now contain the products of the shock η_t and the state of the system, ds_{t-1}, leading to both state-dependence and conditional heteroskedasticity.

There is a nonzero mean to $d^2 s_t$ and $(ds_t)^2$, which can be calculated as

$$E\left[\begin{array}{c} d^2 s_t \\ (ds_t)^2 \end{array}\right] = \left\{ I - \left[\begin{array}{cc} m_s & m_{ss} \\ 0 & m_s^2 \end{array}\right] \right\}^{-1} \left[\begin{array}{c} m_{\eta\eta} E\left[\eta_t^2\right] \\ m_\eta^2 E\left[\eta_t^2\right] \end{array}\right] ,$$

so that the second-order approximation no longer displays certainty equivalence in the mean.