NBER WORKING PAPER SERIES

THE PRUNED STATE-SPACE SYSTEM FOR NON-LINEAR DSGE MODELS: THEORY AND EMPIRICAL APPLICATIONS

Martin M. Andreasen
Jesús Fernández-Villaverde
Juan Rubio-Ramírez

Working Paper 18983
http://www.nber.org/papers/w18983

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
April 2013

We acknowledge access to computer facilities provided by the Danish Center for Scientific Computing (DCSC). We appreciate the financial support of the Center for Research in Econometric Analysis of Time Series (CREATES), funded by the Danish National Research Foundation and the National Science Foundation (NSF). The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

The Pruned State-Space System for Non-Linear DSGE Models: Theory and Empirical Applications
Martin M. Andreasen, Jesús Fernández-Villaverde, and Juan Rubio-Ramírez
NBER Working Paper No. 18983
April 2013
JEL No. C15,C53,E30

ABSTRACT

This paper studies the pruned state-space system for higher-order approximations to the solutions of DSGE models. For second- and third-order approximations, we derive the statistical properties of this system and provide closed-form expressions for first and second unconditional moments and impulse response functions. Thus, our analysis introduces GMM estimation for DSGE models approximated up to third-order and provides the foundation for indirect inference and SMM when simulation is required. We illustrate the usefulness of our approach by estimating a New Keynesian model with habits and Epstein-Zin preferences by GMM when using first and second unconditional moments of macroeconomic and financial data and by SMM when using additional third and fourth unconditional moments and non-Gaussian innovations.

Martin M. Andreasen
Aarhus University
Fuglesangs Allé 4
8210 Aarhus V
Denmark
mandreasen@econ.au.dk

Jesús Fernández-Villaverde
University of Pennsylvania
160 McNeil Building
3718 Locust Walk
Philadelphia, PA 19104
and NBER
jesusfv@econ.upenn.edu

Juan Rubio-Ramírez
Duke University
P.O. Box 90097
Durham, NC 27708
juan.rubio-ramirez@duke.edu

An online appendix is available at:
http://www.nber.org/data-appendix/w18983

1 Introduction

Perturbation methods allow researchers to build higher-order Taylor approximations to the solution of dynamic stochastic general equilibrium (DSGE) models (see Judd and Guu (1997), Schmitt-Grohé and Uribe (2004), Aruoba, Fernandez-Villaverde and Rubio-Ramirez (2006), among others). Higher-order Taylor approximations are growing in popularity for two reasons. First, researchers want to find approximated solutions that are more accurate than those obtained from a linear approximation. Second, there is much interest in questions that are inherently non-linear, such as the consequences of uncertainty shocks or macroeconomic determinants behind risk premia that cannot be studied with linear methods (see Fernández-Villaverde, Guerrón-Quintana, Rubio-Ramírez and Uribe (2011) or Binsbergen, Fernandez-Villaverde, Koijen and Rubio-Ramirez (2012)).

Although higher-order approximations are intuitive and straightforward to compute, the simulated sample paths that they generate often explode. This is true even when the corresponding linear approximation is stable. The presence of explosive behavior complicates any model evaluation because no unconditional moments would exist based on this approximation. It also means that any unconditional moment-matching estimation methods, such as the generalized method of moments (GMM) or the simulated method of moments (SMM), are inapplicable in this context as they rely on finite moments from stationary and ergodic probability distributions. For a review of GMM in the context of DSGE models, see Ruge-Murcia (2013).[1]

For second-order approximations, Kim, Kim, Schaumburg and Sims (2008) have proposed to address the problem of explosive sample paths by applying the pruning method. Pruning means leaving out terms in the solution that have higher-order effects than the approximation order. The intuition is as follows. Suppose we have a solution for capital k_t that depends on a quadratic function of k_{t-1} (as we will typically have with a second-order approximation). If we substitute k_{t-1} for its own quadratic function of k_{t-2}, then we obtain an expression for k_t that depends on k_{t-2}, k_{t-2}^2, k_{t-2}^3, and k_{t-2}^4. Pruning omits terms related to k_{t-2}^3 and k_{t-2}^4 because they are third- and fourth-order effects, respectively. Kim, Kim, Schaumburg and Sims (2008) show that leaving out these effects ensures that the pruned approximation does not explode.

The purpose of this paper is to extend pruning beyond second-order approximations and to

[1] Non-explosive paths are also necessary to implement full likelihood methods, for instance, with a non-linear filter such as the particle filter presented in Fernández-Villaverde and Rubio-Ramírez (2007).

explore its econometric implications. We proceed in two steps. First, we show how to apply pruning to an approximation of any arbitrary order. We do so by exploiting what we refer to as the pruned state-space system. We pay special attention to second and third-order approximations, those most likely to be used by researchers. Second, we derive the statistical properties of the pruned state-space system. Under very general technical conditions, we show that first and second unconditional moments exist. Then, we provide closed-form expressions for first and second unconditional moments and impulse response functions (IRFs). Conditions for the existence of some higher unconditional moments, such as skewness and kurtosis, are derived as well.

The econometric implications of these closed-form expressions are significant as most of the existing unconditional moment-based estimation methods for linearized DSGE models now carry over to non-linear approximations in a simple way. For approximations up to third order, this includes GMM estimation based on first and second unconditional moments and matching of model-implied IRFs to their empirical counterpart, as in Christiano, Eichenbaum and Evans (2005) for a linearized model. Both estimation methods can be implemented without the need for simulation and thereby avoiding numerical simulation errors. Following Kim (2002), the unconditional moment conditions in optimal GMM estimation can be used to build a so-called limited information likelihood function, from which Bayesian inference may be carried out. Hence, our closed-form expressions are also useful to researchers interested in the Bayesian tradition. Finally, when simulations are needed to calculate higher unconditional moments, our analysis provides the foundation for different types of indirect inference as in Smith (1993) and SMM as in Duffie and Singleton (1993).

The suggested GMM estimation approach, its Bayesian equivalent, and IRF matching seem particular promising because we can compute first and second unconditional moments or any IRFs in a trivial amount of time for medium-sized DSGE models approximated up to third-order. For the model described in section 7 with seven state variables, it takes less than one second to find all first and second unconditional moments and slightly more than one second to compute all the IRFs using `Matlab`. Moreover, our moment approach would enable us to determine the stochastic specification of the innovations in a semi-parametric manner, although due to space limitations, we do not pursue that route explicitly in this paper. To make our results easily assessable, our `Matlab` codes are publicly available on the authors' home pages.[2]

[2] Ongoing work deals with implementing our findings in Dynare.

Our results are also relevant to researchers who do not rely on formal estimation but, instead, prefer to calibrate their models. In the traditional calibration strategy with linear approximations, structural parameters are determined by matching the steady-state values of the model to mean values in the data (see Cooley and Prescott (1995) for a canonical exposition). However, when higher-order approximations are considered, the first unconditional moments of variables in the model will, in general, differ from their steady-state values as non-linear terms correct for uncertainty. Thus, following the traditional calibration strategy when higher-order approximations are considered might result in misleading structural parameter values. Instead, the closed-form solutions for first unconditional moments provided in this paper can be used to select the correct values for the structural parameters for models solved with higher-order approximations.[3]

In an empirical application, we illustrate some of the econometric techniques that our paper makes available for non-linear DSGE models. Our focus is on a rich New Keynesian model with habits and Epstein-Zin preferences, which we estimate on key moments for the U.S. yield curve and five macroeconomic variables. Using first and second unconditional moments, we estimate the model by GMM when using second- and third-order approximations. When additional third and fourth unconditional moments are included, we resort to simulation and estimate the model by SMM. Thanks to our results, the estimation is efficiently done and the results are transparent and easy to interpret.

To be concrete, in the case of the GMM estimation using first and second unconditional moments, we find sizable habits, a low Frisch elasticity of labor supply, a high relative risk aversion, a low intertemporal elasticity of substitution for consumption, and a high degree of price stickiness. With respect to the fit of the model, our economy matches all mean values extremely well, in particular the short- and long-term interest rates of 5.6 percent and 6.9 percent, respectively, and the mean inflation rate of 3.7 percent. The fit of the variability, correlations, and autocorrelations in the data is still good, although less impressive than the fit of the first moments. Consequently, a standard specification J-test cannot reject the model. Interestingly, when we use the third-order approximation, we find a nominal term premium with a sizable mean of 174 annualized basis points

[3] Some papers in the literature have accounted for the difference between the deterministic steady state and the mean of the ergodic distribution by means of simulation; see, for instance, Fernández-Villaverde, Guerrón-Quintana, Rubio-Ramírez and Uribe (2011). The formulas in this paper provide a much more efficient way to carry out the calibration approach in non-linear approximations.

and a quite realistic standard deviation of 20 annualized basis points. For the SMM estimation with third and fourth unconditional moments, most parameters are unchanged and the specification test still cannot reject the model. Moreover, we find evidence of non-Gaussianity in technology shocks but not in preference shocks.

Based on our analysis, we view the pruning method as a most useful tool. It makes available a large set of well-known and intuitive estimators when approximations up to third-order are considered. This is because any sample path is covariance stationary, implying that explosive sample paths do not occur (almost surely).

The rest of the paper is structured as follows. Section 2 introduces the problem. Section 3 presents the pruning method and derives what we call the pruned state-space system for DSGE models whose solution has been approximated to any order. The statistical properties of the pruned state-space system associated with second- and third-order approximations, including closed-form expressions for first and second unconditional moments, are analyzed in section 4. Section 5 derives closed-form expressions for IRFs. Section 6 presents a standard New Keynesian model with habits and Epstein-Zin preferences that we use to examine the numerical accuracy of pruning. The econometric implications of the pruned state-space system are discussed in section 7. There we also estimate a New Keynesian DSGE model on post-war US data. Concluding comments are provided in section 8. Detailed derivations and proofs are deferred to the appendix at the end of the paper. In addition, a longer technical appendix is available on the authors' home pages or on request.

2 The State-Space System

We consider the following class of DSGE models. Let $\mathbf{x}_t \in \mathbb{R}^{n_x}$ be a vector of predetermined state variables, $\mathbf{y}_t \in \mathbb{R}^{n_y}$ a vector of non-predetermined control variables, and $\sigma \geq 0$ an auxiliary perturbation parameter. The exact solution to the DSGE model is given by a set of decision rules for the control variables

$$\mathbf{y}_t = \mathbf{g}\left(\mathbf{x}_t, \sigma\right), \tag{1}$$

and for the state variables

$$\mathbf{x}_{t+1} = \mathbf{h}\left(\mathbf{x}_t, \sigma\right) + \sigma \boldsymbol{\eta} \boldsymbol{\epsilon}_{t+1}, \tag{2}$$

where $\boldsymbol{\epsilon}_{t+1}$ contains the n_ϵ exogenous zero-mean innovations. Initially, we only assume that $\boldsymbol{\epsilon}_{t+1}$ is independent and identically distributed with finite second moments, meaning that no distributional assumption is imposed and the innovations may therefore be non-Gaussian. This is denoted by $\boldsymbol{\epsilon}_{t+1} \sim IID(\mathbf{0}, \mathbf{I})$. Additional moment restrictions will be imposed in the following sections.[4] The perturbation parameter σ scales the square root of the covariance matrix for the innovations $\boldsymbol{\eta}$, which has dimension $n_x \times n_\epsilon$.[5] Equation (1) is typically called the observation equation and equation (2) the state equation. We will refer to (1)-(2) as the exact state-space system.

DSGE models do not, in general, have closed-form solutions and the functions $\mathbf{g}(\mathbf{x}_t, \sigma)$ and $\mathbf{h}(\mathbf{x}_t, \sigma)$ cannot be found explicitly. The perturbation method is a popular way to obtain Taylor-series expansions to these functions around the deterministic steady state, that is, at $\mathbf{x}_t = \mathbf{x}_{t+1} = \mathbf{x}_{ss}$ and $\sigma = 0$ (to simplify the notation, all variables are expressed in deviation from the steady state, i.e., $\mathbf{x}_{ss} = 0$.) Given these Taylor-series expansions, we can set up an approximated state-space system and simulate the economy, compute conditional or unconditional moments, and evaluate the likelihood function given some observed data.

When the functions $\mathbf{g}(\mathbf{x}_t, \sigma)$ and $\mathbf{h}(\mathbf{x}_t, \sigma)$ are approximated up to first-order, the approximated state-space system is simply obtained by replacing $\mathbf{g}(\mathbf{x}_t, \sigma)$ and $\mathbf{h}(\mathbf{x}_t, \sigma)$ with $\mathbf{g}_\mathbf{x}\mathbf{x}_t$ and $\mathbf{h}_\mathbf{x}\mathbf{x}_t$ in (1) and (2), respectively. Here, $\mathbf{g}_\mathbf{x}$ is an $n_y \times n_x$ matrix with first-order derivatives of $\mathbf{g}(\mathbf{x}_t, \sigma)$ with respect to $\mathbf{x}_t$ and $\mathbf{h}_\mathbf{x}$ is an $n_x \times n_x$ matrix with first-order derivatives of $\mathbf{h}(\mathbf{x}_t, \sigma)$ with respect to $\mathbf{x}_t$ (by certainty equivalence, the first-order derivatives $\mathbf{g}_\sigma$ and $\mathbf{h}_\sigma$ of $\mathbf{g}(\mathbf{x}_t, \sigma)$ and $\mathbf{h}(\mathbf{x}_t, \sigma)$ with respect to σ are zero.) Given our assumptions about $\boldsymbol{\epsilon}_{t+1}$, this system has finite first and second moments if all eigenvalues of $\mathbf{h}_\mathbf{x}$ have modulus less than one. Furthermore, the approximated state-space system fluctuates around the deterministic steady state, which also corresponds to its mean value. It is therefore straightforward to calibrate the parameters of the model based on first and second moments or to estimate them using standard econometric tools such as Bayesian methods, maximum likelihood estimation (MLE), GMM, SMM, etc. (see Ruge-Murcia (2007)).

When the decision rules are approximated beyond first-order, we can, in principle, apply the

[4] Throughout the rest of the paper, unless stated otherwise, the moment requirements refer to the unconditional moments.

[5] The assumption that innovations enter linearly in (2) may appear restrictive, but it is without loss of generality because the state vector can be extended to deal with non-linearities between $\mathbf{x}_t$ and $\boldsymbol{\epsilon}_{t+1}$. Appendix B provides the details and an illustration based on the neoclassical growth model with heteroscedastic innovations in the process for technology shocks.

same method to construct the approximated state-space system. That is, we can replace $\mathbf{g}(\mathbf{x}_t, \sigma)$ and $\mathbf{h}(\mathbf{x}_t, \sigma)$ in (1) and (2) with their higher-order Taylor-series expansions. However, the resulting approximated state-space system cannot, in general, be shown to have any finite moments and it often displays explosive dynamics. This occurs, for instance, when we simulate simple versions of the New Keynesian model with just a few endogenous state variables. Hence, it is difficult to use the approximated state-space system to calibrate or to estimate the parameters of the model.

Thus, it is useful to construct an alternative approximated state-space system with well-defined statistical properties to analyze DSGE models. Section 3 explains how this can be done with the pruning method. Sections 4 and 5 derive the statistical properties of the resulting state-space system and compute closed-form expressions for first and second moments and IRFs.

3 The Pruning Method

In the context of DSGE models whose solution has been approximated up to second-order, Kim, Kim, Schaumburg and Sims (2008) suggest using a pruning method when constructing the alternative approximated state-space system. They show that this system is stable because it preserves only first- and second-order effects when the system is iterated forward in time. Any other effects are omitted because they do not in general correspond to higher-order terms in a more accurate Taylor-series expansion. Applying the pruning method is thus different from simply replacing $\mathbf{g}(\mathbf{x}_t, \sigma)$ and $\mathbf{h}(\mathbf{x}_t, \sigma)$ with their Taylor-series expansions in (1) and (2) because the latter procedure includes all such higher-order effects.

In section 3.1, we review the pruning method and explain its logic for second-order approximations. Section 3.2 extends this method to third-order approximations. The general procedure for constructing the pruned state-space system associated with any order of approximation is then outlined in section 3.3. Finally, section 3.4 relates our approach to the existing literature.

3.1 Second-Order Approximation

The first step when constructing the pruned state-space system associated with the second-order approximation is to decompose the state variables into first-order effects $\mathbf{x}_t^f$ and second-order effects

$\mathbf{x}_t^s$. This can be done starting from the second-order Taylor-series expansion to the state equation

$$\mathbf{x}_{t+1}^{(2)} = \mathbf{h_x x}_t^{(2)} + \frac{1}{2}\mathbf{H_{xx}}\left(\mathbf{x}_t^{(2)} \otimes \mathbf{x}_t^{(2)}\right) + \frac{1}{2}\mathbf{h}_{\sigma\sigma}\sigma^2 + \sigma\boldsymbol{\eta}\boldsymbol{\epsilon}_{t+1}, \tag{3}$$

where we use the notation $\mathbf{x}_t^{(2)}$ to represent the unpruned second-order approximation to the state variable.[6] Here, $\mathbf{H_{xx}}$ is an $n_x \times n_x^2$ matrix with the derivatives of $\mathbf{h}(\mathbf{x}_t, \sigma)$ with respect to $(\mathbf{x}_t, \mathbf{x}_t)$ and $\mathbf{h}_{\sigma\sigma}$ is an $n_x \times 1$ matrix containing derivatives taken with respect to (σ, σ).[7] Substituting $\mathbf{x}_t^{(2)}$ by $\mathbf{x}_t^f + \mathbf{x}_t^s$ into the right-hand side of (3) gives

$$\mathbf{h_x}\left(\mathbf{x}_t^f + \mathbf{x}_t^s\right) + \frac{1}{2}\mathbf{H_{xx}}\left(\left(\mathbf{x}_t^f + \mathbf{x}_t^s\right) \otimes \left(\mathbf{x}_t^f + \mathbf{x}_t^s\right)\right) + \frac{1}{2}\mathbf{h}_{\sigma\sigma}\sigma^2 + \sigma\boldsymbol{\eta}\boldsymbol{\epsilon}_{t+1}. \tag{4}$$

A law of motion for $\mathbf{x}_{t+1}^f$ is derived by preserving only first-order effects in (4). We, therefore, keep the first-order effects from the previous period $\mathbf{h_x x}_t^f$ and the innovations $\sigma\boldsymbol{\eta}\boldsymbol{\epsilon}_{t+1}$ to obtain

$$\mathbf{x}_{t+1}^f = \mathbf{h_x x}_t^f + \sigma\boldsymbol{\eta}\boldsymbol{\epsilon}_{t+1}. \tag{5}$$

Hence, the expression for $\mathbf{x}_{t+1}^f$ is the standard first-order approximation to the state equation. Note also that $\mathbf{x}_{t+1}^f$ is a first-order polynomial of $\{\boldsymbol{\epsilon}_s\}_{s=1}^{t+1}$. The first-order approximation to the observation equation is also standard and given by

$$\mathbf{y}_t^f = \mathbf{g_x x}_t^f. \tag{6}$$

Accordingly, the pruned state-space system for a first-order approximation to the solution of the model is given by (5) and (6), meaning that the pruned and unpruned state-space systems are identical in this case.

A law of motion for $\mathbf{x}_{t+1}^s$ is derived by preserving only second-order effects in (4). Here, we include the second-order effects from the previous period $\mathbf{h_x x}_t^s$, the squared first-order effects in

[6] Note that (3) adopts the standard assumption that the model has a unique stable first-order approximation, which implies that all second- and higher-order terms are also unique (see Judd and Guu (1997)).

[7] The matrix $\mathbf{H_{xx}}$ can be obtained in `Matlab` by using the *reshape* function, $\mathbf{H_{xx}} = reshape(\mathbf{h_{xx}}, n_x, n_x^2)$, where $\mathbf{h_{xx}}$ is an array of dimension $n_x \times n_x \times n_x$. For instance, the i-th row of $\mathbf{H_{xx}}$ contains all the second-order terms for the i-th state variable.

the previous period $\frac{1}{2}\mathbf{H}_{\mathbf{xx}}\left(\mathbf{x}_t^f \otimes \mathbf{x}_t^f\right)$, and the correction $\frac{1}{2}\mathbf{h}_{\sigma\sigma}\sigma^2$. Hence,

$$\mathbf{x}_{t+1}^s = \mathbf{h}_{\mathbf{x}}\mathbf{x}_t^s + \frac{1}{2}\mathbf{H}_{\mathbf{xx}}\left(\mathbf{x}_t^f \otimes \mathbf{x}_t^f\right) + \frac{1}{2}\mathbf{h}_{\sigma\sigma}\sigma^2. \tag{7}$$

We do not include terms with $\mathbf{x}_t^f \otimes \mathbf{x}_t^s$ and $\mathbf{x}_t^s \otimes \mathbf{x}_t^s$ because they reflect third- and fourth-order effects, respectively. It is also worth mentioning that we are treating σ as a variable and not as a constant when deriving (7), meaning that $\frac{1}{2}\mathbf{h}_{\sigma\sigma}\sigma^2$ is a second-order effect. Note that $\mathbf{x}_{t+1}^s$ is a second-order polynomial in $\{\boldsymbol{\epsilon}_s\}_{s=1}^t$.

The final step in setting up the pruned state-space system is to derive the expression for the observation equation. Using the same approach, we start from the second-order Taylor-series expansion to the observation equation

$$\mathbf{y}_t^{(2)} = \mathbf{g}_{\mathbf{x}}\mathbf{x}_t^{(2)} + \frac{1}{2}\mathbf{G}_{\mathbf{xx}}\left(\mathbf{x}_t^{(2)} \otimes \mathbf{x}_t^{(2)}\right) + \frac{1}{2}\mathbf{g}_{\sigma\sigma}\sigma^2, \tag{8}$$

where $\mathbf{y}_t^{(2)}$ denotes the unpruned second-order approximation to the control variable. Here, $\mathbf{G}_{\mathbf{xx}}$ is an $n_y \times n_x^2$ matrix with the corresponding derivatives of $\mathbf{g}(\mathbf{x}_t, \sigma)$ with respect to $(\mathbf{x}_t, \mathbf{x}_t)$ and $\mathbf{g}_{\sigma\sigma}$ is an $n_y \times 1$ matrix containing derivatives with respect to (σ, σ). We only want to preserve effects up to second-order, meaning that

$$\mathbf{y}_t^s = \mathbf{g}_{\mathbf{x}}\left(\mathbf{x}_t^f + \mathbf{x}_t^s\right) + \frac{1}{2}\mathbf{G}_{\mathbf{xx}}\left(\mathbf{x}_t^f \otimes \mathbf{x}_t^f\right) + \frac{1}{2}\mathbf{g}_{\sigma\sigma}\sigma^2. \tag{9}$$

Here, we leave out terms with $\mathbf{x}_t^f \otimes \mathbf{x}_t^s$ and $\mathbf{x}_t^s \otimes \mathbf{x}_t^s$ because they reflect third- and fourth-order effects, respectively.[8]

Accordingly, the pruned state-space system for a second-order approximation is given by (5), (7), and (9). The state vector in this system is thus extended to $\begin{bmatrix} \left(\mathbf{x}_t^f\right)' & \left(\mathbf{x}_t^s\right)' \end{bmatrix}'$ as we separately need to track first- and second-order effects. Another important observation is that all variables in this system are second-order polynomials of the innovations. The pruning method applied to a second-order approximation may, therefore, be interpreted as imposing the additional requirement that all variables in the pruned state-space system are second-order polynomials in the innovations.

[8] We are treating $\mathbf{y}_t^s$ as the sum of the first- and second-order effects, while $\mathbf{x}_t^s$ is only the second-order effects. This allows us to simplify the notation.

The unpruned state-space system for a second-order approximation is given by (3) and (8).

3.2 Third-Order Approximation

We now construct the pruned state-space system for third-order approximations. Following the steps outlined above, we start by decomposing the state variables into first-order effects $\mathbf{x}_t^f$, second-order effects $\mathbf{x}_t^s$, and third-order effects $\mathbf{x}_t^{rd}$. The laws of motions for $\mathbf{x}_t^f$ and $\mathbf{x}_t^s$ are the same as in the previous section, and only the recursion for $\mathbf{x}_t^{rd}$ remains to be derived. The third-order Taylor-series expansion to the state equation is (see Andreasen (2012))

$$\begin{aligned}\mathbf{x}_{t+1}^{(3)} &= \mathbf{h_x}\mathbf{x}_t^{(3)} + \frac{1}{2}\mathbf{H_{xx}}\left(\mathbf{x}_t^{(3)} \otimes \mathbf{x}_t^{(3)}\right) + \frac{1}{6}\mathbf{H_{xxx}}\left(\mathbf{x}_t^{(3)} \otimes \mathbf{x}_t^{(3)} \otimes \mathbf{x}_t^{(3)}\right) \\ &\quad + \frac{1}{2}\mathbf{h}_{\sigma\sigma}\sigma^2 + \frac{3}{6}\mathbf{h}_{\sigma\sigma\mathbf{x}}\sigma^2\mathbf{x}_t^{(3)} + \frac{1}{6}\mathbf{h}_{\sigma\sigma\sigma}\sigma^3 + \sigma\boldsymbol{\eta}\boldsymbol{\epsilon}_{t+1}. \end{aligned} \tag{10}$$

where $\mathbf{x}_t^{(3)}$ represents the unpruned third-order approximation to the state variable. Now, $\mathbf{H_{xxx}}$ denotes an $n_x \times n_x^3$ matrix containing derivatives of $\mathbf{h}(\mathbf{x}_t, \sigma)$ with respect to $(\mathbf{x}_t, \mathbf{x}_t, \mathbf{x}_t)$, $\mathbf{h}_{\sigma\sigma\mathbf{x}}$ is an $n_x \times n_x$ matrix with derivatives with respect to $(\sigma, \sigma, \mathbf{x}_t)$, and $\mathbf{h}_{\sigma\sigma\sigma}$ is an $n_x \times 1$ matrix containing derivatives related to (σ, σ, σ). We adopt the same procedure as in the previous subsection and substitute $\mathbf{x}_t^f + \mathbf{x}_t^s + \mathbf{x}_t^{rd}$ into the right-hand side of (10) to obtain

$$\begin{aligned}&\mathbf{h_x}\left(\mathbf{x}_t^f + \mathbf{x}_t^s + \mathbf{x}_t^{rd}\right) + \frac{1}{2}\mathbf{H_{xx}}\left(\left(\mathbf{x}_t^f + \mathbf{x}_t^s + \mathbf{x}_t^{rd}\right) \otimes \left(\mathbf{x}_t^f + \mathbf{x}_t^s + \mathbf{x}_t^{rd}\right)\right) \\ &+ \frac{1}{6}\mathbf{H_{xxx}}\left(\left(\mathbf{x}_t^f + \mathbf{x}_t^s + \mathbf{x}_t^{rd}\right) \otimes \left(\mathbf{x}_t^f + \mathbf{x}_t^s + \mathbf{x}_t^{rd}\right) \otimes \left(\mathbf{x}_t^f + \mathbf{x}_t^s + \mathbf{x}_t^{rd}\right)\right) \\ &+ \frac{1}{2}\mathbf{h}_{\sigma\sigma}\sigma^2 + \frac{3}{6}\mathbf{h}_{\sigma\sigma\mathbf{x}}\sigma^2\left(\mathbf{x}_t^f + \mathbf{x}_t^s + \mathbf{x}_t^{rd}\right) + \frac{1}{6}\mathbf{h}_{\sigma\sigma\sigma}\sigma^3 + \sigma\boldsymbol{\eta}\boldsymbol{\epsilon}_{t+1}. \end{aligned} \tag{11}$$

A law of motion for the third-order effects is then derived by preserving only third-order effects in (11). Using the same line of reasoning as above, we get

$$\mathbf{x}_{t+1}^{rd} = \mathbf{h_x}\mathbf{x}_t^{rd} + \mathbf{H_{xx}}\left(\mathbf{x}_t^f \otimes \mathbf{x}_t^s\right) + \frac{1}{6}\mathbf{H_{xxx}}\left(\mathbf{x}_t^f \otimes \mathbf{x}_t^f \otimes \mathbf{x}^f\right) + \frac{3}{6}\mathbf{h}_{\sigma\sigma\mathbf{x}}\sigma^2\mathbf{x}_t^f + \frac{1}{6}\mathbf{h}_{\sigma\sigma\sigma}\sigma^3. \tag{12}$$

As in the derivation of the law of motion for $\mathbf{x}_t^s$ in (7), σ is interpreted as a variable when constructing (12). This means that $\frac{3}{6}\mathbf{h}_{\sigma\sigma\mathbf{x}}\sigma^2\mathbf{x}_t^s$ and $\frac{3}{6}\mathbf{h}_{\sigma\sigma\mathbf{x}}\sigma^2\mathbf{x}_t^{rd}$ represent fourth- and fifth-order effects, respectively, and are therefore omitted. Haan and Wind (2010) adopt the opposite approach and consider σ as a constant when deriving the pruned state-space system, meaning that they also include $\frac{3}{6}\mathbf{h}_{\sigma\sigma\mathbf{x}}\sigma^2\mathbf{x}_t^{rd}$ in (12). We prefer to treat σ as a variable because it is consistent with the assumed functional forms for $\mathbf{g}(\mathbf{x}_t,\sigma)$ and $\mathbf{h}(\mathbf{x}_t,\sigma)$ and the way the perturbation approximation is derived. As section 4 shows, interpreting σ as a variable is essential for proving that the pruned state-space system is, in general, stable and has well-defined statistical properties. Finally, note that $\mathbf{x}_{t+1}^{rd}$ is a third-order polynomial in $\{\boldsymbol{\epsilon}_s\}_{s=1}^{t}$.

