Econometrics and Statistics 6 (2018) 44–56

Contents lists available at ScienceDirect

Econometrics and Statistics

journal homepage: www.elsevier.com/locate/ecosta

Higher-order statistics for DSGE models☆