The final step is to set up the expression for the observation equation. Using results in Andreasen (2012), the third-order Taylor-series expansion to the observation may be written as

$$\begin{aligned}\mathbf{y}_t^{(3)} \quad = \quad & \mathbf{g}_{\mathbf{x}}\mathbf{x}_t^{(3)} + \frac{1}{2}\mathbf{G}_{\mathbf{xx}}\left(\mathbf{x}_t^{(3)}\otimes\mathbf{x}_t^{(3)}\right) + \frac{1}{6}\mathbf{G}_{\mathbf{xxx}}\left(\mathbf{x}_t^{(3)}\otimes\mathbf{x}_t^{(3)}\otimes\mathbf{x}_t^{(3)}\right) \\ & + \frac{1}{2}\mathbf{g}_{\sigma\sigma}\sigma^2 + \frac{3}{6}\mathbf{g}_{\sigma\sigma\mathbf{x}}\sigma^2\mathbf{x}_t^{(3)} + \frac{1}{6}\mathbf{g}_{\sigma\sigma\sigma}\sigma^3, \end{aligned} \tag{13}$$

where $\mathbf{y}_t^{(3)}$ represents the unpruned third-order approximation to the control variable. In (13), $\mathbf{G}_{\mathbf{xxx}}$ denotes an $n_y \times n_x^3$ matrix containing derivatives of $\mathbf{g}(\mathbf{x}_t,\sigma)$ with respect to $(\mathbf{x}_t,\mathbf{x}_t,\mathbf{x}_t)$, $\mathbf{g}_{\sigma\sigma\mathbf{x}}$ is an $n_y \times n_x$ matrix with derivatives with respect to $(\sigma,\sigma,\mathbf{x}_t)$, and $\mathbf{g}_{\sigma\sigma\sigma}$ is an $n_x \times 1$ matrix containing derivatives related to (σ,σ,σ). Hence, preserving effects up to third-order gives[9]

$$\begin{aligned}\mathbf{y}_t^{rd} \quad = \quad & \mathbf{g}_{\mathbf{x}}\left(\mathbf{x}_t^{f}+\mathbf{x}_t^{s}+\mathbf{x}_t^{rd}\right) + \frac{1}{2}\mathbf{G}_{\mathbf{xx}}\left(\left(\mathbf{x}_t^{f}\otimes\mathbf{x}_t^{f}\right)+2\left(\mathbf{x}_t^{f}\otimes\mathbf{x}_t^{s}\right)\right) \\ & + \frac{1}{6}\mathbf{G}_{\mathbf{xxx}}\left(\mathbf{x}_t^{f}\otimes\mathbf{x}_t^{f}\otimes\mathbf{x}_t^{f}\right) + \frac{1}{2}\mathbf{g}_{\sigma\sigma}\sigma^2 + \frac{3}{6}\mathbf{g}_{\sigma\sigma\mathbf{x}}\sigma^2\mathbf{x}_t^{f} + \frac{1}{6}\mathbf{g}_{\sigma\sigma\sigma}\sigma^3. \end{aligned} \tag{14}$$

Thus, the pruned state-space system associated with the third-order approximation is given by (5), (7), (12), and (14). The state vector in this system is further extended to $\left[\begin{array}{ccc}\left(\mathbf{x}_t^{f}\right)' & \left(\mathbf{x}_t^{s}\right)' & \left(\mathbf{x}_t^{rd}\right)'\end{array}\right]'$, as we need to separately track first-, second-, and third-order effects. Also, all variables in this system are third-order polynomials of the innovations. Hence, the pruning method applied to a third-order approximation may be interpreted as imposing the requirement that all variables in the pruned state-space system are third-order polynomials in the innovations. The unpruned state-space system for a third-order approximation is given by (10) and (13).

[9] To again simplify notation, we treat $\mathbf{y}_t^{rd}$ as the sum of the first-, second-, and third-order effects, while $\mathbf{x}_t^{rd}$ is only the third-order effects.

3.3 Higher-Order Approximations

It is straightforward to apply the pruning method and obtain the pruned state-space system for the kth-order approximation to the model's solution. Based on the kth-order Taylor-series expansions of $\mathbf{g}(\mathbf{x}_t, \sigma)$ and $\mathbf{h}(\mathbf{x}_t, \sigma)$, the three steps are:

1. Decompose the state variables into first-, second-, ... , and kth-order effects.
2. Set up laws of motions for the state variables capturing only first-, second-, ... , and kth-order effects.
3. Construct the expression for control variables by preserving only effects up to kth-order.

In comparison, the unpruned state-space system for a kth-order approximation is given by the kth-order Taylor-series expansions of $\mathbf{g}(\mathbf{x}_t, \sigma)$ and $\mathbf{h}(\mathbf{x}_t, \sigma)$.

3.4 Related Literature

Besides Kim, Kim, Schaumburg and Sims (2008), there are a few papers that analyze the pruning method. Lombardo and Sutherland (2007) use a reasoning similar to ours to derive a second-order accurate expression for all squared terms in a DSGE model. With this result and standard solution routines for linearized models, they compute the second-order derivatives of the functions $\mathbf{g}(\mathbf{x}_t, \sigma)$ and $\mathbf{h}(\mathbf{x}_t, \sigma)$. Hence, at least up to second-order, Lombardo and Sutherland (2007) show that pruning is fully consistent with the perturbation method.

Lan and Meyer-Gohde (2011) employ perturbation to derive nonlinear approximations of $\mathbf{g}(\mathbf{x}_t, \sigma)$ and $\mathbf{h}(\mathbf{x}_t, \sigma)$, where these functions are expressed in terms of past innovations $\left\{\boldsymbol{\epsilon}_s, \boldsymbol{\epsilon}_s^2, \boldsymbol{\epsilon}_s^3, ...\right\}_{s=1}^t$, which delivers a stable approximation. As noted above, the pruned approximated state-space system may also be expressed as an infinite moving average in terms of past innovations and is in this sense similar to the approximation in Lan and Meyer-Gohde (2011).

Finally, Haan and Wind (2010) highlight two potential disadvantages of pruning. First, the state vector is larger. Second, a pruned n-th order approximation cannot exactly fit the true solution if it happens to be an n-th-order polynomial. We do not consider the extended state vector to be a disadvantage because we find it informative to assess how important each of the higher-order effects is relative to the first-order effects and current computing power makes memory considerations less of a constraint. We view the second disadvantage to be minor because an exact fit can be obtained by rasing the approximation beyond order n (this point is also noted in Haan and Wind (2010)). As mentioned above, Haan and Wind (2010) consider σ as a constant when deriving the pruned state-space system, meaning that they would also include $\frac{3}{6}\mathbf{h}_{\sigma\sigma\mathbf{x}}\sigma^2\mathbf{x}_t^{rd}$ in (12). We treat σ as a state variable because it is consistent with the assumed functional forms for $\mathbf{g}\left(\mathbf{x}_t, \sigma\right)$ and $\mathbf{h}\left(\mathbf{x}_t, \sigma\right)$ and the way the perturbation approximation is derived. As we will show below, interpreting σ as a state variable is essential for showing that the pruned state-space system in general is stable and has well-defined statistical properties.

4 Statistical Properties of the Pruned System

This section shows that the pruned state-space system has well-defined statistical properties and that closed-form expressions for first and second moments can be easily computed. We proceed as follows. Section 4.1 extends the analysis in Kim, Kim, Schaumburg and Sims (2008) for second-order approximations. Section 4.2 conducts a similar analysis for third-order approximations. Applying the steps below to higher-order approximations is conceptually straightforward.

4.1 Second-Order Approximation

It is convenient to consider a more compact representation of the pruned state-space system than the one in section 3.1. To do so, we introduce the vector

$$\mathbf{z}_t^{(2)} \equiv \begin{bmatrix} \left(\mathbf{x}_t^f\right)' & \left(\mathbf{x}_t^s\right)' & \left(\mathbf{x}_t^f \otimes \mathbf{x}_t^f\right)' \end{bmatrix}',$$

where the superscript denotes the approximation order. The first n_x elements in $\mathbf{z}_t^{(2)}$ are the first-order effects, while the remaining part of $\mathbf{z}_t^{(2)}$ contains second-order effects. The laws of motion for $\mathbf{x}_t^f$ and $\mathbf{x}_t^s$ are stated above and the evolution for $\mathbf{x}_t^f \otimes \mathbf{x}_t^f$ is easily derived from (5). Given $\mathbf{z}_t^{(2)}$, the laws of motion for the first- and second-order effects in (5) and (7) can be rewritten as a linear law of motion in $\mathbf{z}_t^{(2)}$:

$$\mathbf{z}_{t+1}^{(2)} = \mathbf{A}^{(2)}\mathbf{z}_t^{(2)} + \mathbf{B}^{(2)}\boldsymbol{\xi}_{t+1}^{(2)} + \mathbf{c}^{(2)}, \tag{15}$$

where

$$\mathbf{c}^{(2)} \equiv \begin{bmatrix} \mathbf{0} \\ \frac{1}{2}\mathbf{h}_{\sigma\sigma}\sigma^2 \\ (\sigma\boldsymbol{\eta} \otimes \sigma\boldsymbol{\eta})\, vec\,(\mathbf{I}_{n_e}) \end{bmatrix} \quad \mathbf{A}^{(2)} \equiv \begin{bmatrix} \mathbf{h_x} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{h_x} & \frac{1}{2}\mathbf{H_{xx}} \\ \mathbf{0} & \mathbf{0} & \mathbf{h_x} \otimes \mathbf{h_x} \end{bmatrix}, \tag{16}$$

and

$$\mathbf{B}^{(2)} \equiv \begin{bmatrix} \sigma\boldsymbol{\eta} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \sigma\boldsymbol{\eta} \otimes \sigma\boldsymbol{\eta} & \sigma\boldsymbol{\eta} \otimes \mathbf{h_x} & \mathbf{h_x} \otimes \sigma\boldsymbol{\eta} \end{bmatrix} \quad \boldsymbol{\xi}_{t+1}^{(2)} \equiv \begin{bmatrix} \boldsymbol{\epsilon}_{t+1} \\ \boldsymbol{\epsilon}_{t+1} \otimes \boldsymbol{\epsilon}_{t+1} - vec\,(\mathbf{I}_{n_e}) \\ \boldsymbol{\epsilon}_{t+1} \otimes \mathbf{x}_t^f \\ \mathbf{x}_t^f \otimes \boldsymbol{\epsilon}_{t+1} \end{bmatrix}. \tag{17}$$

Also, the law of motion in (9) can be written as a linear function of $\mathbf{z}_t^{(2)}$:

$$\mathbf{y}_t^s = \mathbf{C}^{(2)}\mathbf{z}_t^{(2)} + \mathbf{d}^{(2)}, \tag{18}$$

where

$$\mathbf{C}^{(2)} \equiv \begin{bmatrix} \mathbf{g_x} & \mathbf{g_x} & \frac{1}{2}\mathbf{G_{xx}} \end{bmatrix} \quad \mathbf{d}^{(2)} \equiv \frac{1}{2}\mathbf{g}_{\sigma\sigma}\sigma^2. \tag{19}$$

Standard properties for the Kronecker product and block matrices imply that the system in (15) is stable with all eigenvalues of $\mathbf{A}^{(2)}$ having modulus less than one, provided the same holds for $\mathbf{h_x}$. This result might also be directly inferred from (5) and (7) because $\mathbf{x}_t^f$ is stable by assumption and $\mathbf{x}_t^s$ is constructed from a stable process and the autoregressive part of $\mathbf{x}_t^s$ is stable. The system has finite first and second moments if $\boldsymbol{\xi}_{t+1}^{(2)}$ has finite second moments. The latter is equivalent to $\boldsymbol{\epsilon}_{t+1}$ having finite fourth moments. These results are summarized in the next proposition, with the proof given in appendix C.

Proposition 1 *If all eigenvalues of $\mathbf{h_x}$ have modulus less than one and $\boldsymbol{\epsilon}_{t+1}$ has finite fourth moments, the pruned state-space system defined by (5), (7), and (9) - or equivalently by (15) and (18) - has finite first and second moments.*

Proposition 1 implies that explosive sample paths do not appear in the pruned state-space system (almost surely). Proposition 1 also holds for models with deterministic and stochastic trends, provided trending variables are appropriately scaled (see King and Rebelo (1999)).

The representation in (15) and (18) makes it straightforward to derive additional statistical properties for the system. Importantly, the pruned state-space system has finite third and fourth moments if $\boldsymbol{\xi}_{t+1}^{(2)}$ has finite third and fourth moments, respectively. The latter is equivalent to $\boldsymbol{\epsilon}_{t+1}$ having finite sixth and eighth moments, respectively. We highlight these results in the next proposition, with the proof provided in appendix D.

Proposition 2 *If all eigenvalues of* $\mathbf{h_x}$ *have modulus less than one and* $\boldsymbol{\epsilon}_{t+1}$ *has finite sixth and eighth moments, the pruned state-space system defined by (5), (7), and (9) - or equivalently by (15) and (18) - has finite third and fourth moments, respectively.*

The next step is to find the expressions for first and second moments of the pruned state-space system in (15) and (18). The innovations $\boldsymbol{\xi}_{t+1}^{(2)}$ are a function of $\mathbf{x}_t^f$, $\boldsymbol{\epsilon}_{t+1}$, and $\boldsymbol{\epsilon}_{t+1} \otimes \boldsymbol{\epsilon}_{t+1}$, and we directly have that $\mathbb{E}\left[\boldsymbol{\xi}_{t+1}^{(2)}\right] = \mathbf{0}$. Hence, the mean of $\mathbf{z}_t^{(2)}$ is

$$\mathbb{E}\left[\mathbf{z}_t^{(2)}\right] = \left(\mathbf{I}_{2n_x+n_x^2} - \mathbf{A}^{(2)}\right)^{-1} \mathbf{c}^{(2)}. \tag{20}$$

We explicitly compute some of the elements in $\mathbb{E}\left[\mathbf{z}_t^{(2)}\right]$ to obtain intuition for the determinants of the mean of the pruned state-space system. The mean of $\mathbf{x}_t^f$ is easily seen to be zero from (5). Equation (7) implies that the mean of $\mathbf{x}_t^s$ is

$$\mathbb{E}\left[\mathbf{x}_t^s\right] = \left(\mathbf{I} - \mathbf{h_x}\right)^{-1} \left(\frac{1}{2}\mathbf{H_{xx}}\mathbb{E}\left[\mathbf{x}_t^f \otimes \mathbf{x}_t^f\right] + \frac{1}{2}\mathbf{h}_{\sigma\sigma}\sigma^2\right). \tag{21}$$

Adding the mean value for the first- and second-order effects, we obtain

$$\mathbb{E}\left[\mathbf{x}_t^f\right] + \mathbb{E}\left[\mathbf{x}_t^s\right], \tag{22}$$

which is the mean value of the state variables in the pruned state-space system associated with the second-order approximation. Equations (21) and (22) show that second-order effects $\mathbb{E}\left[\mathbf{x}_t^s\right]$ correct the mean of the first-order effects to adjust for risk. The adjustment comes from the second derivative of the perturbation parameter $\mathbf{h}_{\sigma\sigma}$ and the mean of $\mathbf{x}_t^f \otimes \mathbf{x}_t^f$. The latter term can be computed from (5) and is given by

$$\mathbb{E}\left[\mathbf{x}_t^f \otimes \mathbf{x}_t^f\right] = \left(\mathbf{I} - \mathbf{h_x} \otimes \mathbf{h_x}\right)^{-1} \left(\sigma\boldsymbol{\eta} \otimes \sigma\boldsymbol{\eta}\right) vec\left(\mathbf{I}_{n_e}\right).$$

The fact that $\mathbb{E}\left[\mathbf{x}_t^f\right] = 0$ and $\mathbb{E}\left[\mathbf{x}_t^s\right] \neq 0$ shows that the mean of a first-order approximation is the steady state, while the mean of the pruned state-space system associated with the second-order approximation is corrected by the second moment of the innovations. Hence, the mean of the ergodic distribution implied by the pruned state-space system for any variable of interest will, in most cases, differ from its value in the deterministic steady state. Thus, we cannot, in general, use the steady state to calibrate and/or estimate the pruned system.

Let us now consider second moments. Standard properties of a VAR(1) system imply that the variance-covariance matrix for $\mathbf{z}_t^{(2)}$ is given by

$$vec\left(\mathbb{V}\left(\mathbf{z}_t^{(2)}\right)\right) = \left(\mathbf{I}_{(2n_x+n_x^2)^2} - \left(\mathbf{A}^{(2)}\otimes\mathbf{A}^{(2)}\right)\right)^{-1} vec\left(\mathbf{B}^{(2)}\mathbb{V}\left(\boldsymbol{\xi}_t^{(2)}\right)\left(\mathbf{B}^{(2)}\right)'\right), \tag{23}$$

or

$$\mathbb{V}\left(\mathbf{z}_t^{(2)}\right) = \mathbf{A}^{(2)}\mathbb{V}\left(\mathbf{z}_t^{(2)}\right)\left(\mathbf{A}^{(2)}\right)' + \mathbf{B}^{(2)}\mathbb{V}\left(\boldsymbol{\xi}_t^{(2)}\right)\left(\mathbf{B}^{(2)}\right)' \tag{24}$$

because $\mathbf{z}_t^{(2)}$ and $\boldsymbol{\xi}_{t+1}^{(2)}$ are uncorrelated as $\boldsymbol{\epsilon}_{t+1}$ is independent across time. Appendix C explains how to calculate $\mathbb{V}\left(\boldsymbol{\xi}_t^{(2)}\right)$, a rather direct yet tedious exercise. Once we know $\mathbb{V}\left(\boldsymbol{\xi}_t^{(2)}\right)$, we solve for $\mathbb{V}\left(\mathbf{z}_t^{(2)}\right)$ by standard methods for discrete Lyapunov equations.

Our procedure for computing $\mathbb{V}\left(\mathbf{z}_t^{(2)}\right)$ differs slightly from the one in Kim, Kim, Schaumburg and Sims (2008). They suggest using a second-order approximation to $\mathbb{V}\left(\boldsymbol{\xi}_t^{(2)}\right)$ by letting the last n_x^2 elements in $\boldsymbol{\xi}_t^{(2)}$ be zero. This eliminates all third- and fourth-order terms related to $\boldsymbol{\epsilon}_{t+1}$ and seems inconsistent with the fact that $\mathbf{A}^{(2)}\otimes\mathbf{A}^{(2)}$ contains third- and fourth-order terms. We therefore prefer to compute $\mathbb{V}\left(\boldsymbol{\xi}_t^{(2)}\right)$ without further approximations, implying that $\mathbb{V}\left(\mathbf{z}_t^{(2)}\right)$ corresponds to the sample moment in a long simulation using the pruned state-space system. Thus, combining first- and second-order effects, we have that the variance of the state variables in the pruned state space system related to the second-order approximation is

$$\mathbb{V}\left(\mathbf{x}_t^f\right) + \mathbb{V}\left(\mathbf{x}_t^s\right) + Cov\left(\mathbf{x}_t^f, \mathbf{x}_t^s\right) + Cov\left(\mathbf{x}_t^s, \mathbf{x}_t^f\right).$$

The auto-covariances for $\mathbf{z}_t^{(2)}$ are easily shown to be

$$Cov\left(\mathbf{z}_{t+l}^{(2)}, \mathbf{z}_t^{(2)}\right) = \left(\mathbf{A}^{(2)}\right)^l \mathbb{V}\left(\mathbf{z}_t^{(2)}\right) \quad \text{for } l = 1, 2, 3, ...$$

because $\mathbf{z}_t^{(2)}$ and $\boldsymbol{\xi}_{t+l}^{(2)}$ are uncorrelated for $l = 1, 2, 3, ...$, given that $\boldsymbol{\epsilon}_{t+1}$ is independent across time.

Finally, closed-form expressions for all corresponding moments related to $\mathbf{y}_t^s$ follow directly from the linear relationship between $\mathbf{y}_t^s$ and $\mathbf{z}_t^{(2)}$ in (18). That is,

$$\mathbb{E}\left[\mathbf{y}_t^s\right] = \mathbf{C}^{(2)}\mathbb{E}\left[\mathbf{z}_t^{(2)}\right] + \mathbf{d}^{(2)},$$

$$\mathbb{V}\left[\mathbf{y}_t^s\right] = \mathbf{C}^{(2)}\mathbb{V}\left[\mathbf{z}_t\right]\left(\mathbf{C}^{(2)}\right)',$$

and

$$Cov\left(\mathbf{y}_t^s, \mathbf{y}_{t+l}^s\right) = \mathbf{C}^{(2)}Cov\left(\mathbf{z}_{t+l}^{(2)}, \mathbf{z}_t^{(2)}\right)\left(\mathbf{C}^{(2)}\right)' \quad \text{for } l = 1, 2, 3, ...$$

4.2 Third-Order Approximation

As we did for the second-order approximation, we start by deriving a more compact representation for the pruned state-space system than the one in section 3.2. This is done based on the vector

$$\mathbf{z}_t^{(3)} \equiv \begin{bmatrix} \left(\mathbf{x}_t^f\right)' & \left(\mathbf{x}_t^s\right)' & \left(\mathbf{x}_t^f \otimes \mathbf{x}_t^f\right)' & \left(\mathbf{x}_t^{rd}\right)' & \left(\mathbf{x}_t^f \otimes \mathbf{x}_t^s\right)' & \left(\mathbf{x}_t^f \otimes \mathbf{x}_t^f \otimes \mathbf{x}_t^f\right)' \end{bmatrix}', \tag{25}$$

where the first part reproduces $\mathbf{z}_t^{(2)}$ and the last three components denote third-order effects. The law of motion for $\mathbf{x}_t^{rd}$ was derived in section 3.2, and recursions for $\mathbf{x}_t^f \otimes \mathbf{x}_t^s$ and $\mathbf{x}_t^f \otimes \mathbf{x}_t^f \otimes \mathbf{x}_t^f$ follow from (5) and (7). Hence, the law of motion for $\mathbf{x}_t^f$, $\mathbf{x}_t^s$, and $\mathbf{x}_t^{rd}$ in (5), (7), and (12), respectively, can be represented by a VAR(1) system

$$\mathbf{z}_{t+1}^{(3)} = \mathbf{A}^{(3)}\mathbf{z}_t^{(3)} + \mathbf{B}^{(3)}\boldsymbol{\xi}_{t+1}^{(3)} + \mathbf{c}^{(3)}. \tag{26}$$

We also have that the control variables are linear in $\mathbf{z}_t^{(3)}$:

$$\mathbf{y}_t^{rd} = \mathbf{C}^{(3)}\mathbf{z}_t^{(3)} + \mathbf{d}^{(3)}. \tag{27}$$

The expressions for $\mathbf{A}^{(3)}$, $\mathbf{b}^{(3)}$, $\boldsymbol{\xi}_{t+1}^{(3)}$, $\mathbf{C}^{(3)}$, and $\mathbf{d}^{(3)}$ are provided in appendix E.

Appendix F shows that the system in (26) is stable, with all eigenvalues of $\mathbf{A}^{(3)}$ having modulus less than one, provided the same holds for $\mathbf{h_x}$. Building on the intuition from the second-order approximation, this result follows from the fact that the new component of the state vector $\mathbf{x}_t^{rd}$ is constructed from stable processes and its autoregressive component is also stable. The stability of $\mathbf{x}_t^{rd}$ relies heavily on σ being treated as a variable when setting up the pruned state-space system. If, instead, we had have followed Haan and Wind (2010) and included the term $\frac{3}{6}\mathbf{h}_{\sigma\sigma\mathbf{x}}\sigma^2\mathbf{x}_t^{rd}$ in the law of motion for $\mathbf{x}_{t+1}^{rd}$, then $\mathbf{x}_{t+1}^{rd}$ would have had the autoregressive matrix $\mathbf{h_x} + \frac{3}{6}\mathbf{h}_{\sigma\sigma\mathbf{x}}\sigma^2$ which may have eigenvalues with modulus greater than one even when $\mathbf{h_x}$ is stable. Moreover, the system in (26) and (27) has finite first and second moments if $\boldsymbol{\xi}_{t+1}^{(3)}$ has finite second moments. The latter is equivalent to $\boldsymbol{\epsilon}_{t+1}$ having finite sixth moments. These results are summarized in the next proposition, with the proof given in appendix F.

Proposition 3 *If all eigenvalues of $\mathbf{h_x}$ have modulus less than one and $\boldsymbol{\epsilon}_{t+1}$ has finite sixth moments, the pruned state-space system defined by (5), (7), (12), and (14) - or equivalently by (26) and (27) - has finite first and second moments.*

The representation in (26) and (27) of the pruned state-space system and the fact that $\boldsymbol{\xi}_{t+1}^{(3)}$ is a function of $\boldsymbol{\epsilon}_{t+1} \otimes \boldsymbol{\epsilon}_{t+1} \otimes \boldsymbol{\epsilon}_{t+1}$ make it straightforward to derive additional statistical properties for the system, with the proof provided in appendix G.

Proposition 4 *If all eigenvalues of $\mathbf{h_x}$ have modulus less than one and $\boldsymbol{\epsilon}_{t+1}$ has finite ninth and twelfth moments, the pruned state-space system defined by (5), (7), (12), and (14) - or equivalently by (26) and (27) - has finite third and fourth moments, respectively.*

The next step is to explicitly compute the first and second moments of the pruned state-space system defined by (26) and (27). According to appendix E, the innovations $\boldsymbol{\xi}_{t+1}^{(3)}$ in (26) are a function of $\mathbf{x}_t^f$, $\mathbf{x}_t^s$, $\boldsymbol{\epsilon}_{t+1}$, $\boldsymbol{\epsilon}_{t+1} \otimes \boldsymbol{\epsilon}_{t+1}$, and $\boldsymbol{\epsilon}_{t+1} \otimes \boldsymbol{\epsilon}_{t+1} \otimes \boldsymbol{\epsilon}_{t+1}$. Hence, we have that $\mathbb{E}\left[\boldsymbol{\xi}_{t+1}^{(3)}\right] = \mathbf{0}$ and

$$\mathbb{E}\left[\mathbf{z}_t^{(3)}\right] = \left(\mathbf{I}_{3n_x+2n_x^2+n_x^3} - \mathbf{A}^{(3)}\right)^{-1}\mathbf{c}^{(3)}. \tag{28}$$

It is interesting to explore the value of $\mathbb{E}\left[\mathbf{x}_t^{rd}\right]$ as it may change the mean of the state variables.

From (12), we immediately have

$$\mathbb{E}\left[\mathbf{x}_t^{rd}\right] = \left(\mathbf{I}_{n_x} - \mathbf{h}_\mathbf{x}\right)^{-1}\left(\mathbf{H}_{\mathbf{xx}}\mathbb{E}\left[\mathbf{x}_t^f \otimes \mathbf{x}_t^s\right] + \frac{1}{6}\mathbf{H}_{\mathbf{xxx}}\mathbb{E}\left[\mathbf{x}_t^f \otimes \mathbf{x}_t^f \otimes \mathbf{x}_t^f\right] + \frac{1}{6}\mathbf{h}_{\sigma\sigma\sigma}\sigma^3\right), \tag{29}$$

and simple algebra gives

$$\mathbb{E}\left[\mathbf{x}_t^f \otimes \mathbf{x}_t^s\right] = \left(\mathbf{I}_{n_x^2} - \left(\mathbf{h}_\mathbf{x} \otimes \mathbf{h}_\mathbf{x}\right)\right)^{-1}\left(\mathbf{h}_\mathbf{x} \otimes \frac{1}{2}\mathbf{H}_{\mathbf{xx}}\right)\mathbb{E}\left[\mathbf{x}_t^f \otimes \mathbf{x}_t^f \otimes \mathbf{x}_t^f\right] \tag{30}$$

and

$$\mathbb{E}\left[\mathbf{x}_t^f \otimes \mathbf{x}_t^f \otimes \mathbf{x}_t^f\right] = \left(\mathbf{I}_{n_x^3} - \left(\mathbf{h}_\mathbf{x} \otimes \mathbf{h}_\mathbf{x} \otimes \mathbf{h}_\mathbf{x}\right)\right)^{-1}\left(\sigma\boldsymbol{\eta} \otimes \boldsymbol{\sigma\eta} \otimes \boldsymbol{\sigma\eta}\right)\mathbb{E}\left[\boldsymbol{\epsilon}_{t+1} \otimes \boldsymbol{\epsilon}_{t+1} \otimes \boldsymbol{\epsilon}_{t+1}\right]. \tag{31}$$

Adding the mean value for the first-, second- and third-order effects, we obtain

$$\mathbb{E}\left[\mathbf{x}_t^f\right] + \mathbb{E}\left[\mathbf{x}_t^s\right] + \mathbb{E}\left[\mathbf{x}_t^{rd}\right], \tag{32}$$

which is the mean value of the state variables in the pruned state-space system associated with the third-order approximation. If we next consider the standard case where all innovations have symmetric probability distributions and hence zero third moments, then $\mathbb{E}\left[\boldsymbol{\epsilon}_{t+1} \otimes \boldsymbol{\epsilon}_{t+1} \otimes \boldsymbol{\epsilon}_{t+1}\right] = \mathbf{0}$, which in turn implies $\mathbb{E}\left[\mathbf{x}_t^f \otimes \mathbf{x}_t^f \otimes \mathbf{x}_t^f\right] = \mathbf{0}$ and $\mathbb{E}\left[\mathbf{x}_t^f \otimes \mathbf{x}_t^s\right] = \mathbf{0}$. Furthermore, $\mathbf{h}_{\sigma\sigma\sigma}$ and $\mathbf{g}_{\sigma\sigma\sigma}$ are also zero based on the results in Andreasen (2012) when all innovations have symmetric probability distributions. Thus $\mathbb{E}\left[\mathbf{x}_t^{rd}\right] = \mathbf{0}$. As a result, the mean of the state vector is not further corrected by the third-order effects when all innovations have zero third moments. A similar property holds for the control variables because they are a linear function of $\mathbf{x}_t^{rd}$, $\mathbf{x}_t^f \otimes \mathbf{x}_t^s$, and $\mathbf{x}_t^f \otimes \mathbf{x}_t^f \otimes \mathbf{x}_t^f$. On the other hand, if one or several innovations have non-symmetric probability distributions, then $\mathbf{h}_{\sigma\sigma\sigma}$ and $\mathbf{g}_{\sigma\sigma\sigma}$ may be non-zero and $\mathbb{E}\left[\mathbf{x}_t^{rd}\right] \neq \mathbf{0}$, implying that the mean has an additional correction for risk by the third-order effects. These results are summarized in the next proposition.

Proposition 5 *If all third moments of $\boldsymbol{\epsilon}_t$ are zero, the mean values in the pruned state-space system defined by (5), (7), (12), and (14) - or equivalently by (26) and (27) - are identical to those in the pruned state-space system defined by (5), (7), and (9) - or equivalently by (15) and (18).*

This proposition is useful when calibrating or estimating DSGE models with symmetric probability distributions because it shows that the third-order effects do not affect the mean of the variables of interest.

Let us now consider second moments. The expression for the variance-covariance matrix of the $\mathbf{z}_t^{(3)}$ is slightly more complicated than the one for $\mathbf{z}_t^{(2)}$ because $\mathbf{z}_t^{(3)}$ is correlated with $\boldsymbol{\xi}_{t+1}^{(3)}$. This correlation arises from terms of the form $\mathbf{x}_t^f \otimes \boldsymbol{\epsilon}_{t+1} \otimes \boldsymbol{\epsilon}_{t+1}$ in $\boldsymbol{\xi}_{t+1}^{(3)}$ which are correlated with elements in $\mathbf{z}_t^{(3)}$. Hence,

$$\begin{aligned}\mathbb{V}\left(\mathbf{z}_t^{(3)}\right) &= \mathbf{A}^{(3)}\mathbb{V}\left(\mathbf{z}_t^{(3)}\right)\left(\mathbf{A}^{(3)}\right)' + \mathbf{B}^{(3)}\mathbb{V}\left(\boldsymbol{\xi}_t^{(3)}\right)\left(\mathbf{B}^{(3)}\right)' \\ &\quad \mathbf{A}^{(3)}Cov\left(\mathbf{z}_t^{(3)}, \boldsymbol{\xi}_{t+1}^{(3)}\right)\left(\mathbf{B}^{(3)}\right)' + \mathbf{B}^{(3)}Cov\left(\boldsymbol{\xi}_{t+1}^{(3)}, \mathbf{z}_t^{(3)}\right)\left(\mathbf{A}^{(3)}\right)'. \end{aligned} \tag{33}$$

The expressions for $\mathbb{V}\left(\boldsymbol{\xi}_t^{(3)}\right)$ and $Cov\left(\boldsymbol{\xi}_{t+1}^{(3)}, \mathbf{z}_t^{(3)}\right)$ are provided in appendix F. Thus, combining first-, second-, and third-order effects, we have that the variance of the state variables in the pruned state space system related to the third-order approximation is:

$$\begin{aligned}&\mathbb{V}\left(\mathbf{x}_t^f\right) + \mathbb{V}\left(\mathbf{x}_t^s\right) + \mathbb{V}\left(\mathbf{x}_t^{rd}\right) + Cov\left(\mathbf{x}_t^f, \mathbf{x}_t^s\right) + Cov\left(\mathbf{x}_t^f, \mathbf{x}_t^{rd}\right) \\ &+Cov\left(\mathbf{x}_t^s, \mathbf{x}_t^f\right) + Cov\left(\mathbf{x}_t^s, \mathbf{x}_t^{rd}\right) + Cov\left(\mathbf{x}_t^{rd}, \mathbf{x}_t^f\right) + Cov\left(\mathbf{x}_t^{rd}, \mathbf{x}_t^s\right).\end{aligned}$$

The auto-covariances for $\mathbf{z}_t^{(3)}$ are

$$Cov\left(\mathbf{z}_{t+s}^{(3)}, \mathbf{z}_t^{(3)}\right) = \left(\mathbf{A}^{(3)}\right)^s \mathbb{V}\left[\mathbf{z}_t^{(3)}\right] + \textstyle\sum_{j=0}^{s-1}\left(\mathbf{A}^{(3)}\right)^{s-1-j}\mathbf{B}^{(3)}Cov\left(\boldsymbol{\xi}_{t+1+j}^{(3)}, \mathbf{z}_t^{(3)}\right)$$

for $s = 1, 2, 3, ...$ The derivation of $Cov\left(\boldsymbol{\xi}_{t+l}^{(3)}, \mathbf{z}_t^{(3)}\right)$ for $l = 1, 2, 3, ...$ is given in appendix F.

Finally, closed-form expressions for all corresponding moments related to $\mathbf{y}_t^{rd}$ follow directly from the linear relationship between $\mathbf{y}_t^{rd}$ and $\mathbf{z}_t^{(3)}$ in (27) and are given by

$$\mathbb{E}\left[\mathbf{y}_t^{rd}\right] = \mathbf{C}^{(3)}\mathbb{E}\left[\mathbf{z}_t^{(3)}\right] + \mathbf{d}^{(3)},$$

$$\mathbb{V}\left[\mathbf{y}_t^{rd}\right] = \mathbf{C}^{(3)}\mathbb{V}\left[\mathbf{z}_t^{(3)}\right]\left(\mathbf{C}^{(3)}\right)',$$

and

$$Cov\left(\mathbf{y}_t^{rd}, \mathbf{y}_{t+l}^{rd}\right) = \mathbf{C}^{(3)}Cov\left(\mathbf{z}_{t+l}^{(3)}, \mathbf{z}_t^{(3)}\right)\left(\mathbf{C}^{(3)}\right)' \quad \text{for } l = 1, 2, 3, ..$$

5 Impulse Response Functions

Another fruitful way to study the properties of DSGE models is to look at their IRFs. For the first-order approximation, these functions have simple expressions where the effects of shocks are scalable, symmetric, and independent of the state of the economy. For higher-order approximations, no simple expressions exist for these functions and simulation is typically required. This section shows that the use of the pruning method allows us to derive closed-form solutions for these functions and hence avoid the use of simulation.

Throughout this section, we consider the generalized impulse response function (GIRF) proposed by Koop, Pesaran and Potter (1996). The GIRF for any variable in the model **var** (either a state or control variable) in period $t+l$ to a shock of size $\boldsymbol{\nu}$ in period $t+1$ is defined as

$$GIRF_{\mathbf{var}}\left(l,\boldsymbol{\nu},\mathbf{w}_t\right) = \mathbb{E}\left[\mathbf{var}_{t+l}|\mathbf{w}_t,\boldsymbol{\epsilon}_{t+1}=\boldsymbol{\nu}\right] - \mathbb{E}\left[\mathbf{var}_{t+l}|\mathbf{w}_t\right], \tag{34}$$

where $\mathbf{w}_t$ denotes the required state variables in period t. As we will see below, the content of $\mathbf{w}_t$ depends on the number of state variables in the model and the approximation order.

5.1 Second-Order Approximation

The derivation uses the pruned state-space system for a second-order approximation given by (5), (7), and (9). We start by deriving the GIRF for the state variables. The GIRF for the first-order effects $\mathbf{x}_t^f$ is just

$$GIRF_{\mathbf{x}^f}\left(l,\boldsymbol{\nu}\right) = \mathbb{E}\left[\mathbf{x}_{t+l}^f|\mathbf{x}_t^f,\boldsymbol{\epsilon}_{t+1}=\boldsymbol{\nu}\right] - \mathbb{E}\left[\mathbf{x}_{t+l}^f|\mathbf{x}_t^f\right] = \mathbf{h}_{\mathbf{x}}^{l-1}\sigma\boldsymbol{\eta}\boldsymbol{\nu}. \tag{35}$$

This $GIRF_{\mathbf{x}^f}$ is unaffected by the state variable $\mathbf{x}_t^f$ because it enters symmetrically in the two conditional expectations in (35).

For the second-order effects $\mathbf{x}_t^s$, we have from (7) that

$$\mathbf{x}_{t+l}^s = \mathbf{h}_{\mathbf{x}}^l\mathbf{x}_t^s + \sum_{j=1}^{l-1}\mathbf{h}_{\mathbf{x}}^{l-1-j}\frac{1}{2}\mathbf{H}_{\mathbf{xx}}\left(\mathbf{x}_{t+j}^f\otimes\mathbf{x}_{t+j}^f\right) + \frac{1}{2}\mathbf{h}_{\sigma\sigma}\sigma^2\sum_{j=0}^{l-1}\mathbf{h}_{\mathbf{x}}^{l-1-j}. \tag{36}$$

The GIRF for $\mathbf{x}_t^f \otimes \mathbf{x}_t^f$ is derived in appendix H, showing that

$$\begin{aligned} GIRF_{\mathbf{x}^f\otimes\mathbf{x}^f}\left(l,\boldsymbol{\nu},\mathbf{x}_t^f\right) &= \mathbf{h}_{\mathbf{x}}^l\mathbf{x}_t^f\otimes\mathbf{h}_{\mathbf{x}}^{l-1}\sigma\boldsymbol{\eta\nu}+\mathbf{h}_{\mathbf{x}}^{l-1}\sigma\boldsymbol{\eta\nu}\otimes\mathbf{h}_{\mathbf{x}}^l\mathbf{x}_t^f \\ &\quad+\left(\mathbf{h}_{\mathbf{x}}^{l-1}\otimes\mathbf{h}_{\mathbf{x}}^{l-1}\right)\left((\sigma\boldsymbol{\eta\nu}\otimes\sigma\boldsymbol{\eta\nu})-E_t\left[\sigma\boldsymbol{\eta\epsilon}_{t+1}\otimes\sigma\boldsymbol{\eta\epsilon}_{t+1}\right]\right). \end{aligned} \tag{37}$$

Using this expression and (36), we then get

$$GIRF_{\mathbf{x}^s}\left(l,\boldsymbol{\nu},\mathbf{x}_t^f\right)=\sum_{j=1}^{l-1}\mathbf{h}_{\mathbf{x}}^{l-1-j}\frac{1}{2}\mathbf{H}_{\mathbf{xx}}GIRF_{\mathbf{x}^f\otimes\mathbf{x}^f}\left(j,\boldsymbol{\nu},\mathbf{x}_t^f\right). \tag{38}$$

Expressions (37)-(38) reveal three interesting implications. First, the GIRF for the second-order effect is not scalable as $GIRF_{\mathbf{x}^f\otimes\mathbf{x}^f}\left(l,\tau\times\boldsymbol{\nu},\mathbf{x}_t^f\right)\neq\tau\times GIRF_{\mathbf{x}^f\otimes\mathbf{x}^f}\left(l,\boldsymbol{\nu},\mathbf{x}_t^f\right)$ for $\tau\in\mathbb{R}$. Second, the term $(\sigma\boldsymbol{\eta\nu}\otimes\sigma\boldsymbol{\eta\nu})$ implies that the IRF for $\mathbf{x}_t^s$ is not symmetric in positive and negative shocks. Third, $\mathbf{x}_t^f$ enters in the expression for $GIRF_{\mathbf{x}^s}$, meaning that it depends on the first-order effects of the state variable. Adding the GIRF for the first- and second-order effects, we obtain

$$GIRF_{\mathbf{x}^f}(l,\boldsymbol{\nu})+GIRF_{\mathbf{x}^s}\left(l,\boldsymbol{\nu},\mathbf{x}_t^f\right), \tag{39}$$

which is the pruned GIRF for the state variables implied by the pruned state-space system associated with the second-order approximation.

Finally, the pruned GIRF for the control variables is easily derived from (8) and previous results

$$\begin{aligned} GIRF_{\mathbf{y}^s}\left(l,\boldsymbol{\nu},\mathbf{x}_t^f\right) &= \mathbf{g}_{\mathbf{x}}\left(GIRF_{\mathbf{x}^f}\left(l,\boldsymbol{\nu},\mathbf{x}_t^f\right)+GIRF_{\mathbf{x}^s}\left(l,\boldsymbol{\nu},\mathbf{x}_t^f\right)\right) \\ &\quad+\frac{1}{2}\mathbf{G}_{\mathbf{xx}}GIRF_{\mathbf{x}^f\otimes\mathbf{x}^f}\left(l,\boldsymbol{\nu},\mathbf{x}_t^f\right). \end{aligned} \tag{40}$$

5.2 Third-Order Approximation

To derive the GIRF's at third order, we use the pruned state-space system in (5), (7), (12), and (14). To obtain GIRF in this case we need the GIRF for $\mathbf{x}_t^{rd}$. Using (12), we first note that

$$\begin{aligned} \mathbf{x}_{t+l}^{rd} &= \mathbf{h}_{\mathbf{x}}^l\mathbf{x}_t^{rd}+\sum_{j=0}^{l-1}\mathbf{h}_{\mathbf{x}}^{l-1-j}\left[\mathbf{H}_{\mathbf{xx}}\left(\mathbf{x}_{t+j}^f\otimes\mathbf{x}_{t+j}^s\right)+\frac{1}{6}\mathbf{H}_{\mathbf{xxx}}\left(\mathbf{x}_{t+j}^f\otimes\mathbf{x}_{t+j}^f\otimes\mathbf{x}_{t+j}^f\right)\right] \\ &\quad+\sum_{j=0}^{l-1}\mathbf{h}_{\mathbf{x}}^{l-1-j}\left[\frac{3}{6}\mathbf{h}_{\sigma\sigma\mathbf{x}}\sigma^2\mathbf{x}_{t+j}^f+\frac{1}{6}\mathbf{h}_{\sigma\sigma\sigma}\sigma^3\right]. \end{aligned} \tag{41}$$

Simple algebra then implies

$$
\begin{aligned}
GIRF_{\mathbf{x}^{rd}}\left(l,\boldsymbol{\nu},\left(\mathbf{x}_t^f,\mathbf{x}_t^s\right)\right) &= \sum_{j=1}^{l-1}\mathbf{h}_{\mathbf{x}}^{l-1-j}\mathbf{H}_{\mathbf{xx}}GIRF_{\mathbf{x}^f\otimes\mathbf{x}^s}\left(j,\boldsymbol{\nu},\left(\mathbf{x}_t^f,\mathbf{x}_t^s\right)\right) \\
&\quad+\sum_{j=1}^{l-1}\mathbf{h}_{\mathbf{x}}^{l-1-j}\frac{1}{6}\mathbf{H}_{\mathbf{xxx}}GIRF_{\mathbf{x}^f\otimes\mathbf{x}^f\otimes\mathbf{x}^f}\left(j,\boldsymbol{\nu},\mathbf{x}_t^f\right) \\
&\quad+\sum_{j=1}^{l-1}\mathbf{h}_{\mathbf{x}}^{l-1-j}\frac{3}{6}\mathbf{h}_{\sigma\sigma\mathbf{x}}\sigma^2 GIRF_{\mathbf{x}^f}\left(j,\boldsymbol{\nu}\right). \qquad (42)
\end{aligned}
$$

All terms are known except for $GIRF_{\mathbf{x}^f\otimes\mathbf{x}^s}\left(j,\boldsymbol{\nu},\left(\mathbf{x}_t^f,\mathbf{x}_t^s\right)\right)$ and $GIRF_{\mathbf{x}^f\otimes\mathbf{x}^f\otimes\mathbf{x}^f}\left(j,\boldsymbol{\nu},\mathbf{x}_t^f\right)$. The derivations of these terms is deferred to appendix I because it is notationally demanding. As it was the case for the second-order approximation, the GIRF for the third-order effect is not scalable, not symmetric, and depends on the first-order effects of the state variable $\mathbf{x}_t^f$. Note that the GIRF for the third-order effects also depends on $\mathbf{x}_t^s$ because it affects $GIRF_{\mathbf{x}^f\otimes\mathbf{x}^s}\left(j,\boldsymbol{\nu},\left(\mathbf{x}_t^f,\mathbf{x}_t^s\right)\right)$. Adding the GIRF for the first-, second-, and third-order effects, we obtain

$$
GIRF_{\mathbf{x}^f}\left(l,\boldsymbol{\nu}\right)+GIRF_{\mathbf{x}^s}\left(l,\boldsymbol{\nu},\mathbf{x}_t^f\right)+GIRF_{\mathbf{x}^{rd}}\left(l,\boldsymbol{\nu},\left(\mathbf{x}_t^f,\mathbf{x}_t^s\right)\right), \qquad (43)
$$

which is the pruned GIRF for the state variables implied by the pruned state-space system associated with the third-order approximation..

The pruned GIRF for the control variables in a third-order approximation is

$$
\begin{aligned}
GIRF_{\mathbf{y}^{rd}}\left(l,\boldsymbol{\nu},\left(\mathbf{x}_t^f,\mathbf{x}_t^s\right)\right) &= \mathbf{g}_{\mathbf{x}}\left(GIRF_{\mathbf{x}^f}\left(l,\boldsymbol{\nu}\right)+GIRF_{\mathbf{x}^s}\left(l,\boldsymbol{\nu},\mathbf{x}_t^f\right)\right)+\mathbf{g}_{\mathbf{x}}GIRF_{\mathbf{x}^{rd}}\left(l,\boldsymbol{\nu},\left(\mathbf{x}_t^f,\mathbf{x}_t^s\right)\right) \\
&\quad+\frac{1}{2}\mathbf{G}_{\mathbf{xx}}\left(GIRF_{\mathbf{x}^f\otimes\mathbf{x}^f}\left(l,\boldsymbol{\nu},\mathbf{x}_t^f\right)+2GIRF_{\mathbf{x}^f\otimes\mathbf{x}^s}\left(l,\boldsymbol{\nu},\left(\mathbf{x}_t^f,\mathbf{x}_t^s\right)\right)\right) \\
&\quad+\frac{1}{6}\mathbf{G}_{\mathbf{xxx}}GIRF_{\mathbf{x}^f\otimes\mathbf{x}^f\otimes\mathbf{x}^f}\left(l,\boldsymbol{\nu},\mathbf{x}_t^f\right) \\
&\quad+\frac{3}{6}\mathbf{g}_{\sigma\sigma\mathbf{x}}\sigma^2 GIRF_{\mathbf{x}^f}\left(l,\boldsymbol{\nu}\right) \qquad (44)
\end{aligned}
$$

where all terms are known.

6 Accuracy of Pruning

Kim, Kim, Schaumburg and Sims (2008) argue that pruning should improve accuracy because the omitted terms do not in general correspond to higher-order terms in a more accurate Taylor-series expansion. However, Kim, Kim, Schaumburg and Sims (2008) do not provide numerical evidence to support this statement. To fill this gap, this section examines the accuracy of the pruning method in the case of a standard New Keynesian model, a basic workhorse of modern macroeconomics and widely used for policy analysis. We proceed as follows. A brief presentation of the model is provided in section 6.1, with additional details given in appendix J. Section 6.2 describes how Euler equation errors are computed. Section 6.3 reports results from our numerical exercise, which show that, in general, pruning improves accuracy.

6.1 A New Keynesian Model

The specific version of the considered New Keynesian model is chosen based on two requirements. First, we want a model with sizable higher-order terms, as this allow us to see any potential differences between simulating the model using the pruned and unpruned state-space system. Second, the model should not generate explosive sample paths when it is simulated using the unpruned state-space system; otherwise any accuracy tests would trivially favor a pruned approximation. To meet the first objective, we include habit formation and Epstein-Zin preferences, which also help the New Keynesian model in matching various macro and financial moments (Christiano, Eichenbaum and Evans (2005), Hordahl, Tristani and Vestin (2008), and Binsbergen, Fernandez-Villaverde, Koijen and Rubio-Ramirez (2012)). The second requirement is met by modelling sticky prices as in Rotemberg (1982). This specification does not introduce an extra endogenous state variable as with Calvo contracts. The absence of this state variable turns out to be essential for generating non-explosive sample paths using the unpruned state-space system. Of course, it should be emphasized that we cannot prove that the model without pruning is stable. Instead, we rely on extensive simulation experiments and discard the (very few) explosive paths. This implies that the results are biased in favor of the unpruned state-space system.

6.1.1 Households

We assume a representative household with Epstein-Zin preferences (Epstein and Zin (1989)). Using the formulation in Rudebusch and Swanson (2012), the value function V_t is given by

$$V_t \equiv \begin{cases} u_t + \beta \left(\mathbb{E}_t \left[V_{t+1}^{1-\phi_3} \right] \right)^{\frac{1}{1-\phi_3}} & \text{if } u_t > 0 \text{ for all } t \\ u_t - \beta \left(\mathbb{E}_t \left[(-V_{t+1})^{1-\phi_3} \right] \right)^{\frac{1}{1-\phi_3}} & \text{if } u_t < 0 \text{ for all } t \end{cases} \tag{45}$$

As Rudebusch and Swanson (2012), we let the periodic utility function display separability between consumption c_t and hours worked h_t

$$u_t \equiv d_t \frac{(c_t - bc_{t-1})^{1-\phi_2}}{1-\phi_2} + (z_t^*)^{(1-\phi_2)} \phi_0 \frac{(1-h_t)^{1-\phi_1}}{1-\phi_1}. \tag{46}$$

Here, b is the parameter governing the internal habit formation and d_t is a preference shock where

$$\log d_{t+1} = \rho_d \log d_t + \epsilon_{d,t+1}$$

and $\epsilon_{d,t} \sim \mathcal{IID}\left(0, \sigma_d^2\right)$. The variable $(z_t^*)^{(1-\phi_2)}$ controls the deterministic trend in the economy and will be further specified below. Scaling the second term in (46) by $(z_t^*)^{(1-\phi_2)}$ ensures the presence of a balanced growth path and can be justified based on home production (see Rudebusch and Swanson (2012) for further details).

The budget constraint at time t reads

$$c_t + \frac{i_t}{\Upsilon_t} + \int D_{t,t+1} x_{t+1} d\omega_{t,t+1} = w_t h_t + r_t^k k_t + \frac{x_t}{\pi_t} + div_t. \tag{47}$$

Resources are spent on consumption, investment i_t, and on a portfolio of state-contingent claims X_{t+1} for the next period, which pay one unit of cash given events $\omega_{t,t+1}$. These claims are sold at prices $D_{t,t+1}$, meaning that their real costs are

$$\frac{\int D_{t,t+1} X_{t+1} d\omega_{t,t+1}}{P_t} = \int D_{t,t+1} x_{t+1} d\omega_{t,t+1}$$

where $x_{t+1} \equiv X_{t+1}/P_t$. The variable Υ_t denotes a deterministic trend in the real relative price of investment, that is, $\log \Upsilon_{t+1} = \log \Upsilon_t + \log \mu_{\Upsilon,ss}$. Letting w_t denote the real wage and r_t^k the real price of capital k_t, resources consist of i) real labor income $w_t h_t$, ii) real income from capital services sold to firms $r_t^k k_t$, iii) real payoffs from state-contingent claims purchased in the previous period x_t/π_t, and iv) dividends from firms div_t. Here, π_t is gross inflation, i.e. $\pi_t \equiv P_t/P_{t-1}$.

The law of motion for k_t is given by

$$k_{t+1} = (1-\delta) k_t + i_t - \frac{\kappa}{2} \left(\frac{i_i}{k_t} - \psi \right)^2 k_t, \tag{48}$$

where $\kappa \geq 0$ introduces capital adjustment costs based on i_i/k_t as in Jermann (1998). The constant ψ ensures that these adjustment costs are zero in the steady state.

6.1.2 Firms

A perfectly competitive representative firm produces final output using a continuum of intermediate goods $y_{i,t}$ and the production function

$$y_t = \left(\int_0^1 y_{i,t}^{(\eta-1)/\eta} di \right)^{\eta/(\eta-1)}$$

with $\eta > 1$. This generates the demand function $y_{i,t} = \left(\frac{P_{i,t}}{P_t} \right)^{-\eta} y_t$, with aggregate price level:

$$P_t = \left[\int_0^1 P_{i,t}^{1-\eta} di \right]^{1/(1-\eta)}.$$

The intermediate goods are produced by monopolistic competitors using the production function

$$y_{i,t} = a_t k_{i,t}^{\theta} \left(z_t h_{i,t} \right)^{1-\theta}. \tag{49}$$

Here, z_t is a deterministic trend that follows $\log z_{t+1} = \log z_t + \log \mu_{z,ss}$, and

$$\log a_{t+1} = \rho_a \log a_t + \epsilon_{a,t+1}$$

where $\epsilon_{a,t} \sim \mathcal{IID}\left(0, \sigma_a^2\right)$. As in Altig, Christiano, Eichenbaum and Linde (2011), we define $z_t^* \equiv \Upsilon_t^{\frac{\theta}{1-\theta}} z_t$, which denotes the degree of technological process in the economy.

The intermediate firms maximize the net present value of real profit with respect to capital, labor, and prices. Following Rotemberg (1982), firms face quadratic price adjustment costs $\xi_p \geq 0$ in relation to steady-state inflation π_{ss}. Hence, the i-th firm solves

$$\max_{h_{i,t},k_{i,t},P_{i,t}} \mathbb{E}_t \sum_{j=0}^{\infty} D_{t,t+j} P_{t+j} \left[\frac{P_{i,t+j}}{P_{t+j}} y_{i,t+j} - r_t^k k_{i,t+j} - w_{t+j} h_{i,t+j} - \frac{\xi_p}{2} \left(\frac{P_{i,t+j}}{P_{i,t+j-1}} \frac{1}{\pi_{ss}} - 1 \right)^2 y_{t+j} \right]$$

subject to $y_{i,t} = \left(\frac{P_{i,t}}{P_t}\right)^{-\eta} y_t$ and (49).

6.1.3 Monetary Authority

A monetary authority conducts monetary policy using a Taylor rule of the form

$$r_{t,1} = (1-\rho_r) r_{ss} + \rho_r r_{t-1,1} + \beta_\pi \log \left(\frac{\pi_t}{\pi_{ss}} \right) + \beta_y \log \left(\frac{y_t}{z_t^* Y_{ss}} \right), \tag{50}$$

where $r_{t,1}$ is the continuously compounded one-period net nominal interest rate. The output gap is here measured in terms of output in deviations from its deterministic trend, which is equal to z_t^* times the output in the normalized steady state Y_{ss} (see appendix J).

6.1.4 Solution

To use the perturbation method, all trending variables must be transformed into stationary variables. We generally adopt the convention of letting capital letters represent transformed variables. For instance, C_t is the transformed consumption level. The only exception is the value function, where the transformed variable is denoted by $\tilde{V}_t$. After appropriate transformations to remove non-stationary variables, the model can be written in 18 equations where the control vector $\mathbf{y}_t$ contains the following 13 variables: C_t, $r_{t,1}$, π_t, h_t, Λ_t, $\mathbb{E}_t \left[\tilde{V}_{t+1}^{1-\phi_3} \right]$, I_t, Y_t, Q_t, $\tilde{V}_t$, W_t, R_t^k, and mc_t. The variables Λ_t, Q_t, and mc_t are Lagrange multipliers related to (47), (48), and (49), respectively. The state vector $\mathbf{x}_t$ has five elements: $r_{t-1,1}$, C_{t-1}, K_t, a_t, and d_t. Finally, the innovation $\boldsymbol{\epsilon}_{t+1}$ has two elements, $\epsilon_{d,t+1}$ and $\epsilon_{a,t+1}$.

6.2 Euler Equation Errors

As shown in appendix J, the equilibrium conditions for our model can be expressed by the vector-function $\mathbf{f}(\cdot)$

$$\mathbb{E}_t \left[\mathbf{f}\left(\mathbf{x}_t, \mathbf{x}_{t+1}, \mathbf{y}_t, \mathbf{y}_{t+1}\right)\right] = \mathbf{0}. \tag{51}$$

If we insert sample paths implied by any approximated decision rules in (51), then each of the equations in (51) will not in general equal zero. This difference is referred to as Euler equation errors and can be used to evaluate the performance of a given approximation. In this section we use the Euler equation errors to compare the accuracy of pruned and unpruned state-space systems associated with different orders of approximation.

In general the errors depend on the value of state variables. A standard strategy when evaluating these Euler equation errors is to plot the errors for different values of the state variables. However, this procedure is not appropriate in our context because, in general, the pruned and unpruned state-space system have different state variables. Instead, the following procedure is adopted. We first fix a sequence of innovations $\{\boldsymbol{\epsilon}_t\}_{t=1}^{T=1,000}$ and use it to simulate the pruned systems $\left\{\mathbf{x}_t^f, \mathbf{x}_t^s, \mathbf{x}_t^{rd}, \mathbf{y}_t^f, \mathbf{y}_t^s, \mathbf{y}_t^{rd}\right\}_{t=1}^{T=1,000}$ and the unpruned systems $\left\{\mathbf{x}_t^{(1)}, \mathbf{x}_t^{(2)}, \mathbf{x}_t^{(3)}, \mathbf{y}_t^{(1)}, \mathbf{y}_t^{(2)}, \mathbf{y}_t^{(3)}\right\}_{t=1}^{T=1,000}$ starting at the steady state after a burn-up. These time series are then substituted into (51) to find the Euler equation errors in a given period, where any conditional expectations are evaluated by simulating $5,000$ realizations using the pruned and unpruned state-space systems, respectively. Note that we do not need to normalize the Euler equation errors into dollar terms because we only want to compare the relative accuracy of the two approximated systems.

6.3 Accuracy Results

Table 1 reports the values of parameters used for our accuracy test. Although most values are fairly standard, table 1 is not a formal calibration, as several parameters are chosen to get strong non-linearities in the model as required to examine the accuracy of the pruning method. A proper statistical evaluation of the model will be performed in the next section. In relation to the calibration, we highlight the following decisions. The price adjustment coefficient ξ_p is chosen to match a Calvo parameter of 0.75 for a linearized model. To select the values for the preference parameters ϕ_2 and ϕ_3, we rely on the results in Swanson (2012*a*), who extends existing formulas for relative risk aversion to account for a variable labor supply. Given our utility function, we have

$$RRA^c = \frac{\phi_2}{\frac{1-b\mu_{z^*}^{-1}}{1-\beta b} + \frac{w_{ss}(1-h_{ss})}{c_{ss}}\frac{\phi_2}{\phi_1}} + \phi_3 \frac{1-\phi_2}{\frac{1-b\mu_{z^*}^{-1}}{1-\beta b} + \frac{w_{ss}(1-h_{ss})}{c_{ss}}\frac{1-\phi_2}{1-\phi_1}}, \tag{52}$$

when measuring household wealth by the present value of life time consumption.[10] Given the values in table 1, $\frac{1-b\mu_{z^*}^{-1}}{1-\beta b} \approx 1$ and $w_{ss}(1-h_{ss})/c_{ss}$ equals approximately 1.3, making ϕ_2 and ϕ_3 the key determinants behind risk aversion. We select a relatively low value for ϕ_2 ($\phi_2 = 2$) and instead introduce strong non-linearities into the model by having $\phi_3 = -100$. This gives a high relative risk aversion of $RRA^c = 60$.

< Table 1 about here >

The first block of rows in table 2 displays the root mean squared Euler equation errors (RMSEs) for each of the nine model equilibrium conditions.[11] At second-order, the RMSEs implied by the pruned system are smaller than the ones implied by the unpruned ones in four of the nine equations. This is emphasized in table 2 by bold figures. When computing the mean RMSEs across all nine equations, we find that pruning delivers better accuracy. The improved performance of pruning is more visible at third-order where it outperforms the unpruned approximation in eight of the nine equations. As a result, the mean RMSEs across all nine equations with pruning are clearly smaller. Table 2 also shows that a third-order approximation delivers higher accuracy than a second-order approximation and a second-order approximation delivers, on average, higher accuracy than a first-order approximation. The remaining part of table 2 shows that these results are robust to considering highly persistent shocks ($\rho_d = \rho_a = 0.98$) and large innovations ($\sigma_d = \sigma_a = 0.03$). Both deviations from the benchmark specification take the sample path further away from the approximation point (the steady state) and therefore constitute more challenging cases than our benchmark specification.

< Table 2 about here >

[10] This definition of relative risk aversion is preferred to the one where wealth is measured by the present value of consumption *and* leisure because Swanson (2012*b*) shows that the measure in (52) displays the highest correlation with the equity premium in a similar model.

[11] We leave out the following log-linear relations that, by construction, all have zero model errors: firm's first-order condition for labor and capital, the Taylor rule, and the production function.

From this analysis, we conclude that pruning does improve accuracy. These results are, by construction, model-specific and dependent on the chosen calibration. However, both the model and the calibration are fairly standard and we therefore conjecture that our results are representative of a large class of DSGE models of interest.

7 Econometric Implications of the Pruning Method: An Application to Asset Pricing

The econometric implications of our derivations in sections 4 and 5 are significant because various moment matching methods used in linearized DSGE models now carry over to non-linear approximations. For approximations up to third-order, this includes GMM estimation (Hansen (1982)) based on first and second moments of the variables. The work by Kim (2002) shows how optimal GMM estimation may be used to build a so-called limited information likelihood (LIL) function. Equipped with priors, we may then carry out standard Bayesian analysis from the LIL function, where the asymptotic distribution for the posterior equals the limiting distribution of GMM. Hence, our closed-form expressions for first and second moments for approximations up to third-order may also be relevant to a Bayesian researcher.

Another possibility is to match model-implied IRFs to their empirical counterparts as is done in Christiano, Eichenbaum and Evans (2005) for a linearized model. For the case of non-linear approximations, the empirical IRFs should not be computed from an SVAR model because it restricts these functions to be scalable, symmetric, and independent of the state of the economy. None of these assumptions are imposed in the flexible projection method by Jorda (2005), which therefore is promising when DSGE models are solved with non-linear terms.

All these econometric methods are attractive because we can now easily compute first and second moments or any IRFs in medium-sized DSGE models whose solution has been approximated to third-order. If we want to use higher-order moments such as skewness and kurtosis in estimation, then simulations are needed. While it is possible to compute closed-form expressions for skewness and kurtosis in DSGE models when the pruning method is applied, the memory requirement for such computations are extremely large and therefore only applicable to small models with a few state variables. But even if we simulate to get around this memory limitations, our analysis provides the foundation for SMM following Duffie and Singleton (1993) and indirect inference as considered in Smith (1993), Dridi, Guay and Renault (2007), Creel and Kristensen (2011), among others.[12]

Our results greatly simplify the computational burden required for the implementation of these econometric methods while using higher-order approximations, but they also make a theoretical contribution. This is because without adopting the pruning method, we do not know if any moments exist and the asymptotic distributions for all the aforementioned estimators cannot, in general, be adopted. To realize this, recall that the limiting distributions for GMM and SMM require stationary processes, or processes that can be transformed to be stationary, but this property cannot be ensured in non-linear approximations without the use of the pruning method.[13]

To illustrate some of these points, we estimate the model in section 6.1 with GMM and SMM using macroeconomic and financial data for the US economy from 1961Q3 to 2007Q4. To make our model slightly more standard, we abandon the quadratic price adjustment costs and instead consider staggered pricing as in Calvo (1983). That is, we now assume that a fraction $\alpha \in [0, 1[$ of randomly chosen firms cannot set the optimal nominal price of the good they produce in each period and instead let $P_{i,t} = \pi_{t-1} P_{i,t-1}$. The rest of the model is as described in section 6.1.

Seven time series are used in the estimation of our quarterly model: i) consumption growth Δc_t, ii) investment growth Δi_t, iii) inflation π_t, iv) the 1-quarter nominal interest rate $r_{t,1}$, v) the 10-year nominal interest rate $r_{t,40}$, vi) the 10-year ex post excess holding period return $xhr_{t,40} \equiv \log\left(P_{t,39}/P_{t-1,40}\right) - r_{t-1,1}$, and vii) log of hours $\log h_t$. To compute $r_{t,40}$ and $xhr_{t,40}$, we solve for nominal zero-coupon bond prices $P_{t,k}$ maturing in period $t+k$ by using the efficient perturbation algorithm in Andreasen and Zabczyk (2010). The presence of a short- and long-term interest rate jointly captures the slope of the yield curve, while the excess holding period return is included as an observable proxy for the term premium. All seven time series are stored in $\mathbf{data}_t$. We describe the empirical data series in appendix K.

[12] See also Ruge-Murcia (2012) for a Monte Carlo study and application of SMM based on the neoclassical growth model solved up to third-order.

[13] See Peralta-Alva and Santos (2012) for a summary of the literature on the relation between estimation methods and numerical errors in simulation.

7.1 Estimation Results Using GMM

We start by exploring whether our model can match the mean, the variance, contemporaneous covariances, and the persistency in the data. Hence, let

$$\mathbf{q}_t \equiv \begin{bmatrix} \{data_{i,t}\}_{i=1}^{7} \\ vech\left(\mathbf{data}_t \mathbf{data}_t'\right) \\ \{data_{i,t} data_{i,t-1}\}_{i=1}^{7} \end{bmatrix}, \tag{53}$$

and denote the structural parameters by $\boldsymbol{\theta}$. The GMM estimator is then given by

$$\boldsymbol{\theta}_{GMM} = \arg\min_{\boldsymbol{\theta}\in\boldsymbol{\Theta}} \left(\frac{1}{T}\sum_{t=1}^{T}\mathbf{q}_t - \mathbb{E}\left[\mathbf{q}_t\left(\boldsymbol{\theta}\right)\right]\right)' \mathbf{W} \left(\frac{1}{T}\sum_{t=1}^{T}\mathbf{q}_t - \mathbb{E}\left[\mathbf{q}_t\left(\boldsymbol{\theta}\right)\right]\right). \tag{54}$$

Here, $\mathbf{W}$ is a positive definite weighting matrix, $\frac{1}{T}\sum_{t=1}^{T}\mathbf{q}_t$ is a row-vector with the empirical moments, and $\mathbb{E}\left[\mathbf{q}_t\left(\boldsymbol{\theta}\right)\right]$ contains the model-implied moments that we compute in closed-form using the above formulas. Following common practice, we apply a diagonal weighting matrix to obtain preliminary consistent estimates that we use to compute the optimal weighting matrix - in our case the Newey-West estimator with 10 lags (all results are robust to using either 5 or 15 lags instead.)

All structural parameters in the model are estimated, except δ, θ, and η, which are harder to identify and therefore assigned the values in table 1. Based on the considered moments in (53), we can estimate the first four moments for the innovations when using a second-order approximation and the first six moments when the model is solved up to third order. In practice, it will typically be difficult to separately identify all the relevant moments, and we therefore impose a distributional specification. Letting the innovations to preference and technology shocks be normally distributed, we are then left with 18 parameters to be estimated based on 42 moments.

We start by estimating the model based on a second-order approximation. Given the normally distributed innovations, the mean values implied by the second- and third-order approximations are identical and only the variances differ. The great advantage of starting with a second-order approximation is that all considered moments for our model with seven state variables can be computed in just 0.03 second when using `Matlab`.[14]

The estimates are reported in the second column of table 3. We find sizable habits ($b = 0.68$),

[14] The execution time is for a 64-bit version of Matlab 12b on an Intel(R) Core(TM) i5-3210 M CPU @ 2.50Ghz.

a low Frisch elasticity of labor supply ($1/\phi_1 = 0.16$), and a moderate degree of curvature in the periodic utility function with respect to consumption ($\phi_2 = 1.59$). The Epstein-Zin parameter is large ($\phi_3 = -196$) and accounts for a high relative risk aversion of 99 which is similar to the finding in Rudebusch and Swanson (2012). Our estimates also imply a low intertemporal elasticity of substitution for consumption of 0.2. As in Justiniano and Primiceri (2008), the degree of price stickiness is high ($\alpha = 0.93$), meaning that the inflation-indexed prices are being re-optimized roughly every second year.[15] The central bank displays preferences for smoothing changes in the policy rate ($\rho_r = 0.68$) and assigns more weight to inflation than output when conducting monetary policy ($\beta_\pi = 3.99$, $\beta_y = 0.56$).

< Table 3 about here >

Table 4 provides empirical and model-implied moments to evaluate the fit of the model. The model matches all mean values extremely well, in particular, the short- and long-term interest rates of 5.6 percent and 6.9 percent, respectively. Also, our model reproduces the mean inflation rate of 3.7 percent, even though the corresponding steady-state level is much higher ($4 \log \pi_{ss} = 15.97$ percent). The difference relates to a large correction due to the highly risk-averse households. The model is also successful in matching the variability in the data, although the standard deviation in consumption growth is below the empirical value (1.4 percent compared to 2.7 percent) and the standard deviations for inflation, the short rate, and the log of hours are somewhat higher than in the data. A satisfying performance is also seen with respect to the first-order autocorrelations, where only consumption growth displays notably higher persistence than the empirical moment (0.7 compared to 0.3).

The final part of table 4 shows the contemporaneous correlations where our model in most cases reproduces the correct sign. The main exception relates to hours, which contrary to empirical evidence, is negatively correlated with consumption and investment growth and displays positive comovement with inflation and the two interest rates.

< Table 4 about here >

[15] The prices in our model change every quarter: (roughly) seven out of eight quarters with a simple indexation rule and every eighth quarter by a full re-optimization. Our model is, then, fully compatible with a high degree of price variability in the micro data.

We next use a third-order approximation and re-estimate the model on the same moments. The optimization problem is still quite tractable, partly because estimates at second-order are good starting values and partly because it only takes 0.8 second to compute the considered moments in `Matlab`. The estimation results in table 3 (third column) show only minor differences when compared to a second-order approximation. We note a small increase in the size of habits (0.69 vs. 0.71), lower curvature in the periodic utility function with respect to consumption (1.57 vs. 1.53), and lower capital adjustment costs (4.11 to 3.59). As a result, the relative risk aversion now falls to 94. This is not a surprise, as higher-order terms allow the model to capture the dynamics of the data with a lower risk aversion. The precision of most estimates as given by the asymptotic distribution greatly improves when moving to a third-order approximation. This is similar to the findings in An and Schorfheide (2007) and Andreasen (Forthcoming).

Table 4 shows that the small changes in the estimated parameters result in broadly the same fit as with a second-order approximation. Of course, a key difference between the two cases is that the third-order approximation allows for time variation in the nominal term premium, whereas it is constant at second order. We use the measure of the term premium suggested by Rudebusch and Swanson (2012) which is based on the risk-neutral bond price $\tilde{P}_{t,k}$ where payments are discounted by $r_{t,1}$, i.e.

$$\tilde{P}_{t,k} = e^{-r_{t,1}} \mathbb{E}_t \left[\tilde{P}_{t+1,k-1} \right].$$

The corresponding yield-to-maturity on this bond is $\tilde{r}_{t,k} = -\frac{1}{k} \log \tilde{P}_{t,k}$, and nominal term premium at maturity k is then given by

$$TP_{t,k} = r_{t,k} - \tilde{r}_{t,k}.$$

At third-order, we find a nominal term premium with a sizable mean of 174 annualized basis points and a quite realistic standard deviation of 20 annualized basis points.

Given that we have more moments (n) than parameters (n_p), it is possible to apply the J-test and conduct an overall specification test of our model. This is done in table 5 using $TQ \sim \chi^2_{n-n_p}$ where T refers to the number of observations and Q is the value of the objective function when using the optimal weighting matrix. Table 5 shows that the P-value for this test is 0.84 at second-order and 0.71 at third-order, meaning that our model is not rejected by the data, that is, the observed differences between empirical and model-implied moments in table 4 are not unusual given the sample variation in the empirical moments. However, this finding should be interpreted with some caution because it is well-known that the J-test displays low power and may therefore fail to reject a misspecified model (Ruge-Murcia (2007)).

<Table 5 about here>

We end this section by reporting IRFs following positive one-standard-deviation shocks to technology and preferences (figures 1 and 2). These functions are computed at first- and third-order using the formulas derived in this paper. In the case of third-order approximation we set the value of the relevant state variables at their unconditional means. With our implementation, it takes only about 1.4 seconds to compute these functions in `Matlab`. All IRFs have the expected pattern and we therefore direct our attention to the effects of higher-order terms, that is, the difference between the marked and unmarked lines. For a technology shock, we see substantial non-linear effects in consumption and investment growth. For a preference shock, notable differences appear for the interest rates and inflation.

< Figure 1 and 2 about here>

7.2 Estimation Results Using SMM

We next explore how well our model matches higher-order moments such as skewness and kurtosis in the data using a third-order approximation. The set of considered moments are therefore extended to[16]

$$\tilde{\mathbf{q}}_t \equiv \begin{bmatrix} \mathbf{q}_t \\ \left\{ data_{i,t}^3 \right\}_{i=1}^{6} \\ \left\{ data_{i,t}^4 \right\}_{i=1}^{6} \end{bmatrix}.$$

To give our model greater flexibility in matching these additional moments, we let innovations to technology and preference shocks follow the normal-inverse Gaussian distribution. This is a very flexible probability distribution with separate parameters controlling the asymmetry of the distribution and its tail heaviness (see Barndorff-Nielsen (1997) for further details). As before, these innovations are assumed to be identical and independent across time.

[16] Higher moments for $\log h_t$ are not included because they are too tightly correlated with existing moments: $corr\left(\log h_t, (\log h_t)^3\right) = 0.9995$ and $corr\left((\log h_t)^2, (\log h_t)^4\right) = 0.9995$. If we were to include them, we could face problems computing the optimal weighting matrix in a robust way because this matrix is unlikely to have full rank.

The presence of third and fourth moments in $\tilde{\mathbf{q}}_t$ implies that all moments cannot be computed in closed-form. Therefore, we resort to simulation. Then, our SMM estimator for the structural parameters is given by

$$\boldsymbol{\theta}_{SMM} = \underset{\boldsymbol{\theta}\in\boldsymbol{\Theta}}{\arg\min} \left(\frac{1}{T} \sum_{t=1}^{T} \tilde{\mathbf{q}}_t - \frac{1}{T\tau} \sum_{s=1}^{T\tau} \tilde{\mathbf{q}}_s \left(\boldsymbol{\theta}\right) \right)' \mathbf{W} \left(\frac{1}{T} \sum_{t=1}^{T} \tilde{\mathbf{q}}_t - \frac{1}{T\tau} \sum_{s=1}^{T\tau} \tilde{\mathbf{q}}_s \left(\boldsymbol{\theta}\right) \right), \tag{55}$$

where the model-implied moments are estimated by using $\frac{1}{T\tau} \sum_{s=1}^{T\tau} \tilde{\mathbf{q}}_s\left(\boldsymbol{\theta}\right)$ for $\tau > 1$. Since the normal-inverse Gaussian distribution has finite moments up to any order (given standard requirements), we know from proposition 4 that all model moments of $\tilde{\mathbf{q}}_t$ are finite.

The estimates are reported in the final column of table 3. Compared to the third-order GMM estimation, most parameters are unchanged. Only the degree of habit formation has increased (0.71 to 0.73) and so has the curvature in the periodic utility function with respect to consumption (1.53 to 1.59). For technology shocks, the skewness parameter is estimated to be 0.23 while the tail heaviness parameter is 1.25. Using standard formulas for skewness and kurtosis in the normal-inverse Gaussian distribution, we find that technology innovations display a positive skewness of 0.45 and a kurtosis of 5.31. For preference shocks, the skewness and tail parameters are estimated very inaccurately. A standard t-test shows that the skewness parameter could be zero and computing kurtosis using the tail heaviness parameter plus one standard deviation gives a kurtosis of 3.14. Hence, we find evidence of non-Gaussianity in technology shocks but not in preference shocks. These results are broadly in line with those of Justiniano and Primiceri (2008), who find time-varying volatility in technology shocks but not in preference shocks.

The final column of table 4 shows our model's ability to match first and second moments with normal-inverse Gaussian shocks is broadly similar to the case with normally distributed shocks (i.e., GMM^{3rd}). We therefore focus on the higher-order moments in table 6. The model correctly generates negative skewness in investment growth and positive skewness in excess holding period returns but is unable to match skewness for any of the other variables. The model is slightly more successful in matching values of kurtosis in consumption growth, investment growth, and excess holding period returns. Although the model is unable to perfectly match all moments, the model specification test in table 5 shows that we cannot reject the model based on our data.

8 Conclusion

This paper shows how to extend the pruning method by Kim, Kim, Schaumburg and Sims (2008) to third- and higher-order approximations. Special attention is devoted to models solved up to third order. Conditions for the existence of first and second moments are derived, and their values are provided in closed-form. The existence of higher-order moments in the form of skewness and kurtosis is also established. We also analyze impulse response functions and provide simple closed-form expressions for these functions.

The econometric implications of our findings are significant as most of the existing moment-based estimation methods for linearized DSGE models now carry over to non-linear approximations. For approximations up to third-order, this includes GMM estimation based on first and second moments and matching model-implied IRFs to their empirical counterparts. When simulations are needed, our analysis also provides the foundation for different types of indirect inference and SMM. These results are not only relevant for classical inference, as the moment conditions in optimal GMM estimation may be used to build a limited information likelihood function, from which Bayesian inference may be carried out. We therefore hope that the tools developed in this paper will be picked by other researchers who estimate DSGE models where non-linearities play a key role.

A Appendix

In this appendix, we provide further details and derivations for the paper. In particular, we will show how our state-space framework can handle a rich set of non-linearities between $\mathbf{x}_t$ and $\boldsymbol{\epsilon}_{t+1}$, we will present the proofs of different results in the main text, discuss the New Keynesian model that we use as an empirical application in depth, and describe the data for our estimation.

B Non-linearities Between State Variables and Innovations

To illustrate how non-linearities between $\mathbf{x}_t$ and $\boldsymbol{\epsilon}_{t+1}$ can be addressed in our framework, let $\mathbf{v}_t \equiv \begin{bmatrix} \mathbf{x}'_{t-1} & \boldsymbol{\epsilon}'_t \end{bmatrix}'$ be an expanded state vector where the innovations now appear as state variables. The new state equation is then given by

$$\mathbf{v}_{t+1} = \begin{bmatrix} \mathbf{h}(\mathbf{v}_t, \sigma) \\ \mathbf{0} \end{bmatrix} + \sigma \begin{bmatrix} \mathbf{0}_{n_x} \\ \mathbf{u}_{t+1} \end{bmatrix},$$

where $\mathbf{u}_{t+1} \sim IID(\mathbf{0}, \mathbf{I})$ is of dimension n_ϵ, and the new observation equation is

$$\mathbf{y}_t = \mathbf{g}(\mathbf{v}_t, \sigma).$$

Thus, any model with non-linearities between state variables and innovations may be rewritten into our notation with only linear innovations.

As an illustration, consider a neoclassical growth model with stochastic volatility. The equilibrium conditions (where we follow standard notation) are given by:

$$\begin{gathered}
c_t^{-\gamma} = \beta \mathbb{E}_t \left[c_{t+1}^{-\gamma} \left(a_{t+1} \alpha k_{t+1}^{\alpha-1} + 1 - \delta \right) \right] \\
c_t + k_{t+1} = a_t k_t^\alpha + (1-\delta) k_t \\
\log a_{t+1} = \rho \log a_t + \sigma_{a,t+1} \epsilon_{a,t+1} \\
\log \left(\frac{\sigma_{a,t+1}}{\sigma_{a,ss}} \right) = \rho_\sigma \log \left(\frac{\sigma_{a,t}}{\sigma_{a,ss}} \right) + \epsilon_{\sigma,t+1}
\end{gathered}$$

We then re–write these conditions as:

$$\begin{gathered}
c_t^{-\gamma} = \mathbb{E}_t \left[\beta c_{t+1}^{-\gamma} \left(\exp^{\left\{ \rho \log a_t + \sigma_{a,ss} \exp^{\left\{ \rho_\sigma \log \left(\frac{\sigma_{a,t}}{\sigma_{a,ss}} \right) + \epsilon_{\sigma,t+1} \right\}} \epsilon_{a,t+1} \right\}} \alpha k_{t+1}^{\alpha-1} + 1 - \delta \right) \right] \\
c_t + k_{t+1} = a_t k_t^\alpha + (1-\delta) k_t \\
\log a_t = \rho \log a_{t-1} + \sigma_{a,t} \epsilon_{a,t} \\
\log \left(\frac{\sigma_{a,t}}{\sigma_{a,ss}} \right) = \rho_\sigma \log \left(\frac{\sigma_{a,t-1}}{\sigma_{a,ss}} \right) + \epsilon_{\sigma,t} \\
\epsilon_{a,t+1} = \sigma u_{a,t+1} \\
\epsilon_{\sigma,t+1} = \sigma u_{\sigma,t+1}
\end{gathered}$$

where the extended state vector is $\mathbf{v}_t \equiv \begin{bmatrix} k_t & a_{t-1} & \sigma_{a,t-1} & \epsilon_{a,t} & \epsilon_{\sigma,t} \end{bmatrix}$ and σ is the perturbation parameter scaling the innovations $u_{a,t+1}$ and $u_{\sigma,t+1}$.

If, instead, the volatility process is specified as a GARCH(1,1) model, then the equilibrium conditions can be expressed as:

$$
\begin{aligned}
c_t^{-\gamma} = \mathbb{E}_t\left[\beta c_{t+1}^{-\gamma}\left(\exp^{\{\rho \log a_t+\sigma_{a,t+1}\epsilon_{a,t+1}\}} \alpha k_{t+1}^{\alpha-1}+1-\delta\right)\right] \\
c_t + k_{t+1} = a_t k_t^{\alpha} + (1-\delta)\, k_t \\
\log a_t = \rho \log a_{t-1} + \sigma_{a,t}\epsilon_{a,t} \\
\sigma_{a,t+1}^2 = \sigma_{a,ss}^2 (1-\rho_1) + \rho_1 \sigma_{a,t}^2 + \rho_2 \sigma_{a,t}^2 \epsilon_{a,t}^2 \\
\epsilon_{a,t+1} = \sigma u_{t+1}
\end{aligned}
$$

where the extended state vector is $\mathbf{v}_t \equiv \left[\begin{array}{llll} k_t & \sigma_{a,t} & a_{t-1} & \epsilon_{a,t} \end{array}\right]$ and σ is the perturbation parameter scaling u_{t+1}. As in Andreasen (2012), the constant term in the GARCH process is scaled by $(1-\rho_1)$ to ensure that $\sigma_{a,t} = \sigma_{a,ss}$ in the deterministic steady state where $\epsilon_{a,t}^2 = 0$.

C Proof of Proposition 1

First, note that all eigenvalues of $\mathbf{A}^{(2)}$ are strictly less than one. To see this, we work with

$$
\begin{aligned}
p(\lambda) &= \left|\mathbf{A} - \lambda \mathbf{I}_{2n_x+n_x^2}\right| \\
&= \left|\left[\begin{array}{ccc} \mathbf{h_x} - \lambda \mathbf{I}_{n_x} & \mathbf{0}_{n_x \times n_x} & \mathbf{0}_{n_x \times n_x^2} \\ \mathbf{0}_{n_x \times n_x} & \mathbf{h_x} - \lambda \mathbf{I}_{n_x} & \frac{1}{2}\mathbf{H_{xx}} \\ \mathbf{0}_{n_x^2 \times n_x} & \mathbf{0}_{n_x^2 \times n_x} & \mathbf{h_x} \otimes \mathbf{h_x} - \lambda \mathbf{I}_{n_x^2} \end{array}\right]\right| \\
&= \left|\begin{array}{cc} \mathbf{B}_{11} & \mathbf{B}_{12} \\ \mathbf{B}_{21} & \mathbf{B}_{22} \end{array}\right| \\
&= |\mathbf{B}_{11}|\,|\mathbf{B}_{22}|
\end{aligned}
$$

where we let

$$
\begin{aligned}
\mathbf{B}_{11} &\equiv \left[\begin{array}{cc} \mathbf{h_x} - \lambda \mathbf{I}_{n_x} & \mathbf{0}_{n_x \times n_x} \\ \mathbf{0}_{n_x \times n_x} & \mathbf{h_x} - \lambda \mathbf{I}_{n_x} \end{array}\right], \\
\mathbf{B}_{12} &\equiv \left[\begin{array}{c} \mathbf{0}_{n_x \times n_x^2} \\ \frac{1}{2}\mathbf{H_{xx}} \end{array}\right] \\
\mathbf{B}_{21} &\equiv \left[\begin{array}{cc} \mathbf{0}_{n_x^2 \times n_x} & \mathbf{0}_{n_x^2 \times n_x} \end{array}\right],
\end{aligned}
$$

and

$$
\mathbf{B}_{22} \equiv \mathbf{h_x} \otimes \mathbf{h_x} - \lambda \mathbf{I}_{n_x^2}.
$$

and we use the fact that

$$
\left|\begin{array}{cc} \mathbf{U} & \mathbf{C} \\ \mathbf{0} & \mathbf{Y} \end{array}\right| = |\mathbf{U}|\,|\mathbf{Y}|
$$

where $\mathbf{U}$ is an $m \times m$ matrix and $\mathbf{Y}$ is an $n \times n$ matrix. Hence,

$$
p(\lambda) = \left|\left[\begin{array}{cc} \mathbf{h_x} - \lambda \mathbf{I}_{n_x} & \mathbf{0}_{n_x \times n_x} \\ \mathbf{0}_{n_x \times n_x} & \mathbf{h_x} - \lambda \mathbf{I}_{n_x} \end{array}\right]\right| \left|\mathbf{h_x} \otimes \mathbf{h_x} - \lambda \mathbf{I}_{n_x^2}\right| = \left|\mathbf{h_x} - \lambda \mathbf{I}_{n_x}\right| \left|\mathbf{h_x} - \lambda \mathbf{I}_{n_x}\right| \left|\mathbf{h_x} \otimes \mathbf{h_x} - \lambda \mathbf{I}_{n_x^2}\right|.
$$

The eigenvalues are clearly determined from $\left|\mathbf{h_x} - \lambda \mathbf{I}_{n_x}\right| = 0$ or $\left|\mathbf{h_x} \otimes \mathbf{h_x} - \lambda \mathbf{I}_{n_x^2}\right| = 0$. The absolute value of all eigenvalues to the first problem are strictly less than one by assumption. That is $|\lambda_i| < 1$ $i = 1, 2, ..., n_x$. This is also the case for the second problem because the eigenvalues to

$\mathbf{h_x} \otimes \mathbf{h_x}$ are $\lambda_i \lambda_j$ for $i = 1, 2, ..., n_x$ and $j = 1, 2, ..., n_x$.

Given that the system is stable, the mean value is given by $\mathbb{E}\left[\mathbf{z}_t^{(2)}\right] = \left(\mathbf{I}_{2n_x+n_x^2} - \mathbf{A}^{(2)}\right)^{-1} \mathbf{c}^{(2)}$, because $\mathbb{E}\left[\boldsymbol{\xi}_{t+1}^{(2)}\right] = 0$. For the variance, we have

$$\mathbb{V}\left(\mathbf{z}_{t+1}^{(2)}\right) = \mathbf{A}^{(2)} \mathbb{V}\left(\mathbf{z}_t^{(2)}\right) \left(\mathbf{A}^{(2)}\right)' + \mathbf{B}^{(2)} \mathbb{V}\left(\boldsymbol{\xi}_{t+1}^{(2)}\right) \left(\mathbf{B}^{(2)}\right)'$$

as

$$\mathbb{E}\left[\mathbf{z}_t^{(2)} \left(\boldsymbol{\xi}_{t+1}^{(2)}\right)'\right] = \mathbb{E}\left[\begin{array}{cccc} \mathbf{x}_t^f \boldsymbol{\epsilon}_{t+1}' & \mathbf{x}_t^f \left(\boldsymbol{\epsilon}_{t+1} \otimes \boldsymbol{\epsilon}_{t+1} - vec\left(\mathbf{I}_{ne}\right)\right)' & \mathbf{x}_t^f \left(\boldsymbol{\epsilon}_{t+1} \otimes \mathbf{x}_t^f\right)' & \mathbf{x}_t^f \left(\mathbf{x}_t^f \otimes \boldsymbol{\epsilon}_{t+1}\right)' \\ \mathbf{x}_t^s \boldsymbol{\epsilon}_{t+1}' & \mathbf{x}_t^s \left(\boldsymbol{\epsilon}_{t+1} \otimes \boldsymbol{\epsilon}_{t+1} - vec\left(\mathbf{I}_{ne}\right)\right)' & \mathbf{x}_t^s \left(\boldsymbol{\epsilon}_{t+1} \otimes \mathbf{x}_t^f\right)' & \mathbf{x}_t^s \left(\mathbf{x}_t^f \otimes \boldsymbol{\epsilon}_{t+1}\right)' \\ \left(\mathbf{x}_t^f \otimes \mathbf{x}_t^f\right) \boldsymbol{\epsilon}_{t+1}' & \left(\mathbf{x}_t^f \otimes \mathbf{x}_t^f\right) \left(\boldsymbol{\epsilon}_{t+1} \otimes \boldsymbol{\epsilon}_{t+1} - vec\left(\mathbf{I}_{ne}\right)\right)' & \left(\mathbf{x}_t^f \otimes \mathbf{x}_t^f\right) \left(\boldsymbol{\epsilon}_{t+1} \otimes \mathbf{x}_t^f\right)' & \left(\mathbf{x}_t^f \otimes \mathbf{x}_t^f\right) \left(\mathbf{x}_t^f \otimes \boldsymbol{\epsilon}_{t+1}\right)' \end{array}\right] = \mathbf{0}$$

Now, we only need to compute $\mathbb{V}\left(\boldsymbol{\xi}_{t+1}^{(2)}\right)$:

$$\begin{aligned} \mathbb{V}\left(\boldsymbol{\xi}_{t+1}^{(2)}\right) &= \mathbb{E}\left[\left[\begin{array}{c} \boldsymbol{\epsilon}_{t+1} \\ \boldsymbol{\epsilon}_{t+1} \otimes \boldsymbol{\epsilon}_{t+1} - vec\left(\mathbf{I}_{n_e}\right) \\ \boldsymbol{\epsilon}_{t+1} \otimes \mathbf{x}_t^f \\ \mathbf{x}_t^f \otimes \boldsymbol{\epsilon}_{t+1} \end{array}\right] \left[\begin{array}{c} \boldsymbol{\epsilon}_{t+1} \\ \boldsymbol{\epsilon}_{t+1} \otimes \boldsymbol{\epsilon}_{t+1} - vec\left(\mathbf{I}_{n_e}\right) \\ \boldsymbol{\epsilon}_{t+1} \otimes \mathbf{x}_t^f \\ \mathbf{x}_t^f \otimes \boldsymbol{\epsilon}_{t+1} \end{array}\right]'\right] \\ &= \left[\begin{array}{cccc} \mathbf{I}_{n_e} & \mathbb{E}\left[\boldsymbol{\epsilon}_{t+1} \left(\boldsymbol{\epsilon}_{t+1} \otimes \boldsymbol{\epsilon}_{t+1}\right)'\right] & \mathbf{0} & \mathbf{0} \\ \mathbb{E}\left[\left(\boldsymbol{\epsilon}_{t+1} \otimes \boldsymbol{\epsilon}_{t+1}\right) \boldsymbol{\epsilon}_{t+1}'\right] & \mathbb{E}\left[\left(\boldsymbol{\epsilon}_{t+1} \otimes \boldsymbol{\epsilon}_{t+1} - vec\left(\mathbf{I}_{n_e}\right)\right) \left(\boldsymbol{\epsilon}_{t+1} \otimes \boldsymbol{\epsilon}_{t+1} - vec\left(\mathbf{I}_{n_e}\right)\right)'\right] & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbb{E}\left[\left(\boldsymbol{\epsilon}_{t+1} \otimes \mathbf{x}_t^f\right) \left(\boldsymbol{\epsilon}_{t+1} \otimes \mathbf{x}_t^f\right)'\right] & \mathbb{E}\left[\left(\boldsymbol{\epsilon}_{t+1} \otimes \mathbf{x}_t^f\right) \left(\mathbf{x}_t^f \otimes \boldsymbol{\epsilon}_{t+1}\right)'\right] \\ \mathbf{0} & \mathbf{0} & \mathbb{E}\left[\left(\mathbf{x}_t^f \otimes \boldsymbol{\epsilon}_{t+1}\right) \left(\boldsymbol{\epsilon}_{t+1} \otimes \mathbf{x}_t^f\right)'\right] & \mathbb{E}\left[\left(\mathbf{x}_t^f \otimes \boldsymbol{\epsilon}_{t+1}\right) \left(\mathbf{x}_t^f \otimes \boldsymbol{\epsilon}_{t+1}\right)'\right] \end{array}\right] \end{aligned}$$

We note that this variance is finite when $\boldsymbol{\epsilon}_{t+1}$ has finite fourth moment. All elements in this matrix can be computed element-by-element. As an illustration, consider

$$\begin{aligned} \mathbb{E}\left[\boldsymbol{\epsilon}_{t+1} \left(\boldsymbol{\epsilon}_{t+1}' \otimes \boldsymbol{\epsilon}_{t+1}'\right)\right] &= \mathbb{E}\left[\boldsymbol{\epsilon}_{t+1} \left(\boldsymbol{\epsilon}_{t+1} \otimes \boldsymbol{\epsilon}_{t+1}\right)'\right] \\ &= \mathbb{E}\left[\left\{\epsilon_{t+1}\left(\phi_1\right)\right\}_{\phi_1=1}^{n_e} \left(\left\{\epsilon_{t+1}\left(\phi_2\right) \left\{\epsilon_{t+1}\left(\phi_3\right)\right\}_{\phi_3=1}^{n_e}\right\}_{\phi_2=1}^{n_e}\right)'\right] \end{aligned}$$

`Quasi-MATLAB` codes for $\mathbb{E}\left[\boldsymbol{\epsilon}_{t+1} \left(\boldsymbol{\epsilon}_{t+1}' \otimes \boldsymbol{\epsilon}_{t+1}'\right)\right]$ therefore read:

```
E_eps_eps2 = zeros(ne,ne^2)
for phi1 = 1:ne
    index2 = 0
```

```
    for phi2 = 1:ne
        for phi3 = 1:ne
            index2 = index2+1
            if (phi1 = phi2 = phi3)
                E_eps_eps2(phi1,index2)=m^3 (ϵ_{t+1} (phi1))
            end
        end
    end
end
```

Here, m^3 denotes the third moment of $\boldsymbol{\epsilon}_{t+1}$. For further details on computing elements in $\mathbb{V}\left(\boldsymbol{\xi}_{t+1}^{(2)}\right)$, see the technical appendix to the paper.

D Proof of Proposition 2

We consider the system $\mathbf{x}_{t+1} = \mathbf{a}+\mathbf{A}\mathbf{x}_t+\mathbf{v}_{t+1}$ where $\mathbf{A}$ is stable and $\mathbf{v}_{t+1}$ are mean-zero innovations. Thus, the pruned state-space representation of DSGE models belong to this class. For notational convenience, the system is expressed in deviation from its mean as $\mathbf{a} = (\mathbf{I}-\mathbf{A})\,\mathbb{E}[\mathbf{x}]$. Therefore

$$\begin{gathered}\mathbf{x}_{t+1} = (\mathbf{I}-\mathbf{A})\,\mathbb{E}[\mathbf{x}] + \mathbf{A}\mathbf{x}_t + \mathbf{v}_{t+1} \Rightarrow \\ \mathbf{x}_{t+1} - \mathbb{E}[\mathbf{x}] = \mathbf{A}(\mathbf{x}_t - E[\mathbf{x}]) + \mathbf{v}_{t+1} \Rightarrow \\ \mathbf{z}_{t+1} = \mathbf{A}\mathbf{z}_t + \mathbf{v}_{t+1}\end{gathered}$$

We then have

$$\begin{aligned}\mathbf{z}_{t+1}\otimes\mathbf{z}_{t+1} &= (\mathbf{A}\mathbf{z}_t+\mathbf{v}_{t+1})\otimes(\mathbf{A}\mathbf{z}_t+\mathbf{v}_{t+1}) \\ &= \mathbf{A}\mathbf{z}_t\otimes\mathbf{A}\mathbf{z}_t + \mathbf{A}\mathbf{z}_t\otimes\mathbf{v}_{t+1} + \mathbf{v}_{t+1}\otimes\mathbf{A}\mathbf{z}_t + \mathbf{v}_{t+1}\otimes\mathbf{v}_{t+1},\end{aligned}$$

$$\begin{aligned}\mathbf{z}_{t+1}\otimes\mathbf{z}_{t+1}\otimes\mathbf{z}_{t+1} = \;& \mathbf{A}\mathbf{z}_t\otimes\mathbf{A}\mathbf{z}_t\otimes\mathbf{A}\mathbf{z}_t + \mathbf{A}\mathbf{z}_t\otimes\mathbf{A}\mathbf{z}_t\otimes\mathbf{v}_{t+1} \\ &+\mathbf{A}\mathbf{z}_t\otimes\mathbf{v}_{t+1}\otimes\mathbf{A}\mathbf{z}_t + \mathbf{A}\mathbf{z}_t\otimes\mathbf{v}_{t+1}\otimes\mathbf{v}_{t+1} \\ &+\mathbf{v}_{t+1}\otimes\mathbf{A}\mathbf{z}_t\otimes\mathbf{A}\mathbf{z}_t + \mathbf{v}_{t+1}\otimes\mathbf{A}\mathbf{z}_t\otimes\mathbf{v}_{t+1} \\ &+\mathbf{v}_{t+1}\otimes\mathbf{v}_{t+1}\otimes\mathbf{A}\mathbf{z}_t + \mathbf{v}_{t+1}\otimes\mathbf{v}_{t+1}\otimes\mathbf{v}_{t+1}\end{aligned}$$

and

$$\begin{aligned}\mathbf{z}_{t+1}\otimes\mathbf{z}_{t+1}\otimes\mathbf{z}_{t+1}\otimes\mathbf{z}_{t+1} = \;& \mathbf{A}\mathbf{z}_t\otimes\mathbf{A}\mathbf{z}_t\otimes\mathbf{A}\mathbf{z}_t\otimes\mathbf{A}\mathbf{z}_t + \mathbf{A}\mathbf{z}_t\otimes\mathbf{A}\mathbf{z}_t\otimes\mathbf{A}\mathbf{z}_t\otimes\mathbf{v}_{t+1} \\ &+\mathbf{A}\mathbf{z}_t\otimes\mathbf{A}\mathbf{z}_t\otimes\mathbf{v}_{t+1}\otimes\mathbf{A}\mathbf{z}_t + \mathbf{A}\mathbf{z}_t\otimes\mathbf{A}\mathbf{z}_t\otimes\mathbf{v}_{t+1}\otimes\mathbf{v}_{t+1} \\ &+\mathbf{A}\mathbf{z}_t\otimes\mathbf{v}_{t+1}\otimes\mathbf{A}\mathbf{z}_t\otimes\mathbf{A}\mathbf{z}_t + \mathbf{A}\mathbf{z}_t\otimes\mathbf{v}_{t+1}\otimes\mathbf{A}\mathbf{z}_t\otimes\mathbf{v}_{t+1} \\ &+\mathbf{A}\mathbf{z}_t\otimes\mathbf{v}_{t+1}\otimes\mathbf{v}_{t+1}\otimes\mathbf{A}\mathbf{z}_t + \mathbf{A}\mathbf{z}_t\otimes\mathbf{v}_{t+1}\otimes\mathbf{v}_{t+1}\otimes\mathbf{v}_{t+1} \\ &+\mathbf{v}_{t+1}\otimes\mathbf{A}\mathbf{z}_t\otimes\mathbf{A}\mathbf{z}_t\otimes\mathbf{A}\mathbf{z}_t + \mathbf{v}_{t+1}\otimes\mathbf{A}\mathbf{z}_t\otimes\mathbf{A}\mathbf{z}_t\otimes\mathbf{v}_{t+1} \\ &+\mathbf{v}_{t+1}\otimes\mathbf{A}\mathbf{z}_t\otimes\mathbf{v}_{t+1}\otimes\mathbf{A}\mathbf{z}_t + \mathbf{v}_{t+1}\otimes\mathbf{A}\mathbf{z}_t\otimes\mathbf{v}_{t+1}\otimes\mathbf{v}_{t+1} \\ &+\mathbf{v}_{t+1}\otimes\mathbf{v}_{t+1}\otimes\mathbf{A}\mathbf{z}_t\otimes\mathbf{A}\mathbf{z}_t + \mathbf{v}_{t+1}\otimes\mathbf{v}_{t+1}\otimes\mathbf{A}\mathbf{z}_t\otimes\mathbf{v}_{t+1} \\ &+\mathbf{v}_{t+1}\otimes\mathbf{v}_{t+1}\otimes\mathbf{v}_{t+1}\otimes\mathbf{A}\mathbf{z}_t + \mathbf{v}_{t+1}\otimes\mathbf{v}_{t+1}\otimes\mathbf{v}_{t+1}\otimes\mathbf{v}_{t+1}\end{aligned}$$

Thus, to solve for $\mathbb{E}\left[\mathbf{z}_{t+1} \otimes \mathbf{z}_{t+1} \otimes \mathbf{z}_{t+1}\right]$, the innovations need to have a finite third moment. At the second order, $\mathbf{v}_{t+1}$ depends on $\boldsymbol{\epsilon}_{t+1} \otimes \boldsymbol{\epsilon}_{t+1}$, meaning that $\boldsymbol{\epsilon}_{t+1}$ must have a finite sixth moment. Similarly, to solve for $\mathbb{E}\left[\mathbf{z}_{t+1} \otimes \mathbf{z}_{t+1} \otimes \mathbf{z}_{t+1} \otimes \mathbf{z}_{t+1}\right]$, the innovations need to have finite fourth moments. At second order, $\mathbf{v}_{t+1}$ depends on $\boldsymbol{\epsilon}_{t+1} \otimes \boldsymbol{\epsilon}_{t+1}$, meaning that $\boldsymbol{\epsilon}_{t+1}$ must have a finite eight moment.

E Coefficients for the Pruned State-Space System at Third-Order

$$
\mathbf{A}^{(3)} \equiv \begin{bmatrix}
\mathbf{h_x} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\
\mathbf{0} & \mathbf{h_x} & \frac{1}{2}\mathbf{H_{xx}} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\
\mathbf{0} & \mathbf{0} & \mathbf{h_x} \otimes \mathbf{h_x} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\
\frac{3}{6}\mathbf{h}_{\sigma\sigma\mathbf{x}}\sigma^2 & \mathbf{0} & \mathbf{0} & \mathbf{h_x} & \mathbf{H_{xx}} & \frac{1}{6}\mathbf{H_{xxx}} \\
\mathbf{h_x} \otimes \frac{1}{2}\mathbf{h}_{\sigma\sigma}\sigma^2 & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{h_x} \otimes \mathbf{h_x} & \mathbf{h_x} \otimes \frac{1}{2}\mathbf{H_{xx}} \\
\mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{h_x} \otimes \mathbf{h_x} \otimes \mathbf{h_x}
\end{bmatrix}
$$

$$
\mathbf{B}^{(3)} \equiv \left[\begin{array}{cccccccccccc}
\sigma\boldsymbol{\eta} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\
\mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\
\mathbf{0} & \sigma\boldsymbol{\eta}\otimes\sigma\boldsymbol{\eta} & \sigma\boldsymbol{\eta} \otimes \mathbf{h_x} & \mathbf{h_x} \otimes \sigma\boldsymbol{\eta} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\
\mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\
\sigma\boldsymbol{\eta} \otimes \frac{1}{2}\mathbf{h}_{\sigma\sigma}\sigma^2 & \mathbf{0} & \mathbf{0} & \mathbf{0} & \sigma\boldsymbol{\eta} \otimes \mathbf{h_x} & \sigma\boldsymbol{\eta} \otimes \frac{1}{2}\mathbf{H_{xx}} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\
\mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \sigma\boldsymbol{\eta} \otimes \mathbf{h_x} \otimes \mathbf{h_x} & \mathbf{h_x} \otimes \mathbf{h_x} \otimes \sigma\boldsymbol{\eta} & \mathbf{h_x} \otimes \sigma\boldsymbol{\eta} \otimes \mathbf{h_x} & \mathbf{h_x} \otimes \sigma\boldsymbol{\eta}\otimes\sigma\boldsymbol{\eta} & \sigma\boldsymbol{\eta} \otimes \mathbf{h_x}\otimes\sigma\boldsymbol{\eta} & \sigma\boldsymbol{\eta}\otimes\sigma\boldsymbol{\eta} \otimes \mathbf{h_x} & \sigma\boldsymbol{\eta}\otimes\sigma\boldsymbol{\eta}\otimes\sigma\boldsymbol{\eta}
\end{array}\right]
$$

$$
\boldsymbol{\xi}_{t+1}^{(3)} \equiv \begin{bmatrix}
\boldsymbol{\epsilon}_{t+1} \\
\boldsymbol{\epsilon}_{t+1} \otimes \boldsymbol{\epsilon}_{t+1} - vec\left(\mathbf{I}_{n_e}\right) \\
\boldsymbol{\epsilon}_{t+1} \otimes \mathbf{x}_t^f \\
\mathbf{x}_t^f \otimes \boldsymbol{\epsilon}_{t+1} \\
\boldsymbol{\epsilon}_{t+1} \otimes \mathbf{x}_t^s \\
\boldsymbol{\epsilon}_{t+1} \otimes \mathbf{x}_t^f \otimes \mathbf{x}_t^f \\
\mathbf{x}_t^f \otimes \mathbf{x}_t^f \otimes \boldsymbol{\epsilon}_{t+1} \\
\mathbf{x}_t^f \otimes \boldsymbol{\epsilon}_{t+1} \otimes \mathbf{x}_t^f \\
\mathbf{x}_t^f \otimes \boldsymbol{\epsilon}_{t+1} \otimes \boldsymbol{\epsilon}_{t+1} \\
\boldsymbol{\epsilon}_{t+1} \otimes \mathbf{x}_t^f \otimes \boldsymbol{\epsilon}_{t+1} \\
\boldsymbol{\epsilon}_{t+1} \otimes \boldsymbol{\epsilon}_{t+1} \otimes \mathbf{x}_t^f \\
\left(\boldsymbol{\epsilon}_{t+1} \otimes \boldsymbol{\epsilon}_{t+1} \otimes \boldsymbol{\epsilon}_{t+1}\right) - \mathbb{E}\left[\left(\boldsymbol{\epsilon}_{t+1} \otimes \boldsymbol{\epsilon}_{t+1} \otimes \boldsymbol{\epsilon}_{t+1}\right)\right]
\end{bmatrix}
$$

$$\mathbf{c}^{(3)} \equiv \begin{bmatrix} \mathbf{0}_{n_x \times 1} \\ \frac{1}{2}\mathbf{h}_{\sigma\sigma}\sigma^2 \\ (\sigma\boldsymbol{\eta} \otimes \sigma\boldsymbol{\eta})\, vec\,(\mathbf{I}_{n_e}) \\ \frac{1}{6}\mathbf{h}_{\sigma\sigma\sigma}\sigma^3 \\ \mathbf{0}_{n_x^2 \times 1} \\ (\sigma\boldsymbol{\eta}\otimes\sigma\boldsymbol{\eta}\otimes\sigma\boldsymbol{\eta})\,\mathbb{E}\left[(\boldsymbol{\epsilon}_{t+1} \otimes \boldsymbol{\epsilon}_{t+1} \otimes \boldsymbol{\epsilon}_{t+1})\right] \end{bmatrix}$$

$$\mathbf{C}^{(3)} \equiv \begin{bmatrix} \mathbf{g}_{\mathbf{x}} + \frac{3}{6}\mathbf{g}_{\sigma\sigma\mathbf{x}}\sigma^2 & \mathbf{g}_{\mathbf{x}} & \frac{1}{2}\mathbf{G}_{\mathbf{xx}} & \mathbf{g}_{\mathbf{x}} & \mathbf{G}_{\mathbf{xx}} & \frac{1}{6}\mathbf{G}_{\mathbf{xxx}} \end{bmatrix}$$

$$\mathbf{d}^{(3)} \equiv \frac{1}{2}\mathbf{g}_{\sigma\sigma}\sigma^2 + \frac{1}{6}\mathbf{g}_{\sigma\sigma\sigma}\sigma^3.$$

F Proof of Proposition 3

To prove stability:

$$\begin{aligned}
p(\lambda) &= \left|\mathbf{A}^{(3)} - \lambda\mathbf{I}\right| \\
&= \left|\begin{bmatrix}
\mathbf{h}_{\mathbf{x}}-\lambda\mathbf{I} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\
\mathbf{0} & \mathbf{h}_{\mathbf{x}}-\lambda\mathbf{I} & \frac{1}{2}\mathbf{H}_{\mathbf{xx}} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\
\mathbf{0} & \mathbf{0} & \mathbf{h}_{\mathbf{x}}\otimes\mathbf{h}_{\mathbf{x}}-\lambda\mathbf{I} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\
\frac{3}{6}\mathbf{h}_{\sigma\sigma\mathbf{x}}\sigma^2 & \mathbf{0} & \mathbf{0} & \mathbf{h}_{\mathbf{x}}-\lambda\mathbf{I} & \mathbf{H}_{\mathbf{xx}} & \frac{1}{6}\mathbf{H}_{\mathbf{xxx}} \\
\mathbf{h}_{\mathbf{x}}\otimes\frac{1}{2}\mathbf{h}_{\sigma\sigma}\sigma^2 & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{h}_{\mathbf{x}}\otimes\mathbf{h}_{\mathbf{x}}-\lambda\mathbf{I} & \mathbf{h}_{\mathbf{x}}\otimes\frac{1}{2}\mathbf{H}_{\mathbf{xx}} \\
\mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{h}_{\mathbf{x}}\otimes\mathbf{h}_{\mathbf{x}}\otimes\mathbf{h}_{\mathbf{x}}-\lambda\mathbf{I}
\end{bmatrix}\right| \\
&= \left|\begin{bmatrix} \mathbf{B}_{11} & \mathbf{B}_{12} \\ \mathbf{B}_{21} & \mathbf{B}_{22} \end{bmatrix}\right|
\end{aligned}$$

where

$$\mathbf{B}_{11} \equiv \begin{bmatrix} \mathbf{h}_{\mathbf{x}}-\lambda\mathbf{I} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{h}_{\mathbf{x}}-\lambda\mathbf{I} & \frac{1}{2}\mathbf{H}_{\mathbf{xx}} \\ \mathbf{0} & \mathbf{0} & \mathbf{h}_{\mathbf{x}}\otimes\mathbf{h}_{\mathbf{x}}-\lambda\mathbf{I} \end{bmatrix} \quad \mathbf{B}_{12} \equiv \begin{bmatrix} \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} \end{bmatrix}$$

$$\mathbf{B}_{21} \equiv \begin{bmatrix} \frac{3}{6}\mathbf{h}_{\sigma\sigma\mathbf{x}}\sigma^2 & \mathbf{0} & \mathbf{0} \\ \mathbf{h}_{\mathbf{x}}\otimes\frac{1}{2}\mathbf{h}_{\sigma\sigma}\sigma^2 & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} \end{bmatrix} \quad \mathbf{B}_{22} \equiv \begin{bmatrix} \mathbf{h}_{\mathbf{x}}-\lambda\mathbf{I} & \mathbf{H}_{\mathbf{xx}} & \frac{1}{6}\mathbf{H}_{\mathbf{xxx}} \\ \mathbf{0} & \mathbf{h}_{\mathbf{x}}\otimes\mathbf{h}_{\mathbf{x}}-\lambda\mathbf{I} & \mathbf{h}_{\mathbf{x}}\otimes\frac{1}{2}\mathbf{H}_{\mathbf{xx}} \\ \mathbf{0} & \mathbf{0} & \mathbf{h}_{\mathbf{x}}\otimes\mathbf{h}_{\mathbf{x}}\otimes\mathbf{h}_{\mathbf{x}}-\lambda\mathbf{I} \end{bmatrix}$$

$$= |\mathbf{B}_{11}|\,|\mathbf{B}_{22}|$$

$$= |\mathbf{h}_{\mathbf{x}} - \lambda\mathbf{I}|\,|\mathbf{h}_{\mathbf{x}} - \lambda\mathbf{I}|\,|\mathbf{h}_{\mathbf{x}}\otimes\mathbf{h}_{\mathbf{x}} - \lambda\mathbf{I}|\,|\mathbf{B}_{22}|$$

(using the result from the proof of proposition 1)

$$= |\mathbf{h}_{\mathbf{x}} - \lambda\mathbf{I}|\,|\mathbf{h}_{\mathbf{x}} - \lambda\mathbf{I}|\,|\mathbf{h}_{\mathbf{x}}\otimes\mathbf{h}_{\mathbf{x}} - \lambda\mathbf{I}|\,|\mathbf{h}_{\mathbf{x}} - \lambda\mathbf{I}|\,|\mathbf{h}_{\mathbf{x}}\otimes\mathbf{h}_{\mathbf{x}}-\lambda\mathbf{I}|\,|\mathbf{h}_{\mathbf{x}}\otimes\mathbf{h}_{\mathbf{x}}\otimes\mathbf{h}_{\mathbf{x}}-\lambda\mathbf{I}|$$

(using the rule on block determinants repeatedly on $\mathbf{B}_{22}$).

The eigenvalue λ solves $p(\lambda) = 0$, which implies

$$|\mathbf{h}_{\mathbf{x}}-\lambda\mathbf{I}| = 0 \text{ or } |\mathbf{h}_{\mathbf{x}}\otimes\mathbf{h}_{\mathbf{x}}-\lambda\mathbf{I}| = 0 \text{ or } |(\mathbf{h}_{\mathbf{x}}\otimes\mathbf{h}_{\mathbf{x}}\otimes\mathbf{h}_{\mathbf{x}})-\lambda\mathbf{I}| = 0$$

The absolute value of all eigenvalues to the first problem is strictly less than one by assumption. That is $|\lambda_i| < 1$, $i = 1, 2, ..., n_x$. This is also the case for the second problem, because the eigenvalues to $\mathbf{h}_{\mathbf{x}}\otimes\mathbf{h}_{\mathbf{x}}$ are $\lambda_i\lambda_j$ for $i = 1, 2, ..., n_x$ and $j = 1, 2, ..., n_x$. The same argument ensures that the absolute values of all eigenvalues to the third problem are also less than one. This shows that all eigenvalues of $\mathbf{A}^{(3)}$ have modulus less than one.

Given that the system is stable, the mean value is given by $\mathbb{E}\left[\mathbf{z}_t^{(3)}\right] = \left(\mathbf{I}_{3n_x+2n_x^2+n_x^3} - \mathbf{A}^{(3)}\right)^{-1}\mathbf{c}^{(3)}$ because $\mathbb{E}\left[\boldsymbol{\xi}_{t+1}^{(3)}\right] = \mathbf{0}$. For the variance, we have

$$
\begin{aligned}
\mathbb{V}\left[\mathbf{z}_{t+1}^{(3)}\right] &= \mathbf{A}^{(3)}\mathbb{V}\left[\mathbf{z}_t^{(3)}\right]\left(\mathbf{A}^{(3)}\right)' + \mathbf{B}^{(3)}\mathbb{V}\left[\boldsymbol{\xi}_{t+1}^{(3)}\right]\left(\mathbf{B}^{(3)}\right)' \\
&\quad + \mathbf{A}^{(3)}Cov\left[\mathbf{z}_t^{(3)}, \boldsymbol{\xi}_{t+1}^{(3)}\right]\left(\mathbf{B}^{(3)}\right)' + \mathbf{B}^{(3)}Cov\left[\boldsymbol{\xi}_{t+1}^{(3)}, \mathbf{z}_t^{(3)}\right]\left(\mathbf{A}^{(3)}\right)'
\end{aligned}
$$

Contrary to a second-order approximation, $Cov\left[\boldsymbol{\xi}_{t+1}^{(3)}, \mathbf{z}_t^{(3)}\right] \neq 0$. This is seen as follows:

$$
\begin{aligned}
\mathbb{E}\left[\mathbf{z}_t^{(3)}\left(\boldsymbol{\xi}_{t+1}^{(3)}\right)'\right] = \mathbb{E}\Bigg[& \begin{bmatrix} \mathbf{x}_t^f \\ \mathbf{x}_t^s \\ \mathbf{x}_t^f \otimes \mathbf{x}_t^f \\ \mathbf{x}_t^{rd} \\ \mathbf{x}_t^f \otimes \mathbf{x}_t^s \\ \mathbf{x}_t^f \otimes \mathbf{x}_t^f \otimes \mathbf{x}_t^f \end{bmatrix} \\
& \times \Big[\ \boldsymbol{\epsilon}_{t+1}' \quad \left(\boldsymbol{\epsilon}_{t+1} \otimes \boldsymbol{\epsilon}_{t+1} - vec\left(\mathbf{I}_{n_e}\right)\right)' \quad \left(\boldsymbol{\epsilon}_{t+1} \otimes \mathbf{x}_t^f\right)' \quad \left(\mathbf{x}_t^f \otimes \boldsymbol{\epsilon}_{t+1}\right)' \quad \left(\boldsymbol{\epsilon}_{t+1} \otimes \mathbf{x}_t^s\right)' \quad \left(\boldsymbol{\epsilon}_{t+1} \otimes \mathbf{x}_t^f \otimes \mathbf{x}_t^f\right)' \\
& \left(\mathbf{x}_t^f \otimes \mathbf{x}_t^f \otimes \boldsymbol{\epsilon}_{t+1}\right)' \quad \left(\mathbf{x}_t^f \otimes \boldsymbol{\epsilon}_{t+1} \otimes \mathbf{x}_t^f\right)' \quad \left(\mathbf{x}_t^f \otimes \boldsymbol{\epsilon}_{t+1} \otimes \boldsymbol{\epsilon}_{t+1}\right)' \quad \left(\boldsymbol{\epsilon}_{t+1} \otimes \mathbf{x}_t^f \otimes \boldsymbol{\epsilon}_{t+1}\right)' \\
& \left(\boldsymbol{\epsilon}_{t+1} \otimes \boldsymbol{\epsilon}_{t+1} \otimes \mathbf{x}_t^f\right)' \quad \left(\left(\boldsymbol{\epsilon}_{t+1} \otimes \boldsymbol{\epsilon}_{t+1} \otimes \boldsymbol{\epsilon}_{t+1}\right) - E\left[\left(\boldsymbol{\epsilon}_{t+1} \otimes \boldsymbol{\epsilon}_{t+1} \otimes \boldsymbol{\epsilon}_{t+1}\right)\right]\right)'\ \Big]\Bigg]
\end{aligned}
$$

$$
= \begin{bmatrix}
0_{n_x \times n_e} & 0_{n_x \times n_e^2} & 0_{n_x \times n_e n_x} & 0_{n_x \times n_x n_e} & 0_{n_x \times n_x n_e} & 0_{n_x \times n_e n_x^2} & 0_{n_x \times n_x^2 n_e} & 0_{n_x \times n_x^2 n_e} & R_{1,1} & R_{1,2} & R_{1,3} & 0_{n_x \times n_e^3} \\
0_{n_x \times n_e} & 0_{n_x \times n_e^2} & 0_{n_x \times n_e n_x} & 0_{n_x \times n_x n_e} & 0_{n_x \times n_x n_e} & 0_{n_x \times n_e n_x^2} & 0_{n_x \times n_x^2 n_e} & 0_{n_x \times n_x^2 n_e} & R_{2,1} & R_{2,2} & R_{2,3} & 0_{n_x \times n_e^3} \\
0_{n_x^2 \times n_e} & 0_{n_x^2 \times n_e^2} & 0_{n_x^2 \times n_e n_x} & 0_{n_x^2 \times n_x n_e} & 0_{n_x^2 \times n_x n_e} & 0_{n_x^2 \times n_e n_x^2} & 0_{n_x^2 \times n_x^2 n_e} & 0_{n_x^2 \times n_x^2 n_e} & R_{3,1} & R_{3,2} & R_{3,3} & 0_{n_x^2 \times n_e^3} \\
0_{n_x \times n_e} & 0_{n_x \times n_e^2} & 0_{n_x \times n_e n_x} & 0_{n_x \times n_x n_e} & 0_{n_x \times n_x n_e} & 0_{n_x \times n_e n_x^2} & 0_{n_x \times n_x^2 n_e} & 0_{n_x \times n_x^2 n_e} & R_{4,1} & R_{4,2} & R_{4,3} & 0_{n_x \times n_e^3} \\
0_{n_x^2 \times n_e} & 0_{n_x^2 \times n_e^2} & 0_{n_x^2 \times n_e n_x} & 0_{n_x^2 \times n_x n_e} & 0_{n_x^2 \times n_x n_e} & 0_{n_x^2 \times n_e n_x^2} & 0_{n_x^2 \times n_x^2 n_e} & 0_{n_x^2 \times n_x^2 n_e} & R_{5,1} & R_{5,2} & R_{5,3} & 0_{n_x^2 \times n_e^3} \\
0_{n_x^3 \times n_e} & 0_{n_x^3 \times n_e^2} & 0_{n_x^3 \times n_e n_x} & 0_{n_x^3 \times n_x n_e} & 0_{n_x^3 \times n_x n_e} & 0_{n_x^3 \times n_e n_x^2} & 0_{n_x^3 \times n_x^2 n_e} & 0_{n_x^3 \times n_x^2 n_e} & R_{6,1} & R_{6,2} & R_{6,3} & 0_{n_x^3 \times n_e^3}
\end{bmatrix}
$$

$$
= \begin{bmatrix} \mathbf{0} & \mathbf{R} & \mathbf{0} \end{bmatrix}
$$

The $\mathbf{R}$ matrix can easily be computed element by element. To compute $\mathbb{V}\left[\boldsymbol{\xi}_{t+1}^{(3)}\right]$, we consider

$$\mathbb{E}\left[\boldsymbol{\xi}_{t+1}^{(3)}\left(\boldsymbol{\xi}_{t+1}^{(3)}\right)'\right]=\mathbb{E}\left[\left[\begin{array}{c}\boldsymbol{\epsilon}_{t+1}\\ \boldsymbol{\epsilon}_{t+1}\otimes\boldsymbol{\epsilon}_{t+1}-vec\left(\mathbf{I}_{n_e}\right)\\ \boldsymbol{\epsilon}_{t+1}\otimes\mathbf{x}_t^f\\ \mathbf{x}_t^f\otimes\boldsymbol{\epsilon}_{t+1}\\ \boldsymbol{\epsilon}_{t+1}\otimes\mathbf{x}_t^s\\ \boldsymbol{\epsilon}_{t+1}\otimes\mathbf{x}_t^f\otimes\mathbf{x}_t^f\\ \mathbf{x}_t^f\otimes\mathbf{x}_t^f\otimes\boldsymbol{\epsilon}_{t+1}\\ \mathbf{x}_t^f\otimes\boldsymbol{\epsilon}_{t+1}\otimes\mathbf{x}_t^f\\ \mathbf{x}_t^f\otimes\boldsymbol{\epsilon}_{t+1}\otimes\boldsymbol{\epsilon}_{t+1}\\ \boldsymbol{\epsilon}_{t+1}\otimes\mathbf{x}_t^f\otimes\boldsymbol{\epsilon}_{t+1}\\ \boldsymbol{\epsilon}_{t+1}\otimes\boldsymbol{\epsilon}_{t+1}\otimes\mathbf{x}_t^f\\ \left(\boldsymbol{\epsilon}_{t+1}\otimes\boldsymbol{\epsilon}_{t+1}\otimes\boldsymbol{\epsilon}_{t+1}\right)-\mathbb{E}\left[\left(\boldsymbol{\epsilon}_{t+1}\otimes\boldsymbol{\epsilon}_{t+1}\otimes\boldsymbol{\epsilon}_{t+1}\right)\right]\end{array}\right]\right.$$

$$\times\left[\begin{array}{cccccc}\boldsymbol{\epsilon}_{t+1}' & \left(\boldsymbol{\epsilon}_{t+1}\otimes\boldsymbol{\epsilon}_{t+1}-vec\left(\mathbf{I}_{n_e}\right)\right)' & \left(\boldsymbol{\epsilon}_{t+1}\otimes\mathbf{x}_t^f\right)' & \left(\mathbf{x}_t^f\otimes\boldsymbol{\epsilon}_{t+1}\right)' & \left(\boldsymbol{\epsilon}_{t+1}\otimes\mathbf{x}_t^s\right)' & \left(\boldsymbol{\epsilon}_{t+1}\otimes\mathbf{x}_t^f\otimes\mathbf{x}_t^f\right)'\\ \left(\mathbf{x}_t^f\otimes\mathbf{x}_t^f\otimes\boldsymbol{\epsilon}_{t+1}\right)' & \left(\mathbf{x}_t^f\otimes\boldsymbol{\epsilon}_{t+1}\otimes\mathbf{x}_t^f\right)' & \left(\mathbf{x}_t^f\otimes\boldsymbol{\epsilon}_{t+1}\otimes\boldsymbol{\epsilon}_{t+1}\right)' & \left(\boldsymbol{\epsilon}_{t+1}\otimes\mathbf{x}_t^f\otimes\boldsymbol{\epsilon}_{t+1}\right)' & & \\ \left(\boldsymbol{\epsilon}_{t+1}\otimes\boldsymbol{\epsilon}_{t+1}\otimes\mathbf{x}_t^f\right)' & \left(\left(\boldsymbol{\epsilon}_{t+1}\otimes\boldsymbol{\epsilon}_{t+1}\otimes\boldsymbol{\epsilon}_{t+1}\right)-\mathbb{E}\left[\left(\boldsymbol{\epsilon}_{t+1}\otimes\boldsymbol{\epsilon}_{t+1}\otimes\boldsymbol{\epsilon}_{t+1}\right)\right]\right)' & & & & \end{array}\right]$$

Note that $\mathbb{V}\left[\boldsymbol{\xi}_{t+1}^{(3)}\right]$ contains $(\boldsymbol{\epsilon}_{t+1}\otimes\boldsymbol{\epsilon}_{t+1}\otimes\boldsymbol{\epsilon}_{t+1})$ squared, meaning that $\boldsymbol{\epsilon}_{t+1}$ must have a finite sixth moment for $\mathbb{V}\left[\boldsymbol{\xi}_{t+1}^{(3)}\right]$ to be finite. Again, all elements in $\mathbb{V}\left[\boldsymbol{\xi}_{t+1}^{(3)}\right]$ can be computed element by element. For further details, we refer to the paper's technical appendix, which also discusses how $V\left[\boldsymbol{\xi}_{t+1}^{(3)}\right]$ can be computed in a more memory-efficient manner.

For the auto-covariance, we have

$$\begin{aligned}Cov\left(\mathbf{z}_{t+1}^{(3)},\mathbf{z}_t^{(3)}\right)&=Cov\left(\mathbf{c}^{(3)}+\mathbf{A}^{(3)}\mathbf{z}_t^{(3)}+\mathbf{B}^{(3)}\boldsymbol{\xi}_{t+1}^{(3)},\mathbf{z}_t^{(3)}\right)\\&=\mathbf{A}^{(3)}Cov\left(\mathbf{z}_t^{(3)},\mathbf{z}_t^{(3)}\right)+\mathbf{B}^{(3)}Cov\left(\boldsymbol{\xi}_{t+1}^{(3)},\mathbf{z}_t^{(3)}\right)\end{aligned}$$

$$\begin{aligned}Cov\left(\mathbf{z}_{t+2}^{(3)},\mathbf{z}_t^{(3)}\right)&=Cov\left(\mathbf{c}^{(3)}+\mathbf{A}^{(3)}\mathbf{z}_{t+1}^{(3)}+\mathbf{B}^{(3)}\boldsymbol{\xi}_{t+2}^{(3)},\mathbf{z}_t^{(3)}\right)\\&=Cov\left(\mathbf{c}^{(3)}+\mathbf{A}^{(3)}\left(\mathbf{c}^{(3)}+\mathbf{A}^{(3)}\mathbf{z}_t^{(3)}+\mathbf{B}^{(3)}\boldsymbol{\xi}_{t+1}^{(3)}\right)+\mathbf{B}^{(3)}\boldsymbol{\xi}_{t+2}^{(3)},\mathbf{z}_t^{(3)}\right)\\&=Cov\left(\mathbf{c}^{(3)}+\mathbf{A}^{(3)}\mathbf{c}^{(3)}+\left(\mathbf{A}^{(3)}\right)^2\mathbf{z}_t^{(3)}+\mathbf{A}^{(3)}\mathbf{B}^{(3)}\boldsymbol{\xi}_{t+1}^{(3)}+\mathbf{B}^{(3)}\boldsymbol{\xi}_{t+2}^{(3)},\mathbf{z}_t^{(3)}\right)\\&=Cov\left(\left(\mathbf{A}^{(3)}\right)^2\mathbf{z}_t^{(3)},\mathbf{z}_t^{(3)}\right)+Cov\left(\mathbf{A}^{(3)}\mathbf{B}^{(3)}\boldsymbol{\xi}_{t+1}^{(3)},\mathbf{z}_t^{(3)}\right)+Cov\left(\mathbf{B}^{(3)}\boldsymbol{\xi}_{t+2}^{(3)},\mathbf{z}_t^{(3)}\right)\\&=\left(\mathbf{A}^{(3)}\right)^2Cov\left(\mathbf{z}_t^{(3)},\mathbf{z}_t^{(3)}\right)+\mathbf{A}^{(3)}\mathbf{B}^{(3)}Cov\left(\boldsymbol{\xi}_{t+1}^{(3)},\mathbf{z}_t^{(3)}\right)+\mathbf{B}^{(3)}Cov\left(\boldsymbol{\xi}_{t+2}^{(3)},\mathbf{z}_t^{(3)}\right)\end{aligned}$$

So, for $s=1,2,3...$

$$Cov\left(\mathbf{z}_{t+s}^{(3)},\mathbf{z}_t^{(3)}\right)=\left(\mathbf{A}^{(3)}\right)^s\mathbb{V}\left[\mathbf{z}_t^{(3)}\right]+\textstyle\sum_{j=0}^{s-1}\left(\mathbf{A}^{(3)}\right)^{s-1-j}\mathbf{B}^{(3)}Cov\left(\boldsymbol{\xi}_{t+1+j}^{(3)},\mathbf{z}_t^{(3)}\right)$$

and we therefore only need to compute $Cov\left(\boldsymbol{\xi}_{t+1+j}^{(3)},\mathbf{z}_t^{(3)}\right)$:

$$\mathbb{E}\left[\mathbf{z}_t^{(3)}\left(\boldsymbol{\xi}_{t+1+j}^{(3)}\right)'\right]=\mathbb{E}\left[\left[\begin{array}{c}\mathbf{x}_t^f\\ \mathbf{x}_t^s\\ \mathbf{x}_t^f\otimes\mathbf{x}_t^f\\ \mathbf{x}_t^{rd}\\ \mathbf{x}_t^f\otimes\mathbf{x}_t^s\\ \mathbf{x}_t^f\otimes\mathbf{x}_t^f\otimes\mathbf{x}_t^f\end{array}\right]\right.$$

$$
\times \Big[\begin{array}{l} \boldsymbol{\epsilon}'_{t+1+j} \quad (\boldsymbol{\epsilon}_{t+1+j} \otimes \boldsymbol{\epsilon}_{t+1+j} - vec\,(\mathbf{I}_{n_e}))' \quad \left(\boldsymbol{\epsilon}_{t+1+j} \otimes \mathbf{x}^f_{t+j}\right)' \\ \left(\mathbf{x}^f_{t+j} \otimes \boldsymbol{\epsilon}_{t+1+j}\right)' \quad \left(\boldsymbol{\epsilon}_{t+1+j} \otimes \mathbf{x}^s_{t+j}\right)' \quad \left(\boldsymbol{\epsilon}_{t+1+j} \otimes \mathbf{x}^f_{t+j} \otimes \mathbf{x}^f_{t+j}\right)' \\ \left(\mathbf{x}^f_{t+j} \otimes \mathbf{x}^f_{t+j} \otimes \boldsymbol{\epsilon}_{t+1+j}\right)' \quad \left(\mathbf{x}^f_{t+j} \otimes \boldsymbol{\epsilon}_{t+1+j} \otimes \mathbf{x}^f_{t+j}\right)' \\ \left(\mathbf{x}^f_{t+j} \otimes \boldsymbol{\epsilon}_{t+1+j} \otimes \boldsymbol{\epsilon}_{t+1+j}\right)' \quad \left(\boldsymbol{\epsilon}_{t+1+j} \otimes \mathbf{x}^f_{t+j} \otimes \boldsymbol{\epsilon}_{t+1+j}\right)' \\ \left(\boldsymbol{\epsilon}_{t+1+j} \otimes \boldsymbol{\epsilon}_{t+1+j} \otimes \mathbf{x}^f_{t+j}\right)' \quad \left((\boldsymbol{\epsilon}_{t+1+j} \otimes \boldsymbol{\epsilon}_{t+1+j} \otimes \boldsymbol{\epsilon}_{t+1+j}) - E\left[(\boldsymbol{\epsilon}_{t+1} \otimes \boldsymbol{\epsilon}_{t+1} \otimes \boldsymbol{\epsilon}_{t+1})\right]\right)' \end{array} \Big]
$$

$$
= \begin{bmatrix}
0_{n_x\times n_e} & 0_{n_x\times n_e^2} & 0_{n_x\times n_e n_x} & 0_{n_x\times n_x n_e} & 0_{n_x\times n_x n_e} & 0_{n_x\times n_e n_x^2} & 0_{n_x\times n_x^2 n_e} & 0_{n_x\times n_x^2 n_e} \\
0_{n_x\times n_e} & 0_{n_x\times n_e^2} & 0_{n_x\times n_e n_x} & 0_{n_x\times n_x n_e} & 0_{n_x\times n_x n_e} & 0_{n_x\times n_e n_x^2} & 0_{n_x\times n_x^2 n_e} & 0_{n_x\times n_x^2 n_e} \\
0_{n_x^2\times n_e} & 0_{n_x^2\times n_e^2} & 0_{n_x^2\times n_e n_x} & 0_{n_x^2\times n_x n_e} & 0_{n_x^2\times n_x n_e} & 0_{n_x^2\times n_e n_x^2} & 0_{n_x^2\times n_x^2 n_e} & 0_{n_x^2\times n_x^2 n_e} \\
0_{n_x\times n_e} & 0_{n_x\times n_e^2} & 0_{n_x\times n_e n_x} & 0_{n_x\times n_x n_e} & 0_{n_x\times n_x n_e} & 0_{n_x\times n_e n_x^2} & 0_{n_x\times n_x^2 n_e} & 0_{n_x\times n_x^2 n_e} \\
0_{n_x^2\times n_e} & 0_{n_x^2\times n_e^2} & 0_{n_x^2\times n_e n_x} & 0_{n_x^2\times n_x n_e} & 0_{n_x^2\times n_x n_e} & 0_{n_x^2\times n_e n_x^2} & 0_{n_x^2\times n_x^2 n_e} & 0_{n_x^2\times n_x^2 n_e} \\
0_{n_x^3\times n_e} & 0_{n_x^3\times n_e^2} & 0_{n_x^3\times n_e n_x} & 0_{n_x^3\times n_x n_e} & 0_{n_x^3\times n_x n_e} & 0_{n_x^3\times n_e n_x^2} & 0_{n_x^3\times n_x^2 n_e} & 0_{n_x^3\times n_x^2 n_e}
\end{bmatrix.
$$

$$
\begin{bmatrix}
R^j_{1,1} & R^j_{1,2} & R^j_{1,3} & 0_{n_x\times n_e^3} \\
R^j_{2,1} & R^j_{2,2} & R^j_{2,3} & 0_{n_x\times n_e^3} \\
R^j_{3,1} & R^j_{3,2} & R^j_{3,3} & 0_{n_x^2\times n_e^3} \\
R^j_{4,1} & R^j_{4,2} & R^j_{4,3} & 0_{n_x\times n_e^3} \\
R^j_{5,1} & R^j_{5,2} & R^j_{5,3} & 0_{n_x^2\times n_e^3} \\
R^j_{6,1} & R^j_{6,2} & R^j_{6,3} & 0_{n_x^3\times n_e^3}
\end{bmatrix}
$$

$$
= \begin{bmatrix} \mathbf{0} & \mathbf{R}^j & \mathbf{0} \end{bmatrix}
$$

The matrix $\mathbf{R}^j$ can then be computed element by element. For further details, see the paper's technical appendix.

G Proof of Proposition 4

The proof proceeds as for proposition 2. At third order, the only difference is that $\mathbf{v}_{t+1}$ also depends on $\boldsymbol{\epsilon}_{t+1} \otimes \boldsymbol{\epsilon}_{t+1} \otimes \boldsymbol{\epsilon}_{t+1}$. Hence, unconditional third moments exists if $\boldsymbol{\epsilon}_{t+1}$ has a finite ninth moment, and the unconditional fourth moment exists if $\boldsymbol{\epsilon}_{t+1}$ has a finite twelfth moment.

H IRFs: Second-Order

We first note that

$$
\mathbf{x}^f_{t+l} \otimes \mathbf{x}^f_{t+l} = \left(\mathbf{h}^l_{\mathbf{x}}\mathbf{x}^f_t + \sum_{j=1}^{l} \mathbf{h}^{l-j}_{\mathbf{x}} \sigma\boldsymbol{\eta}\boldsymbol{\epsilon}_{t+j}\right) \otimes \left(\mathbf{h}^l_{\mathbf{x}}\mathbf{x}^f_t + \sum_{j=1}^{l} \mathbf{h}^{l-j}_{\mathbf{x}} \sigma\boldsymbol{\eta}\boldsymbol{\epsilon}_{t+j}\right)
$$

$$
= \mathbf{h}^l_{\mathbf{x}}\mathbf{x}^f_t \otimes \mathbf{h}^l_{\mathbf{x}}\mathbf{x}^f_t + \mathbf{h}^l_{\mathbf{x}}\mathbf{x}^f_t \otimes \sum_{j=1}^{l} \mathbf{h}^{l-j}_{\mathbf{x}} \sigma\boldsymbol{\eta}\boldsymbol{\epsilon}_{t+j}
$$

$$
+ \sum_{j=1}^{l} \mathbf{h}^{l-j}_{\mathbf{x}} \sigma\boldsymbol{\eta}\boldsymbol{\epsilon}_{t+j} \otimes \mathbf{h}^l_{\mathbf{x}}\mathbf{x}^f_t + \sum_{j=1}^{l} \mathbf{h}^{l-j}_{\mathbf{x}} \sigma\boldsymbol{\eta}\boldsymbol{\epsilon}_{t+j} \otimes \sum_{j=1}^{l} \mathbf{h}^{l-j}_{\mathbf{x}} \sigma\boldsymbol{\eta}\boldsymbol{\epsilon}_{t+j}
$$

Next, let

$$\tilde{\mathbf{x}}^f_{t+l} \otimes \tilde{\mathbf{x}}^f_{t+l} = \mathbf{h}^l_{\mathbf{x}}\mathbf{x}^f_t \otimes \mathbf{h}^l_{\mathbf{x}}\mathbf{x}^f_t + \mathbf{h}^l_{\mathbf{x}}\mathbf{x}^f_t \otimes \sum_{j=1}^{l} \mathbf{h}^{l-j}_{\mathbf{x}} \sigma\boldsymbol{\eta}\boldsymbol{\delta}_{t+j}$$
$$+ \sum_{j=1}^{l} \mathbf{h}^{l-j}_{\mathbf{x}} \sigma\boldsymbol{\eta}\boldsymbol{\delta}_{t+j} \otimes \mathbf{h}^l_{\mathbf{x}}\mathbf{x}^f_t + \sum_{j=1}^{l} \mathbf{h}^{l-j}_{\mathbf{x}} \sigma\boldsymbol{\eta}\boldsymbol{\delta}_{t+j} \otimes \sum_{j=1}^{l} \mathbf{h}^{l-j}_{\mathbf{x}} \sigma\boldsymbol{\eta}\boldsymbol{\delta}_{t+j}$$

where we define $\boldsymbol{\delta}_{t+j}$ such that $\boldsymbol{\delta}_{t+j} = \boldsymbol{\nu}$ for $j = 1$ and $\boldsymbol{\delta}_{t+j} = \boldsymbol{\epsilon}_{t+j}$ for for $j \neq 1$. This means that:

$$GIRF_{\mathbf{x}^f \otimes \mathbf{x}^f}\left(l, \boldsymbol{\nu}, \mathbf{x}^f_t\right) = \mathbb{E}\left[\mathbf{x}^f_{t+l} \otimes \mathbf{x}^f_{t+l} | \mathbf{x}^f_t, \boldsymbol{\epsilon}_{t+1} = 1\right] - \mathbb{E}\left[\mathbf{x}^f_{t+l} \otimes \mathbf{x}^f_{t+l} | \mathbf{x}^f_t\right]$$
$$= \mathbb{E}\left[\tilde{\mathbf{x}}^f_{t+l} \otimes \tilde{\mathbf{x}}^f_{t+l} | \mathbf{x}^f_t\right] - \mathbb{E}\left[\mathbf{x}^f_{t+l} \otimes \mathbf{x}^f_{t+l} | \mathbf{x}^f_t\right]$$
$$= \mathbb{E}[\mathbf{h}^l_{\mathbf{x}}\mathbf{x}^f_t \otimes \mathbf{h}^{l-1}_{\mathbf{x}} \sigma\boldsymbol{\eta}\boldsymbol{\delta}_{t+1} + \mathbf{h}^{l-1}_{\mathbf{x}} \sigma\boldsymbol{\eta}\boldsymbol{\delta}_{t+1} \otimes \mathbf{h}^l_{\mathbf{x}}\mathbf{x}^f_t$$
$$+ \sum_{j=1}^{l} \mathbf{h}^{l-j}_{\mathbf{x}} \sigma\boldsymbol{\eta}\boldsymbol{\delta}_{t+j} \otimes \sum_{j=1}^{l} \mathbf{h}^{l-j}_{\mathbf{x}} \sigma\boldsymbol{\eta}\boldsymbol{\delta}_{t+j}$$
$$- \sum_{j=1}^{l} \mathbf{h}^{l-j}_{\mathbf{x}} \sigma\boldsymbol{\eta}\boldsymbol{\epsilon}_{t+j} \otimes \sum_{j=1}^{l} \mathbf{h}^{l-j}_{\mathbf{x}} \sigma\boldsymbol{\eta}\boldsymbol{\epsilon}_{t+j} | \mathbf{x}^f_t]$$
$$= \mathbb{E}[\mathbf{h}^l_{\mathbf{x}}\mathbf{x}^f_t \otimes \mathbf{h}^{l-1}_{\mathbf{x}} \sigma\boldsymbol{\eta}\boldsymbol{\delta}_{t+1} + \mathbf{h}^{l-1}_{\mathbf{x}} \sigma\boldsymbol{\eta}\boldsymbol{\delta}_{t+1} \otimes \mathbf{h}^l_{\mathbf{x}}\mathbf{x}^f_t$$
$$+ \left(\mathbf{h}^{l-1}_{\mathbf{x}} \sigma\boldsymbol{\eta}\boldsymbol{\nu} + \sum_{j=2}^{l} \mathbf{h}^{l-j}_{\mathbf{x}} \sigma\boldsymbol{\eta}\boldsymbol{\epsilon}_{t+j}\right) \otimes \left(\mathbf{h}^{l-1}_{\mathbf{x}} \sigma\boldsymbol{\eta}\boldsymbol{\nu} + \sum_{j=2}^{l} \mathbf{h}^{l-j}_{\mathbf{x}} \sigma\boldsymbol{\eta}\boldsymbol{\epsilon}_{t+j}\right)$$
$$- \left(\mathbf{h}^{l-1}_{\mathbf{x}} \sigma\boldsymbol{\eta}\boldsymbol{\epsilon}_{t+1} + \sum_{j=2}^{l} \mathbf{h}^{l-j}_{\mathbf{x}} \sigma\boldsymbol{\eta}\boldsymbol{\epsilon}_{t+j}\right) \otimes \left(\mathbf{h}^{l-1}_{\mathbf{x}} \sigma\boldsymbol{\eta}\boldsymbol{\epsilon}_{t+1} + \sum_{j=2}^{l} \mathbf{h}^{l-j}_{\mathbf{x}} \sigma\boldsymbol{\eta}\boldsymbol{\epsilon}_{t+j}\right) | \mathbf{x}^f_t]$$
$$= \mathbb{E}[\mathbf{h}^l_{\mathbf{x}}\mathbf{x}^f_t \otimes \mathbf{h}^{l-1}_{\mathbf{x}} \sigma\boldsymbol{\eta}\boldsymbol{\nu} + \mathbf{h}^{l-1}_{\mathbf{x}} \sigma\boldsymbol{\eta}\boldsymbol{\nu} \otimes \mathbf{h}^l_{\mathbf{x}}\mathbf{x}^f_t + \mathbf{h}^{l-1}_{\mathbf{x}} \sigma\boldsymbol{\eta}\boldsymbol{\nu} \otimes \mathbf{h}^{l-1}_{\mathbf{x}} \sigma\boldsymbol{\eta}\boldsymbol{\nu} - \mathbf{h}^{l-1}_{\mathbf{x}} \sigma\boldsymbol{\eta}\boldsymbol{\epsilon}_{t+1} \otimes \mathbf{h}^{l-1}_{\mathbf{x}} \sigma\boldsymbol{\eta}\boldsymbol{\epsilon}_{t+1} | \mathbf{x}^f_t]$$
$$= \mathbf{h}^l_{\mathbf{x}}\mathbf{x}^f_t \otimes \mathbf{h}^{l-1}_{\mathbf{x}} \sigma\boldsymbol{\eta}\boldsymbol{\nu} + \mathbf{h}^{l-1}_{\mathbf{x}} \sigma\boldsymbol{\eta}\boldsymbol{\nu} \otimes \mathbf{h}^l_{\mathbf{x}}\mathbf{x}^f_t + \left(\mathbf{h}^{l-1}_{\mathbf{x}} \otimes \mathbf{h}^{l-1}_{\mathbf{x}}\right)\left((\sigma\boldsymbol{\eta}\boldsymbol{\nu} \otimes \sigma\boldsymbol{\eta}\boldsymbol{\nu}) - \mathbb{E}\left[\sigma\boldsymbol{\eta}\boldsymbol{\epsilon}_{t+1} \otimes \sigma\boldsymbol{\eta}\boldsymbol{\epsilon}_{t+1}\right]\right).$$

I IRFs: Third-Order

I.1 Deriving $GIRF_{\mathbf{x}^f \otimes \mathbf{x}^f \otimes \mathbf{x}^f}\left(j, \boldsymbol{\nu}, \mathbf{x}^f_t\right)$

We first note that

$$\mathbf{x}^f_{t+l} \otimes \mathbf{x}^f_{t+l} \otimes \mathbf{x}^f_{t+l}$$
$$= \mathbf{h}^l_{\mathbf{x}}\mathbf{x}^f_t \otimes \mathbf{h}^l_{\mathbf{x}}\mathbf{x}^f_t \otimes \mathbf{h}^l_{\mathbf{x}}\mathbf{x}^f_t + \mathbf{h}^l_{\mathbf{x}}\mathbf{x}^f_t \otimes \sum_{j=1}^{l} \mathbf{h}^{l-j}_{\mathbf{x}} \sigma\boldsymbol{\eta}\boldsymbol{\epsilon}_{t+j} \otimes \mathbf{h}^l_{\mathbf{x}}\mathbf{x}^f_t$$
$$+ \sum_{j=1}^{l} \mathbf{h}^{l-j}_{\mathbf{x}} \sigma\boldsymbol{\eta}\boldsymbol{\epsilon}_{t+j} \otimes \mathbf{h}^l_{\mathbf{x}}\mathbf{x}^f_t \otimes \mathbf{h}^l_{\mathbf{x}}\mathbf{x}^f_t + \sum_{j=1}^{l} \mathbf{h}^{l-j}_{\mathbf{x}} \sigma\boldsymbol{\eta}\boldsymbol{\epsilon}_{t+j} \otimes \sum_{j=1}^{l} \mathbf{h}^{l-j}_{\mathbf{x}} \sigma\boldsymbol{\eta}\boldsymbol{\epsilon}_{t+j} \otimes \mathbf{h}^l_{\mathbf{x}}\mathbf{x}^f_t$$
$$+ \mathbf{h}^l_{\mathbf{x}}\mathbf{x}^f_t \otimes \mathbf{h}^l_{\mathbf{x}}\mathbf{x}^f_t \otimes \sum_{j=1}^{l} \mathbf{h}^{l-j}_{\mathbf{x}} \sigma\boldsymbol{\eta}\boldsymbol{\epsilon}_{t+j} + \mathbf{h}^l_{\mathbf{x}}\mathbf{x}^f_t \otimes \sum_{j=1}^{l} \mathbf{h}^{l-j}_{\mathbf{x}} \sigma\boldsymbol{\eta}\boldsymbol{\epsilon}_{t+j} \otimes \sum_{j=1}^{l} \mathbf{h}^{l-j}_{\mathbf{x}} \sigma\boldsymbol{\eta}\boldsymbol{\epsilon}_{t+j}$$
$$+ \sum_{j=1}^{l} \mathbf{h}^{l-j}_{\mathbf{x}} \sigma\boldsymbol{\eta}\boldsymbol{\epsilon}_{t+j} \otimes \mathbf{h}^l_{\mathbf{x}}\mathbf{x}^f_t \otimes \sum_{j=1}^{l} \mathbf{h}^{l-j}_{\mathbf{x}} \sigma\boldsymbol{\eta}\boldsymbol{\epsilon}_{t+j} + \sum_{j=1}^{l} \mathbf{h}^{l-j}_{\mathbf{x}} \sigma\boldsymbol{\eta}\boldsymbol{\epsilon}_{t+j} \otimes \sum_{j=1}^{l} \mathbf{h}^{l-j}_{\mathbf{x}} \sigma\boldsymbol{\eta}\boldsymbol{\epsilon}_{t+j} \otimes \sum_{j=1}^{l} \mathbf{h}^{l-j}_{\mathbf{x}} \sigma\boldsymbol{\eta}\boldsymbol{\epsilon}_{t+j}$$

Using the definition of $\boldsymbol{\delta}_{t+j}$ from appendix H, we have

$$\tilde{\mathbf{x}}^f_{t+l} \otimes \tilde{\mathbf{x}}^f_{t+l} \otimes \tilde{\mathbf{x}}^f_{t+l} = \mathbf{h}^l_{\mathbf{x}}\mathbf{x}^f_t \otimes \mathbf{h}^l_{\mathbf{x}}\mathbf{x}^f_t \otimes \mathbf{h}^l_{\mathbf{x}}\mathbf{x}^f_t + \mathbf{h}^l_{\mathbf{x}}\mathbf{x}^f_t \otimes \sum_{j=1}^{l} \mathbf{h}^{l-j}_{\mathbf{x}} \sigma\boldsymbol{\eta}\boldsymbol{\delta}_{t+j} \otimes \mathbf{h}^l_{\mathbf{x}}\mathbf{x}^f_t$$
$$+ \sum_{j=1}^{l} \mathbf{h}^{l-j}_{\mathbf{x}} \sigma\boldsymbol{\eta}\boldsymbol{\delta}_{t+j} \otimes \mathbf{h}^l_{\mathbf{x}}\mathbf{x}^f_t \otimes \mathbf{h}^l_{\mathbf{x}}\mathbf{x}^f_t$$
$$+ \sum_{j=1}^{l} \mathbf{h}^{l-j}_{\mathbf{x}} \sigma\boldsymbol{\eta}\boldsymbol{\delta}_{t+j} \otimes \sum_{j=1}^{l} \mathbf{h}^{l-j}_{\mathbf{x}} \sigma\boldsymbol{\eta}\boldsymbol{\delta}_{t+j} \otimes \mathbf{h}^l_{\mathbf{x}}\mathbf{x}^f_t$$

$$+\mathbf{h}_{\mathbf{x}}^{l}\mathbf{x}_t^f\otimes\mathbf{h}_{\mathbf{x}}^{l}\mathbf{x}_t^f\otimes\sum_{j=1}^{l}\mathbf{h}_{\mathbf{x}}^{l-j}\sigma\boldsymbol{\eta}\boldsymbol{\delta}_{t+j}$$
$$+\mathbf{h}_{\mathbf{x}}^{l}\mathbf{x}_t^f\otimes\sum_{j=1}^{l}\mathbf{h}_{\mathbf{x}}^{l-j}\sigma\boldsymbol{\eta}\boldsymbol{\delta}_{t+j}\otimes\sum_{j=1}^{l}\mathbf{h}_{\mathbf{x}}^{l-j}\sigma\boldsymbol{\eta}\boldsymbol{\delta}_{t+j}$$
$$+\sum_{j=1}^{l}\mathbf{h}_{\mathbf{x}}^{l-j}\sigma\boldsymbol{\eta}\boldsymbol{\delta}_{t+j}\otimes\mathbf{h}_{\mathbf{x}}^{l}\mathbf{x}_t^f\otimes\sum_{j=1}^{l}\mathbf{h}_{\mathbf{x}}^{l-j}\sigma\boldsymbol{\eta}\boldsymbol{\delta}_{t+j}$$
$$+\sum_{j=1}^{l}\mathbf{h}_{\mathbf{x}}^{l-j}\sigma\boldsymbol{\eta}\boldsymbol{\delta}_{t+j}\otimes\sum_{j=1}^{l}\mathbf{h}_{\mathbf{x}}^{l-j}\sigma\boldsymbol{\eta}\boldsymbol{\delta}_{t+j}\otimes\sum_{j=1}^{l}\mathbf{h}_{\mathbf{x}}^{l-j}\sigma\boldsymbol{\eta}\boldsymbol{\delta}_{t+j}$$

Simple algebra then gives

$$GIRF_{\mathbf{x}^f\otimes\mathbf{x}^f\otimes\mathbf{x}^f}\left(j,\boldsymbol{\nu},\mathbf{x}_t^f\right)=\mathbb{E}_t\left[\tilde{\mathbf{x}}_{t+l}^f\otimes\tilde{\mathbf{x}}_{t+l}^f\otimes\tilde{\mathbf{x}}_{t+l}^f|\mathbf{x}_t^f\right]-\mathbb{E}_t\left[\mathbf{x}_{t+l}^f\otimes\mathbf{x}_{t+l}^f\otimes\mathbf{x}_{t+l}^f|\mathbf{x}_t^f\right]$$
$$=\mathbf{h}_{\mathbf{x}}^{l}\mathbf{x}_t^f\otimes\mathbf{h}_{\mathbf{x}}^{l-1}\sigma\boldsymbol{\eta}\boldsymbol{\nu}\otimes\mathbf{h}_{\mathbf{x}}^{l}\mathbf{x}_t^f$$
$$+\mathbf{h}_{\mathbf{x}}^{l-1}\sigma\boldsymbol{\eta}\boldsymbol{\nu}\otimes\mathbf{h}_{\mathbf{x}}^{l}\mathbf{x}_t^f\otimes\mathbf{h}_{\mathbf{x}}^{l}\mathbf{x}_t^f$$
$$+\mathbf{h}_{\mathbf{x}}^{l}\mathbf{x}_t^f\otimes\mathbf{h}_{\mathbf{x}}^{l}\mathbf{x}_t^f\otimes\mathbf{h}_{\mathbf{x}}^{l-1}\sigma\boldsymbol{\eta}\boldsymbol{\nu}$$
$$+\left(\mathbf{h}_{\mathbf{x}}^{l-1}\otimes\mathbf{h}_{\mathbf{x}}^{l-1}\right)\left((\sigma\boldsymbol{\eta}\boldsymbol{\nu}\otimes\sigma\boldsymbol{\eta}\boldsymbol{\nu})-\mathbb{E}\left[\sigma\boldsymbol{\eta}\boldsymbol{\epsilon}_{t+1}\otimes\sigma\boldsymbol{\eta}\boldsymbol{\epsilon}_{t+1}\right]\right)\otimes\mathbf{h}_{\mathbf{x}}^{l}\mathbf{x}_t^f$$
$$+\mathbf{h}_{\mathbf{x}}^{l}\mathbf{x}_t^f\otimes\left(\mathbf{h}_{\mathbf{x}}^{l-1}\otimes\mathbf{h}_{\mathbf{x}}^{l-1}\right)\left((\sigma\boldsymbol{\eta}\boldsymbol{\nu}\otimes\sigma\boldsymbol{\eta}\boldsymbol{\nu})-\mathbb{E}\left[\sigma\boldsymbol{\eta}\boldsymbol{\epsilon}_{t+1}\otimes\sigma\boldsymbol{\eta}\boldsymbol{\epsilon}_{t+1}\right]\right)$$
$$+\mathbf{h}_{\mathbf{x}}^{l-1}\sigma\boldsymbol{\eta}\boldsymbol{\nu}\otimes\mathbf{h}_{\mathbf{x}}^{l}\mathbf{x}_t^f\otimes\mathbf{h}_{\mathbf{x}}^{l-1}\sigma\boldsymbol{\eta}\boldsymbol{\nu}-\mathbb{E}\left[\mathbf{h}_{\mathbf{x}}^{l-1}\sigma\boldsymbol{\eta}\boldsymbol{\epsilon}_{t+1}\otimes\mathbf{h}_{\mathbf{x}}^{l}\mathbf{x}_t^f\otimes\mathbf{h}_{\mathbf{x}}^{l-1}\sigma\boldsymbol{\eta}\boldsymbol{\epsilon}_{t+1}\right]$$
$$+\mathbf{h}_{\mathbf{x}}^{l-1}\sigma\boldsymbol{\eta}\boldsymbol{\nu}\otimes\mathbf{h}_{\mathbf{x}}^{l-1}\sigma\boldsymbol{\eta}\boldsymbol{\nu}\otimes\mathbf{h}_{\mathbf{x}}^{l-1}\sigma\boldsymbol{\eta}\boldsymbol{\nu}$$
$$+\sum_{j=2}^{l}\mathbf{h}_{\mathbf{x}}^{l-1}\sigma\boldsymbol{\eta}\boldsymbol{\nu}\otimes\left(\mathbf{h}_{\mathbf{x}}^{l-j}\otimes\mathbf{h}_{\mathbf{x}}^{l-j}\right)\mathbb{E}\left[\sigma\boldsymbol{\eta}\boldsymbol{\epsilon}_{t+1}\otimes\sigma\boldsymbol{\eta}\boldsymbol{\epsilon}_{t+1}\right]$$
$$+\sum_{j=2}^{l}\mathbb{E}\left[\sigma\boldsymbol{\eta}\boldsymbol{\epsilon}_{t+1}\otimes\sigma\boldsymbol{\eta}\boldsymbol{\epsilon}_{t+1}\right]\left(\mathbf{h}_{\mathbf{x}}^{l-j}\otimes\mathbf{h}_{\mathbf{x}}^{l-j}\right)\otimes\mathbf{h}_{\mathbf{x}}^{l-1}\sigma\boldsymbol{\eta}\boldsymbol{\nu}$$
$$+\sum_{j=2}^{l}\mathbb{E}\left[\mathbf{h}_{\mathbf{x}}^{l-j}\sigma\boldsymbol{\eta}\boldsymbol{\epsilon}_{t+1}\otimes\mathbf{h}_{\mathbf{x}}^{l-1}\sigma\boldsymbol{\eta}\boldsymbol{\nu}\otimes\mathbf{h}_{\mathbf{x}}^{l-j}\sigma\boldsymbol{\eta}\boldsymbol{\epsilon}_{t+1}\right]$$
$$+\mathbb{E}\left[\mathbf{h}_{\mathbf{x}}^{l-1}\sigma\boldsymbol{\eta}\boldsymbol{\epsilon}_{t+1}\otimes\mathbf{h}_{\mathbf{x}}^{l-1}\sigma\boldsymbol{\eta}\boldsymbol{\epsilon}_{t+1}\otimes\mathbf{h}_{\mathbf{x}}^{l-1}\sigma\boldsymbol{\eta}\boldsymbol{\epsilon}_{t+1}\right]$$

I.2 Deriving $GIRF_{\mathbf{x}^f\otimes\mathbf{x}^s}\left(j,\boldsymbol{\nu},\left(\mathbf{x}_t^f,\mathbf{x}_t^s\right)\right)$

Using the law of motion for $\mathbf{x}_t^f\otimes\mathbf{x}_t^s$, we first note that

$$\mathbf{x}_{t+l}^f\otimes\mathbf{x}_{t+l}^s=(\mathbf{h}_{\mathbf{x}}\otimes\mathbf{h}_{\mathbf{x}})^l\left(\mathbf{x}_t^f\otimes\mathbf{x}_t^s\right)+\sum_{i=0}^{l-1}(\mathbf{h}_{\mathbf{x}}\otimes\mathbf{h}_{\mathbf{x}})^{l-1-i}\left(\mathbf{h}_{\mathbf{x}}\otimes\tfrac{1}{2}\mathbf{H}_{\mathbf{xx}}\right)\left(\mathbf{x}_{t+i}^f\otimes\mathbf{x}_{t+i}^f\otimes\mathbf{x}_{t+i}^f\right)$$
$$+\sum_{i=0}^{l-1}(\mathbf{h}_{\mathbf{x}}\otimes\mathbf{h}_{\mathbf{x}})^{l-1-i}\left(\mathbf{h}_{\mathbf{x}}\otimes\tfrac{1}{2}\mathbf{h}_{\sigma\sigma}\sigma^2\right)\mathbf{x}_{t+i}^f$$
$$+\sum_{i=0}^{l-1}(\mathbf{h}_{\mathbf{x}}\otimes\mathbf{h}_{\mathbf{x}})^{l-1-i}\left(\sigma\boldsymbol{\eta}\otimes\tfrac{1}{2}\mathbf{h}_{\sigma\sigma}\sigma^2\right)\boldsymbol{\epsilon}_{t+1+i}$$
$$+\sum_{i=0}^{l-1}(\mathbf{h}_{\mathbf{x}}\otimes\mathbf{h}_{\mathbf{x}})^{l-1-i}\left(\sigma\boldsymbol{\eta}\otimes\mathbf{h}_{\mathbf{x}}\right)\left(\boldsymbol{\epsilon}_{t+1+i}\otimes\mathbf{x}_{t+i}^s\right)$$
$$+\sum_{i=0}^{l-1}(\mathbf{h}_{\mathbf{x}}\otimes\mathbf{h}_{\mathbf{x}})^{l-1-i}\left(\sigma\boldsymbol{\eta}\otimes\tfrac{1}{2}\mathbf{H}_{\mathbf{xx}}\right)\left(\boldsymbol{\epsilon}_{t+1+i}\otimes\mathbf{x}_{t+i}^f\otimes\mathbf{x}_{t+i}^f\right)$$

Using the definition of $\boldsymbol{\delta}_{t+j}$ from appendix H, we obtain

$$\tilde{\mathbf{x}}_{t+l}^f\otimes\tilde{\mathbf{x}}_{t+l}^s=(\mathbf{h}_{\mathbf{x}}\otimes\mathbf{h}_{\mathbf{x}})^l\left(\mathbf{x}_t^f\otimes\mathbf{x}_t^s\right)+\sum_{i=0}^{l-1}(\mathbf{h}_{\mathbf{x}}\otimes\mathbf{h}_{\mathbf{x}})^{l-1-i}\left(\mathbf{h}_{\mathbf{x}}\otimes\tfrac{1}{2}\mathbf{H}_{\mathbf{xx}}\right)\left(\tilde{\mathbf{x}}_{t+i}^f\otimes\tilde{\mathbf{x}}_{t+i}^f\otimes\tilde{\mathbf{x}}_{t+i}^f\right)$$
$$+\sum_{i=0}^{l-1}(\mathbf{h}_{\mathbf{x}}\otimes\mathbf{h}_{\mathbf{x}})^{l-1-i}\left(\mathbf{h}_{\mathbf{x}}\otimes\tfrac{1}{2}\mathbf{h}_{\sigma\sigma}\sigma^2\right)\tilde{\mathbf{x}}_{t+i}^f$$

$$+\sum_{i=0}^{l-1}(\mathbf{h_x}\otimes\mathbf{h_x})^{l-1-i}\left(\sigma\boldsymbol{\eta}\otimes\tfrac{1}{2}\mathbf{h}_{\sigma\sigma}\sigma^2\right)\boldsymbol{\delta}_{t+1+i}$$
$$+\sum_{i=0}^{l-1}(\mathbf{h_x}\otimes\mathbf{h_x})^{l-1-i}\left(\sigma\boldsymbol{\eta}\otimes\mathbf{h_x}\right)\left(\boldsymbol{\delta}_{t+1+i}\otimes\tilde{\mathbf{x}}^s_{t+i}\right)$$
$$+\sum_{i=0}^{l-1}(\mathbf{h_x}\otimes\mathbf{h_x})^{l-1-i}\left(\sigma\boldsymbol{\eta}\otimes\tfrac{1}{2}\mathbf{H_{xx}}\right)\left(\boldsymbol{\delta}_{t+1+i}\otimes\tilde{\mathbf{x}}^f_{t+i}\otimes\tilde{\mathbf{x}}^f_{t+i}\right)$$

Simple algebra then implies

$$GIRF_{\mathbf{x}^f\otimes\mathbf{x}^s}\left(j,\boldsymbol{\nu},\left(\mathbf{x}^f_t,\mathbf{x}^s_t\right)\right)=\sum_{i=1}^{l-1}(\mathbf{h_x}\otimes\mathbf{h_x})^{l-1-i}\left(\mathbf{h_x}\otimes\tfrac{1}{2}\mathbf{H_{xx}}\right)GIRF_{\mathbf{x}^f\otimes\mathbf{x}^f\otimes\mathbf{x}^f}\left(i,\boldsymbol{\nu},\mathbf{x}^f_t\right)$$
$$+\sum_{i=1}^{l-1}(\mathbf{h_x}\otimes\mathbf{h_x})^{l-1-i}\left(\mathbf{h_x}\otimes\tfrac{1}{2}\mathbf{h}_{\sigma\sigma}\sigma^2\right)GIRF_{\mathbf{x}^f}(i,\boldsymbol{\nu})$$
$$+(\mathbf{h_x}\otimes\mathbf{h_x})^{l-1}\left(\sigma\boldsymbol{\eta}\boldsymbol{\nu}\otimes\left(\mathbf{h_x}\mathbf{x}^s_t+\tfrac{1}{2}\mathbf{H_{xx}}\left(\mathbf{x}^f_t\otimes\mathbf{x}^f_t\right)+\tfrac{1}{2}\mathbf{h}_{\sigma\sigma}\sigma^2\right)\right)$$

J The New Keynesian Model

J.1 Household

The problem of the household is:

$$\max_{c_t,x_{t+1},h_t,k_{t+1},i_t\ \forall t\geq 0} V_t=\frac{d_t}{1-\phi_2}(c_t-bc_{t-1})^{1-\phi_2}+(z^*_t)^{(1-\phi_2)}\phi_0\frac{(1-h_t)^{1-\phi_1}}{1-\phi_1}+\beta\left(\mathbb{E}_t\left[V^{1-\phi_3}_{t+1}\right]\right)^{\frac{1}{1-\phi_3}}$$
$$\text{s.t. } k_{t+1}=(1-\delta)k_t+i_t-\frac{\kappa}{2}\left(\frac{i_t}{k_t}-\psi\right)^2k_t$$
$$\int D_{t,t+1}x_{t+1}d\omega_{t,t+1}+c_t+\Upsilon^{-1}_t i_t=\frac{x_t}{\pi_t}+h_tw_t+r^k_tk_t+div_t$$

plus a standard no-Ponzi-game condition.

We follow Rudebusch and Swanson (2012) and consider a finite number of states in each period.[17] As in Rudebusch and Swanson (2012), the optimization problem is formulated as a Lagrange problem maximizing

$$\begin{aligned}V_t\ :\ \mathcal{L}=V_t+\mathbb{E}_t\sum_{l=0}^{\infty}\beta^l\gamma_{t+l}&\left[\frac{d_{t+l}}{1-\phi_2}(c_{t+l}-bc_{t+l-1})^{1-\phi_2}+\left(z^*_{t+l}\right)^{(1-\phi_2)}\phi_0\frac{(1-h_{t+l})^{1-\phi_1}}{1-\phi_1}\right]\\
+\mathbb{E}_t\sum_{l=0}^{\infty}\beta^l\gamma_{t+l}&\left[\beta\left(E_t\left[V^{1-\phi_3}_{t+l+1}\right]\right)^{\frac{1}{1-\phi_3}}-V_{t+l}\right]\\
+\mathbb{E}_t\sum_{l=0}^{\infty}\beta^l\lambda_{t+l}&[\frac{x_{t+l}}{\pi_{t+l}}+h_{t+l}w_{t+l}+r^k_{t+l}k_{t+l}+div_{t+l}-\int D_{t+l,t+l+1}x_{t+l+1}d\omega_{t+l,t+1+l}-c_{t+l}-\Upsilon^{-1}_{t+l}i_{t+l}]\\
+\mathbb{E}_t\sum_{l=0}^{\infty}\beta^lq_{t+l}\lambda_{t+l}&\left[(1-\delta)k_{t+l}+i_{t+l}-\frac{\kappa}{2}\left(\frac{i_{t+l}}{k_{t+l}}-\psi\right)^2k_{t+l}-k_{t+1+l}\right]\end{aligned}$$

[17] This assumption is without loss of generality as shown by Epstein and Zin (1989). Moveover, this section covers the case where the value function is positive. The derivations for a negative value function are almost identical and therefore not included.

The first-order conditions for an interior solution are given by:

$$\frac{\partial \mathcal{L}}{\partial c_t} = \gamma_t d_t \left(c_t - bc_{t-1}\right)^{-\phi_2} - b\beta \mathbb{E}_t \gamma_{t+1} d_{t+1} \left(c_{t+1} - bc_t\right)^{-\phi_2} - \lambda_t = 0$$

$$\frac{\partial \mathcal{L}}{\partial x_{t+1}(s)} = prob_t(s)\left[\beta \lambda_{t+1}(s) \frac{1}{\pi_{t+1}(s)} - D_{t,t+1}(s)\lambda_t\right] = 0$$

Then

$$D_{t,t+1}(s) = \beta \frac{\lambda_{t+1}(s)}{\lambda_t}\frac{1}{\pi_{t+1}(s)}$$

$$\frac{\partial \mathcal{L}}{\partial h_t} = -\gamma_t \left(z_t^*\right)^{(1-\phi_2)} \phi_0 \left(1-h_t\right)^{-\phi_1} + \lambda_t w_t = 0$$

$$\frac{\partial \mathcal{L}}{\partial k_{t+1}} = \lambda_t q_t(-1) + \mathbb{E}_t \beta \lambda_{t+1}[r_{t+1}^k + q_{t+1}(1-\delta) - q_{t+1}\frac{\kappa}{2}\left(\frac{i_{t+1}}{k_{t+1}} - \psi\right)^2 + q_{t+1}\kappa\left(\frac{i_{t+1}}{k_{t+1}} - \psi\right)\frac{i_{i+1}}{k_{t+1}^2} k_{t+1}] = 0$$

$$\frac{\partial \mathcal{L}}{\partial i_t} = -\left(\Upsilon_t\right)^{-1}\lambda_t + q_t \lambda_t \left(1 - \kappa\left(\frac{i_t}{k_t} - \psi\right)\right) = 0$$

$$\frac{\partial \mathcal{L}}{\partial V_{t+1}(s)} = \gamma_t(s)\left(\beta \frac{1}{1-\phi_3}\left(E_t\left[V_{t+1}^{1-\phi_3}\right]\right)^{\frac{1}{1-\phi_3} - \frac{1-\phi_3}{1-\phi_3}}(1-\phi_3)V_{t+1}(s)^{-\phi_3} prob_t(s)\right) - \gamma_{t+1}(s) prob_t(s)\beta = 0$$

and

$$\gamma_{t+1}(s) = \gamma_t(s)\left(E_t\left[V_{t+1}^{1-\phi_3}\right]\right)^{\frac{\phi_3}{1-\phi_3}} V_{t+1}(s)^{-\phi_3} \text{ for all states.}$$

The price of a one-period bond is given by

$$\lambda_t = \beta \exp\{r_{t,1}\}\mathbb{E}_t\left[\frac{\lambda_{t+1}}{\pi_{t+1}}\right].$$

J.1.1 Final Good Producers

The final good producer chooses $y_{i,t}$ for $i \in [0,1]$ to solve:

$$\max_{y_{t,i}} P_t y_t - \int_0^1 P_{i,t} y_{i,t} di$$

$$\text{s.t. } y_t = \left(\int_0^1 y_{i,t}^{\frac{\eta-1}{\eta}} di\right)^{\frac{\eta}{\eta-1}}$$

The first-order condition of the problem is given by

$$y_{i,t} = \left(\frac{P_{i,t}}{P_t}\right)^{-\eta} y_t$$

To find the expression for the aggregate price level we use the zero-profit condition

$$P_t y_t = \int_0^1 P_{i,t} y_{i,t} di = \int_0^1 P_{i,t}\left(\frac{P_{i,t}}{P_t}\right)^{-\eta} y_t di = y_t P_t^{\eta}\int_0^1 \left(P_{i,t}\right)^{1-\eta} di$$

and then:

$$P_t = \left[\int_0^1 (P_{i,t})^{1-\eta}\, di\right]^{\frac{1}{1-\eta}}$$

J.1.2 Intermediate Good Producer

The problem of the intermediate good producer is

$$\max_{k_{i,t},h_{i,t},P_{i,t}} \mathbb{E}_t \sum_{j=0}^{\infty} D_{t,t+j} P_{t+j} \left[\frac{P_{i,t+j}}{P_{t+j}}\left(\frac{P_{i,t+j}}{P_{t+j}}\right)^{-\eta} y_{t+j} - w_{t+j}h_{i,t+j} - r^k_{t+j}k_{i,t+j} - \frac{\xi_p}{2}\left(\frac{P_{i,t+j}}{P_{i,t+j-1}}\frac{1}{\pi_{ss}} - 1\right)^2 y_{t+j}\right]$$

$$\text{s.t. } a_t k^{\theta}_{i,t} \left(z_t h_{i,t}\right)^{1-\theta} = \left(\frac{P_{i,t}}{P_t}\right)^{-\eta} y_t$$

plus a no-Ponzi-game condition.

The Lagrange function for this problem reads:

$$\begin{aligned}
\mathcal{L} \;=\; & \mathbb{E}_t \sum_{j=0}^{\infty} D_{t,t+j} P_{t+j} \left[\left(\frac{P_{i,t+j}}{P_{t+j}}\right)^{1-\eta} y_{t+j} - w_{t+j}h_{i,t+j} - r^k_{t+j}k_{i,t+j} - \frac{\xi_p}{2}\left(\frac{P_{i,t+j}}{P_{i,t+j-1}}\frac{1}{\pi_{ss}} - 1\right)^2 y_{t+j}\right] \\
& + \mathbb{E}_t \sum_{j=0}^{\infty} D_{t,t+j} P_{t+j} mc_{i,t+j} \left[a_{t+j} k^{\theta}_{i,t+j} \left(z_t h_{i,t+j}\right)^{1-\theta} - \left(\frac{P_{i,t+j}}{P_{t+j}}\right)^{-\eta} y_{t+j}\right]
\end{aligned}$$

The first-order conditions for an interior solution are:

$$\frac{\partial \mathcal{L}}{\partial k_{i,t}} = P_t \left(mc_{i,t}\theta a_t k^{\theta-1}_{i,t} \left(z_t h_{i,t}\right)^{1-\theta} - r^k_t\right) = 0$$

$$\frac{\partial \mathcal{L}}{\partial h_{i,t}} = P_t \left(-w_t + mc_{i,t}\left(1-\theta\right) z_t^{1-\theta} a_t k^{\theta}_{i,t} h^{-\theta}_{i,t}\right) = 0$$

$$\begin{aligned}
\frac{\partial \mathcal{L}}{\partial P_{i,t}} \;=\; & P_t \mathbb{E}_t \{(1-\eta)\left(\frac{P_{i,t}}{P_t}\right)^{-\eta} \frac{1}{P_t} y_t - \xi_p \left(\frac{P_{i,t}}{P_{i,t-1}}\frac{1}{\pi_{ss}} - 1\right)\frac{1}{P_{i,t-1}}\frac{1}{\pi_{ss}} y_t \\
& - D_{t,t+1}\frac{P_{t+1}}{P_t}\xi_p \left(\frac{P_{i,t+1}}{P_{i,t}}\frac{1}{\pi_{ss}} - 1\right)\left(\frac{-P_{i,t+1}}{P^2_{i,t}}\frac{1}{\pi_{ss}}\right) y_{t+1} \\
& + mc_{i,t}\eta \left(\frac{P_{i,t}}{P_t}\right)^{-\eta-1} \frac{y_t}{P_t}\} = 0
\end{aligned}$$

or

$$\begin{aligned}
\xi_p \left(\frac{P_{i,t}}{P_{i,t-1}}\frac{1}{\pi_{ss}} - 1\right)\frac{y_t}{\pi_{ss}P_{i,t-1}} \;=\; & (1-\eta)\left(\frac{P_{i,t}}{P_t}\right)^{-\eta}\frac{y_t}{P_t} \\
& + \mathbb{E}_t\left[D^{\text{real}}_{t,t+1}\xi_p \left(\frac{P_{i,t+1}}{P_{i,t}}\frac{1}{\pi_{ss}} - 1\right)\frac{P_{i,t+1}}{P^2_{i,t}}\frac{y_{t+1}}{\pi_{ss}}\right] \\
& + mc_{i,t}\eta \left(\frac{P_{i,t}}{P_t}\right)^{-\eta-1}\frac{y_t}{P_t}
\end{aligned}$$

as $P_t > 0$ and $D_{t,t+1} = D^{\text{real}}_{t,t+1}\frac{1}{\pi_{t+1}} = D^{\text{real}}_{t,t+1}\frac{P_t}{P_{t+1}}$ and $D^{\text{real}}_{t,t+1} = \beta\frac{\lambda_{t+1}}{\lambda_t}$.

Since all firms face the same problem, the three conditions simplify to

$$mc_t\theta a_t k_t^{\theta-1}\left(z_t h_t\right)^{1-\theta} = r_t^k$$
$$w_t = mc_t\left(1-\theta\right) z_t^{1-\theta} a_t k_t^{\theta} h_t^{-\theta}$$
$$mc_t = \frac{(\eta-1)}{\eta} - E_t\left[\beta\frac{\lambda_{t+1}}{\lambda_t}\frac{\xi_p}{\eta}\left(\frac{\pi_{t+1}}{\pi_{ss}}-1\right)\frac{\pi_{t+1}y_{t+1}}{\pi_{ss}y_t}\right] + \frac{\xi_p}{\eta}\left(\frac{\pi_t}{\pi_{ss}}-1\right)\frac{\pi_t}{\pi_{ss}}$$

J.2 Aggregation and Market Clearing Conditions

The resource constraint is:

$$y_t = c_t + i_t\Upsilon_t^{-1}$$

The labor market clears when

$$h_t = \int_0^1 h_{i,t}di = \int_0^1\left(\frac{y_{i,t}}{z_t^{1-\theta}a_t k_t^{\theta}}\right)^{\frac{1}{1-\theta}}di = \left(\frac{y_t}{z_t^{1-\theta}a_t k_t^{\theta}}\right)^{\frac{1}{1-\theta}}\int_0^1\left(\frac{P_{i,t}}{P_t}\right)^{\frac{-\eta}{1-\theta}}di$$

or

$$y_t = \left[\int_0^1\left(\frac{P_{i,t}}{P_t}\right)^{\frac{-\eta}{1-\theta}}di\right]^{-(1-\theta)} a_t k_t^{\theta}\left(z_t h_t\right)^{1-\theta}$$

The term in square brackets is the index of price dispersion across firms. Under the assumption of Rotemberg costs, firms charge the same price and therefore this index is equal to one. Hence,

$$y_t = a_t k_t^{\theta}\left(z_t h_t\right)^{1-\theta}$$

The final goods market clears when

$$y_t = c_t + \left(\Upsilon_t\right)^{-1} i_t$$

J.3 Equilibrium Conditions

For convenience, we enumerate the equilibrium conditions of the model,

	Household
1.	$V_t = \frac{d_t}{1-\phi_2}\left(c_t - bc_{t-1}\right)^{1-\phi_2} + \left(z_t^*\right)^{(1-\phi_2)}\phi_0 \frac{(1-h_t)^{1-\phi_1}}{1-\phi_1} + \beta\left(\mathbb{E}_t\left[V_{t+1}^{1-\phi_3}\right]\right)^{\frac{1}{1-\phi_3}}$
2.	$\lambda_t = \gamma_t d_t \left(c_t - bc_{t-1}\right)^{-\phi_2} - b\beta\mathbb{E}_t\gamma_{t+1}d_{t+1}\left(c_{t+1}-bc_t\right)^{-\phi_2}$
3.	$q_t\lambda_t = \mathbb{E}_t\beta\lambda_{t+1}[r^k_{t+1} + q_{t+1}\left(1-\delta\right) - q_{t+1}\frac{\kappa}{2}\left(\frac{i_{t+1}}{k_{t+1}}-\psi\right)^2 + q_{t+1}\kappa\left(\frac{i_{t+1}}{k_{t+1}}-\psi\right)\frac{i_{t+1}}{k_{t+1}}]$
4.	$\gamma_t \left(z_t^*\right)^{1-\phi_2}\phi_0\left(1-h_t\right)^{-\phi_1} = \lambda_t w_t$
5.	$1 = q_t\Upsilon_t\left(1-\kappa\left(\frac{i_i}{k_t}-\psi\right)\right)$
6.	$\lambda_t = \beta\exp\left\{r_{t,1}\right\}\mathbb{E}_t\left[\frac{\lambda_{t+1}}{\pi_{t+1}}\right]$
	Firms
7.	$mc_t a_t z_t\left(1-\theta\right)\left(\frac{z_t h_t}{k_t}\right)^{-\theta} = w_t$
8.	$a_t mc_t\theta\left(\frac{z_t h_t}{k_t}\right)^{1-\theta} = r_t^k$
9.	$mc_t = \frac{(\eta-1)}{\eta} - \mathbb{E}_t\left[\beta\frac{\lambda_{t+1}}{\lambda_t}\frac{\xi_p}{\eta}\left(\frac{\pi_{t+1}}{\pi_{ss}}-1\right)\frac{\pi_{t+1}y_{t+1}}{\pi_{ss}y_t}\right] + \frac{\xi_p}{\eta}\left(\frac{\pi_t}{\pi_{ss}}-1\right)\frac{\pi_t}{\pi_{ss}}$
	Central Bank
10.	$r_{t,1} = \left(1-\rho_r\right)r_{ss} + \rho_r r_{t-1,1} + \beta_\pi\log\left(\frac{\pi_t}{\pi_{ss}}\right) + \beta_y\ln\left(\frac{y_t}{z_t^*Y_{ss}}\right)$
	Other Relations
11.	$a_t k_t^\theta\left(z_t h_t\right)^{1-\theta} = y_t$
12.	$y_t = c_t + \left(\Upsilon_t\right)^{-1} i_t$
13.	$k_{t+1} = \left(1-\delta\right)k_t + i_t - \frac{\kappa}{2}\left(\frac{i_i}{k_t}-\psi\right)^2 k_t$
14.	$z_t^* \equiv \Upsilon_t^{\frac{\theta}{1-\theta}} z_t$ and $\mu_{z^*,t} \equiv \mu_{\Upsilon,t}^{\theta/(1-\theta)}\mu_{z,t}$
15.	$\gamma_{t+1}\left(s\right) = \gamma_t\left(s\right)\left(\mathbb{E}_t\left[V_{t+1}^{1-\phi_3}\right]\right)^{\frac{\phi_3}{1-\phi_3}} V_{t+1}\left(s\right)^{-\phi_3}$ for all states
	Exogenous Shocks
16.	$\ln\left(\mu_{z,t}\right) = \ln\left(\mu_{z,ss}\right)$ and $z_{t+1} \equiv z_t\mu_{z,t+1}$
17.	$\ln\left(\mu_{\Upsilon,t}\right) = \ln\mu_{\Upsilon,ss}$ and $\Upsilon_{t+1} \equiv \Upsilon_t\mu_{\Upsilon,t+1}$
18.	$\ln a_{t+1} = \rho_a\ln a_t + \epsilon_{a,t+1}$
19.	$\ln d_{t+1} = \rho_d\ln d_t + \epsilon_{d,t+1}$

J.4 A Transformation of the DSGE Model

As mentioned in the main text, we need to transform the model into a stationary one to efficiently compute its solution by the perturbation method. To do so, we define: $C_t \equiv \frac{c_t}{z_t^*}$, $R_t^k \equiv \Upsilon_t r_t^k$, $Q_t \equiv \Upsilon_t q_t$, $I_t \equiv \frac{i_t}{\Upsilon_t z_t^*}$, $W_t \equiv \frac{w_t}{z_t^*}$, $Y_t \equiv \frac{y_t}{z_t^*}$, $K_{t+1} \equiv \frac{k_{t+1}}{\Upsilon_t^{\frac{1}{1-\theta}} z_t} = \frac{k_{t+1}}{\Upsilon_t z_t^*}$, $\widetilde{V}_t \equiv \frac{V_t}{(z_t^*)^{1-\phi_2}}$, and $\Lambda_t \equiv \frac{\lambda_t}{\gamma_t (z_t^*)^{-\phi_2}}$. It is useful to note that

$$\begin{aligned}
\mu_{\lambda,t+1} &\equiv \frac{\lambda_{t+1}}{\lambda_t} = \frac{\Lambda_{t+1}\gamma_{t+1}\left(z_{t+1}^*\right)^{-\phi_2}}{\Lambda_t\gamma_t\left(z_t^*\right)^{-\phi_2}} \\
&= \frac{\Lambda_{t+1}}{\Lambda_t}\mu_{z^*,t+1}^{-\phi_2}\left(\mathbb{E}_t\left[V_{t+1}^{1-\phi_3}\right]\right)^{\frac{\phi_3}{1-\phi_3}} V_{t+1}^{-\phi_3}
\end{aligned}$$

The value of ψ that ensures that the capital adjustment costs do not affect the steady state is given by

$$\psi \equiv \frac{I_{ss}}{K_{ss}}\mu_{\Upsilon,ss}\mu_{z^*,ss}$$

The new equilibrium conditions are then given by:

	Household
1.	$\widetilde{V}_t = \left[\frac{d_t}{1-\phi_2}\left(C_t - bC_{t-1}\mu_{z^*,t}^{-1}\right)^{1-\phi_2} + \phi_0\frac{(1-h_t)^{1-\phi_1}}{1-\phi_1}\right] + \beta\mu_{z^*,ss}^{(1-\phi_2)}\left(\mathbb{E}_t\left[\widetilde{V_{t+1}}^{1-\phi_3}\right]\right)^{\frac{1}{1-\phi_3}}$
2.	$\Lambda_t = d_t\left(C_t - bC_{t-1}\mu_{z^*,t}^{-1}\right)^{-\phi_2} - b\beta\mathbb{E}_t\{\left[\left(\frac{\left[\mathbb{E}_t\left[\widetilde{V_{t+1}}^{1-\phi_3}\right]\right]^{\frac{1}{1-\phi_3}}}{\widetilde{V_{t+1}(s)}}\right)^{\phi_3}\right]$
	$\times d_{t+1}\left(C_{t+1} - bC_t\mu_{z^*,t+1}^{-1}\right)^{-\phi_2}\left(\mu_{z^*,t+1}\right)^{-\phi_2}\}$
3.	$Q_t = \mathbb{E}_t\frac{\beta\mu_{\lambda,t+1}}{\mu_{\Upsilon,t+1}}[R_{t+1}^k + Q_{t+1}(1-\delta) - Q_{t+1}\frac{\kappa}{2}\left(\frac{I_{t+1}}{K_{t+1}}\mu_{\Upsilon,t+1}\mu_{z^*,t+1} - \frac{I_{ss}}{K_{ss}}\mu_{\Upsilon,ss}\mu_{z^*,ss}\right)^2$
	$+Q_{t+1}\kappa\left(\frac{I_{t+1}}{K_{t+1}}\mu_{\Upsilon,t+1}\mu_{z^*,t+1} - \frac{I_{ss}}{K_{ss}}\mu_{\Upsilon,ss}\mu_{z^*,ss}\right)\frac{I_{t+1}}{K_{t+1}}\mu_{\Upsilon,t+1}\mu_{z^*,t+1}]$
4.	$\phi_0(1-h_t)^{-\phi_1} = \Lambda_t W_t$
5.	$1 = Q_t\left(1 - \kappa\left(\frac{I_t}{K_t}\mu_{\Upsilon,t}\mu_{z^*,t} - \frac{I_{ss}}{K_{ss}}\mu_{\Upsilon,ss}\mu_{z^*,ss}\right)\right)$
6.	$1 = \mathbb{E}_t\left[\beta\mu_{\lambda\, t+1}\frac{\exp\{r_{t,1}\}}{\pi_{t+1}}\right]$
	Firms
7.	$mc_t(1-\theta)a_t\mu_{\Upsilon,t}^{\frac{-\theta}{1-\theta}}K_t^\theta\left(\mu_{z,t}h_t\right)^{-\theta} = W_t$
8.	$mc_t a_t\theta\mu_{\Upsilon,t}K_t^{\theta-1}\left(\mu_{z,t}h_t\right)^{1-\theta} = R_t^k$
9.	$mc_t = \frac{(\eta-1)}{\eta} + \frac{\xi_p}{\eta}\left(\frac{\pi_t}{\pi_{ss}} - 1\right)\frac{\pi_t}{\pi_{ss}} - \mathbb{E}_t\left[\beta\mu_{\lambda,t+1}\frac{\xi_p}{\eta}\left(\frac{\pi_{t+1}}{\pi_{ss}} - 1\right)\frac{\pi_{t+1}Y_{t+1}}{\pi_{ss}Y_t}\mu_{z^*,t+1}\right]$
	Central Bank
10.	$r_{t,1} = (1-\rho_r)r_{ss} + \rho_r r_{t-1,1} + \beta_\pi\log\left(\frac{\pi_t}{\pi_{ss}}\right) + \beta_y\log\left(\frac{Y_t}{Y_{ss}}\right)$
	Other Relations
11.	$a_t\left(K_t\mu_{\Upsilon,t}^{\frac{-1}{1-\theta}}\mu_{z,t}^{-1}\right)^\theta h_t^{1-\theta} = Y_t$
12.	$Y_t = C_t + I_t$
13.	$K_{t+1} = (1-\delta)K_t\left(\mu_{\Upsilon,t}\mu_{z^*,t}\right)^{-1} + I_t - \frac{\kappa}{2}\left(\frac{I_i}{K_t}\mu_{\Upsilon,t}\mu_{z^*,t} - \frac{I_{ss}}{K_{ss}}\mu_{\Upsilon,ss}\mu_{z^*,ss}\right)^2 K_t\left(\mu_{\Upsilon,t}\mu_{z^*,t}\right)^{-1}$
14.	$\mu_{z^*,t} \equiv \mu_{\Upsilon,t}^{\theta/(1-\theta)}\mu_{z,t}$
15.	$\mu_{\lambda,t+1}(s) = \frac{\Lambda_{t+1}(s)}{\Lambda_t}\mu_{z^*,t+1}^{-\phi_2}(s)\left(\mathbb{E}_t\left[V_{t+1}^{1-\phi_3}\right]\right)^{\frac{\phi_3}{1-\phi_3}}V_{t+1}(s)^{-\phi_3}$ for all states
	Exogenous Shocks
16.	$\ln\left(\mu_{z,t}\right) = \ln\left(\mu_{z,ss}\right)$ and $z_{t+1} \equiv z_t\mu_{z,t+1}$
17.	$\ln\left(\mu_{\Upsilon,t}\right) = \ln\mu_{\Upsilon,ss}$ and $\Upsilon_{t+1} \equiv \Upsilon_t\mu_{\Upsilon,t+1}$
18.	$\ln a_{t+1} = \rho_a\ln a_t + \epsilon_{a,t+1}$
19.	$\ln d_{t+1} = \rho_d\ln d_t + \epsilon_{d,t+1}$

From these equilibrium conditions, it is straightforward to derive a closed-form solution for the steady state of the model and, therefore, we skip the details.

K Data for the Application

We use data from the Federal Reserve Bank of St. Louis covering the period 1961Q3 - 2007Q4. The annualized growth rate in consumption is calculated from real consumption expenditures (PCECC96). The series for real private fixed investment (FPIC96) is used to calculate the growth rate in investment. Both growth rates are expressed in per capita terms based on the total population in the US. The annual inflation rate is for consumer prices. The 3-month nominal interest rate is measured by the rate in the secondary market (TB3MS), and the 10-year nominal rate is from Gürkaynak, Sack and Wright (2007). As in Rudebusch and Swanson (2012), observations for the 10-year interest rate from 1961 Q3 to 1971 Q3 are calculated by extrapolation of the estimated curves in Gürkaynak, Sack and Wright (2007). All moments related to interest rates are expressed in annualized terms. Finally, we use average weekly hours of production and non-supervisory employees in manufacturing (AWHMAN) as provided by the Bureau of Labor Statistics. The series is normalized by dividing it by five times 24 hours, giving a mean level of 0.34.

References

Altig, D., Christiano, L. J., Eichenbaum, M. and Linde, J. (2011), 'Firm-specific capital, nominal rigidities and the business cycle', *Review of Economic Dynamics* **14**, 225–247.

An, S. and Schorfheide, F. (2007), 'Bayesian analysis of DSGE models', *Econometric Review* **26(2-4)**, 113–172.

Andreasen, M. M. (2012), 'On the effects of rare disasters and uncertainty shocks for risk premia in non-linear DSGE models', *Review of Economic Dynamics* **15**, 295–316.

Andreasen, M. M. (Forthcoming), 'Non-linear DSGE models and the central difference kalman filter', *Journal of Applied Econometrics* .

Andreasen, M. M. and Zabczyk, P. (2010), 'An efficient method of computing higher order bond price perturbation approximations', *Working Paper* .

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

Barndorff-Nielsen, O. E. (1997), 'Normal inverse gaussian distributions and stochastic volatility modelling', *Scandinavian Journal of Statistics* **24**, 1–13.

Binsbergen, J. H. V., Fernandez-Villaverde, J., Koijen, R. S. and Rubio-Ramirez, J. (2012), 'The term structure of interest rates in a DSGE model with recursive preferences', *Journal of Monetary Economics* **59**, 634–648.

Calvo, G. A. (1983), 'Staggered prices in a utility-maximizing framework', *Journal of Monetary Economics* **12**, 383–398.

Christiano, L. J., Eichenbaum, M. and Evans, C. L. (2005), 'Nominal rigidities and the dynamic effects of a shock to monetary policy', *Journal of Political Economy* **113**, 1–45.

Cooley, T. F. and Prescott, E. (1995), *Economic Growth and Business Cycles*, Chapter 1 in Frontiers of Business Cycle Research, Ed. T. F. Cooley, Princeton University Press.

Creel, M. and Kristensen, D. (2011), 'Indirect likelihood inference', *Barcelona GSE Working Paper Series* .

Dridi, R., Guay, A. and Renault, E. (2007), 'Indirect inference and calibration of dynamic stochastic general equilibrium models', *Journal of Econometrics* **136**, 397–430.

Duffie, D. and Singleton, K. J. (1993), 'Simulated moments estimation of markov models of asset prices', *Econometrica* **61**(4), 929–952.

Epstein, L. G. and Zin, S. E. (1989), 'Substitution, risk aversion, and the temporal behavior of consumption and asset returns: A theoretical framework', *Econometrica* **57**(4), 937–969.

Fernández-Villaverde, J., Guerrón-Quintana, P., Rubio-Ramírez, J. F. and Uribe, M. (2011), 'Risk matters: The real effects of volatility shocks', *American Economic Review* **101**, 2530–2561.

Fernández-Villaverde, J. and Rubio-Ramírez, J. F. (2007), 'Estimating macroeconomic models: A likelihood approach', *Review of Economic Studies* **74**, 1–46.

Gürkaynak, R., Sack, B. and Wright, J. (2007), 'The u.s. treasury yield curve: 1961 to the present', *Journal of Monetary Economics* **54**, 2291–2304.

Haan, W. J. D. and Wind, J. D. (2010), 'Nonlinear and stable perturbation-based approximations', *Working Paper* .

Hansen, L. P. (1982), 'Large sample properties of generalized method of moments estimators', *Econometrica* **50**(4), 1029–1054.

Hordahl, P., Tristani, O. and Vestin, D. (2008), 'The yield curve and macroeconomic dynamics', *The Economic Journal* **118**, 1937–1970.

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

Jorda, O. (2005), 'Estimation and inference of impulse rresponse by local projections', *American Economic Review* **95**(1), 161–182.

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

Justiniano, A. and Primiceri, G. E. (2008), 'The time-varying volatility of macroeconomic fluctuations', *American Economic Review* **98(3)**, 604–641.

Kim, J., Kim, S., Schaumburg, E. and Sims, C. A. (2008), 'Calculating and using second-order accurate solutions of discrete time dynamic equilibrium models', *Journal of Economic Dynamics and Control* **32**, 3397–3414.

Kim, J.-Y. (2002), 'Limited information likelihood and bayesian analysis', *Journal of Econometrics* **107**, 175–193.

King, R. G. and Rebelo, S. T. (1999), 'Resuscitating real business cycles', *Handbook of Macroeconomics* **1**, 927–1007.

Koop, G., Pesaran, M. H. and Potter, S. M. (1996), 'Impulse response analysis in nonlinear multivariate models', *Journal of Econometrics* **74**, 119–147.

Lan, H. and Meyer-Gohde, A. (2011), 'Solving dsge models with a nonlinear moving average', *SFB Discussion Paper* .

Lombardo, G. and Sutherland, A. (2007), 'Computing second-order accurate solutions for rational expectation models using linear solution methods', *Journal of Economic Dynamics and Control* **31**, 515–530.

Peralta-Alva, A. and Santos, M. S. (2012), 'Analysis of numerical errors', *Federal Reserve Bank of St. Louis, Working Paper Series* .

Rotemberg, J. J. (1982), 'Monopolistic price adjustment and aggregate output', *Review of Economic Studies* **49**, 517–531.

Rudebusch, G. D. and Swanson, E. T. (2012), 'The bond premium in a DSGE model with long-run real and nominal risks', *American Economic Journal: Macroeconomics* **4**(1), 1–43.

Ruge-Murcia, F. (2012), 'Estimating nonlinear dsge models by the simulated method of moments: With an application to business cycles', *Journal of Economic Dynamic and Control* **35**, 914–938.

Ruge-Murcia, F. (2013), *Generalized Method of Moments Estimation of DSGE models*, In Handbook of Research Methods and Applications in Empirical Macroeconomics, N. Hashimzade and M. Thorton (Eds.), Edward Elgar Publishing: Cambley, United Kingdom.

Ruge-Murcia, F. J. (2007), 'Methods to estimate dynamic stochastic general equilibrium models', *Journal of Economic Dynamics and Control* **31**, 2599–2636.

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

Smith, J. A. A. (1993), 'Estimating nonlinear time-series models using simulated vector autoregressions', *Journal of Applied Econometrics* **8, Supplement: Special Issue on Econometric Inference Using Simulation Techniques**, S63–S84.

Swanson, E. (2012*a*), 'Risk aversion and the labor margin in dynamic equilibrium models', *American Economic Review* **102**(4), 1663–1691.

Swanson, E. T. (2012*b*), 'Risk aversion, risk premia, and the labor margin with generalized recursive preferences', *Working Paper* .

Figure 1: Impulse Response Functions: A Positive Technology Shock
The shock (one standard deviation) hits the economy in period one. All responses are displayed in deviation from the steady state.

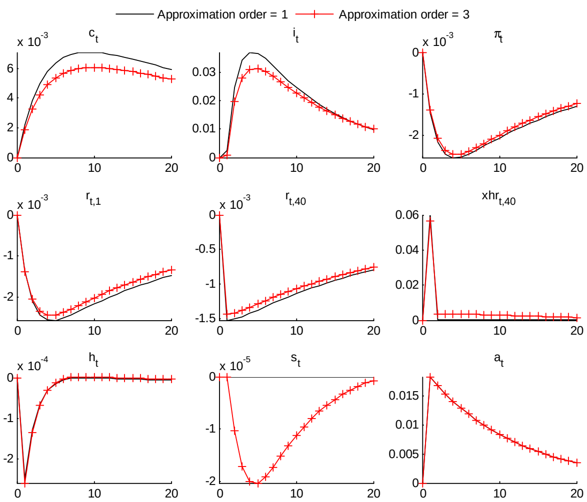

Figure 2: Impulse Response Functions: A Positive Preference Shock
The shock (one standard deviation) hits the economy in period one. All responses are displayed in deviation from the steady state.

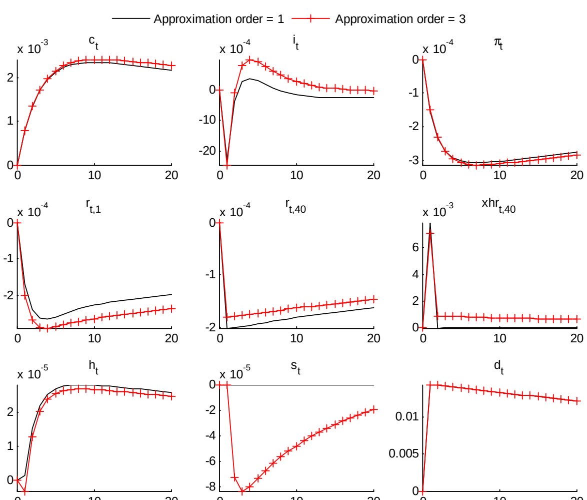

Table 1: Structural Parameter Values

β	0.9925	ρ_r	0.85
h_{ss}	0.33	β_π	1.5
b	0.65	β_y	0.25
ϕ_1	3	$\mu_{\Upsilon,ss}$	1.0017
ϕ_2	2	$\mu_{z,ss}$	1.0052
ϕ_3	-100	π_{ss}	1.01
κ	2.2	ρ_a	0.95
δ	0.025	ρ_d	0.95
η	6	σ_a	0.015
ξ_p	60	σ_d	0.015

Table 2: Euler-Equation Errors

The root mean squared errors using pruning (RMSE^{P}) and no pruning (RMSE) for a simulated time series of length 1,000. Expected values in the model are computed by simulation using 5,000 draws from the pruned and unpruned state-space system. Bold figures indicate the smallest RMSEs among pruning and no pruning. All figures in the table are in pct. Unless stated otherwise, all values are as in table 1.

	First-order	Second-order		Third-order	
	RMSE	RMSE^{P}	RMSE	RMSE^{P}	RMSE
Benchmark					
Household's value function	1.0767	0.9382	**0.5998**	0.4801	**0.3269**
Household's FOC for consumption	49.8994	1.8884	**1.8069**	**0.7840**	1.1971
Household's FOC for capital	5.5079	0.2006	**0.1976**	**0.0907**	0.1473
Household's FOC for labor	0.4605	**0.1769**	0.5305	**0.0406**	0.1831
Household's FOC for investment	0.0570	**0.0092**	0.1023	**0.0023**	0.0553
Euler-eq. for one-period interest rate	4.9705	0.1930	**0.1922**	**0.0770**	0.1320
Firm's FOC for prices	3.8978	0.2026	**0.1709**	**0.0903**	0.0845
Income identity	0.1028	**0.1004**	0.1840	**0.0405**	0.1333
Law of motion for capital	0.1295	**0.0239**	0.2499	**0.0056**	0.1327
Average error	7.3447	**0.4148**	0.4482	**0.1790**	0.2658
High persistence: $\rho_a = \rho_d = 0.98$					
Household's value function	1.8825	6.6221	**3.0231**	5.6384	**2.2791**
Household's FOC for consumption	181.8721	9.1865	**8.3405**	**8.9145**	15.4374
Household's FOC for capital	19.9495	**0.9429**	1.0339	**1.0243**	1.8109
Household's FOC for labor	0.7648	**0.4696**	1.3413	**0.1498**	0.4961
Household's FOC for investment	0.0663	**0.0282**	0.2665	**0.0150**	0.1821
Euler-eq. for one-period interest rate	18.7110	**0.9156**	1.0238	**0.8911**	1.7093
Firm's FOC for prices	9.4727	0.7172	**0.4347**	**0.4474**	0.5480
Income identity	0.1110	**0.3803**	0.4586	**0.2607**	0.5144
Law of motion for capital	0.1512	**0.0692**	0.7438	**0.0354**	0.4929
Average error	25.8868	2.1480	**1.8518**	**1.9307**	2.6078
Large innovations: $\sigma_a = \sigma_d = 0.03$					
Household's value function	4.4465	10.1303	**7.3526**	8.1361	**5.5356**
Household's FOC for consumption	218.0656	**15.8641**	22.9393	**6.8636**	15.1922
Household's FOC for capital	23.9441	**1.5927**	2.8050	**0.8592**	1.9245
Household's FOC for labor	1.9394	**1.3750**	6.4531	**0.6253**	3.0742
Household's FOC for investment	0.2342	**0.0900**	1.2748	**0.0388**	0.9580
Euler-eq. for one-period interest rate	21.4173	**1.5066**	2.7476	**0.5999**	1.6973
Firm's FOC for prices	20.1843	**1.7089**	1.8530	1.4076	**1.3676**
Income identity	0.4258	**1.1276**	1.7264	**0.7266**	2.9088
Law of motion for capital	0.5276	**0.2533**	4.0972	**0.1029**	2.8683
Average error	32.3539	**3.7387**	5.6943	**2.1511**	3.9474

Table 3: Estimation Results

The objective function is computed using the optimal weighting matrix with 10 lags in the Newey-West estimator. The objective function for SMM is computed using 200.000 simulated observations.

	GMM^{2nd}	GMM^{3rd}	SMM^{3rd}
β	0.9925 (0.0021)	0.9926 (0.0002)	0.9926 (0.0023)
b	0.6889 (0.0194)	0.7137 (0.0004)	0.7332 (0.0085)
h_{ss}	0.3402 (0.0010)	0.3401 (0.0004)	0.3409 (0.0065)
ϕ_1	6.1405 (1.2583)	6.1252 (0.0002)	6.1169 (0.0040)
ϕ_2	1.5730 (0.1400)	1.5339 (0.0008)	1.5940 (0.0009)
ϕ_3	-196.31 (51.90)	-197.36 (0.01)	-194.22 (0.01)
κ	4.1088 (0.7213)	3.5910 (0.0160)	3.5629 (0.1085)
α	0.9269 (0.0044)	0.9189 (0.0026)	0.9195 (0.0024)
ρ_r	0.6769 (0.6086)	0.6759 (0.0723)	0.6635 (0.1464)
β_π	3.9856 (8.2779)	3.6974 (0.7892)	3.6216 (1.8555)
β_y	0.5553 (1.5452)	0.50691 (0.1465)	0.5027 (0.3685)
$\mu_{\Upsilon,ss}$	1.0018 (0.0012)	1.0017 (0.0007)	1.0016 (0.0006)
$\mu_{z,ss}$	1.0050 (0.0005)	1.0051 (0.0004)	1.0052 (0.0003)
ρ_a	0.9192 (0.0081)	0.9165 (0.0030)	0.9139 (0.0036)
ρ_d	0.9915 (0.0023)	0.9914 (0.0005)	0.9911 (0.0019)
π_{ss}	1.0407 (0.0134)	1.0419 (0.0022)	1.0432 (0.0057)
σ_α	0.0171 (0.0006)	0.0183 (0.0005)	0.0183 (0.0003)
σ_d	0.0144 (0.0017)	0.0144 (0.0005)	0.0143 (0.0018)
skew_a	–	–	0.2296 (0.0298)
tail_a	–	–	1.2526 (0.0437)
skew_d	–	–	0.0693 (0.4530)
tail_d	–	–	1.1329 (3.4724)

Table 4: Model Fit

Except for the log of hours, all variables are expressed in annualized terms.

	Data	GMM^{2nd}	GMM^{3rd}	SMM^{3rd}
Means				
$\Delta c_t \times 100$	2.439	2.399	2.429	2.435
$\Delta i_t \times 100$	3.105	3.111	3.099	3.088
$\pi_t \times 100$	3.757	3.681	3.724	3.738
$r_{t,1} \times 100$	5.605	5.565	5.548	5.582
$r_{t,40} \times 100$	6.993	6.925	6.955	6.977
$xhr_{t,40} \times 100$	1.724	1.689	1.730	1.717
$\log h_t$	-1.084	-1.083	-1.083	-1.083
Standard deviations (in pct)				
Δc_t	2.685	1.362	1.191	1.127
Δi_t	8.914	8.888	8.878	8.944
π_t	2.481	3.744	3.918	3.897
$r_{t,1}$	2.701	4.020	4.061	4.060
$r_{t,40}$	2.401	2.325	2.326	2.308
$xhr_{t,40}$	22.978	22.646	22.883	22.949
$\log h_t$	1.676	3.659	3.740	3.721
Auto-correlations: 1 lag				
$corr\,(\Delta c_t, \Delta c_{t-1})$	0.254	0.702	0.726	0.7407
$corr\,(\Delta i_t, \Delta i_{t-1})$	0.506	0.493	0.480	0.4817
$corr\,(\pi_t, \pi_{t-1})$	0.859	0.988	0.986	0.9861
$corr\,(r_{t,1}, r_{t-1,1})$	0.942	0.989	0.987	0.987
$corr\,(r_{t,40}, r_{t-1,40})$	0.963	0.969	0.969	0.968
$corr\,(xhr_{t,40}, xhr_{t-1,40})$	-0.024	0.000	-0.003	-0.003
$corr\,(\log h_t, \log h_{t-1})$	0.792	0.726	0.678	0.6706

Table 4: Model Fit (continued)

	Data	GMM^{2nd}	GMM^{3rd}	SMM^{3rd}
$corr\,(\Delta c_t, \Delta i_t)$	0.594	0.590	0.579	0.582
$corr\,(\Delta c_t, \pi_t)$	-0.362	-0.238	-0.296	-0.310
$corr\,(\Delta c_t, r_{t,1})$	-0.278	-0.210	-0.274	-0.290
$corr\,(\Delta c_t, r_{t,40})$	-0.178	-0.3337	-0.355	-0.366
$corr\,(\Delta c_t, xhr_{t,40})$	0.271	0.691	0.655	0.641
$corr\,(\Delta c_t, \log h_t)$	0.065	-0.677	-0.670	-0.674
$corr\,(\Delta i_t, \pi_t)$	-0.242	-0.075	-0.098	-0.098
$corr\,(\Delta i_t, r_{t,1})$	-0.265	-0.058	-0.084	-0.088
$corr\,(\Delta i_t, r_{t,40})$	-0.153	-0.130	-0.133	-0.135
$corr\,(\Delta i_t, xhr_{t,40})$	0.021	0.015	0.024	0.027
$corr\,(\Delta i_t, \log h_t)$	0.232	-0.398	-0.406	-0.418
$corr\,(\pi_t, r_{t,1})$	0.628	0.994	0.997	0.997
$corr\,(\pi_t, r_{t,40})$	0.479	0.990	0.988	0.987
$corr\,(\pi_t, xhr_{t,40})$	-0.249	-0.130	-0.142	-0.141
$corr\,(\pi_t, \log h_t)$	-0.467	0.132	0.128	0.154
$corr\,(r_{t,1}, r_{t,40})$	0.861	0.986	0.991	0.991
$corr\,(r_{t,1}, xhr_{t,40})$	-0.233	-0.122	-0.137	-0.138
$corr\,(r_{t,1}, \log h_t)$	-0.369	0.177	0.153	0.180
$corr\,(r_{t,40}, xhr_{t,40})$	-0.121	-0.247	-0.248	-0.249
$corr\,(r_{t,40}, \log h_t)$	-0.409	0.229	0.238	0.268
$corr\,(xhr_{t,40}, \log h_t)$	-0.132	-0.644	-0.680	-0.690

Table 5: Model Specification Test

The objective function is computed using the optimal weighting matrix with 10 lags in the Newey-West estimator. We have 186 observations. The objective function for SMM is computed using 200.000 simulated observations.

	GMM^{2nd}	GMM^{3rd}	SMM^{3rd}
Objective function: Q	0.0920	0.1055	0.0958
Number of moments	42	42	54
Number of parameters	18	18	22
P-value	0.8437	0.7183	0.9797

Table 6: Higher-Order Moments

	Data	GMM^{2nd}	GMM^{3rd}	SMM^{3rd}
Skewness				
Δc_t	-0.679	0.024	0.034	0.193
Δi_t	-0.762	-0.191	-0.254	-0.122
π_t	1.213	0.013	0.014	-0.054
r_t	1.053	0.012	0.011	-0.051
$r_{t,40}$	0.967	0.014	0.017	-0.043
$xhr_{t,40}$	0.364	-0.026	-0.028	0.368
Kurtosis				
Δc_t	5.766	3.011	3.015	3.547
Δi_t	5.223	3.157	3.279	4.425
π_t	4.232	2.987	2.985	3.040
r_t	4.594	2.968	2.975	3.033
$r_{t,40}$	3.602	2.987	2.979	3.028
$xhr_{t,40}$	5.121	3.003	3.006	5.167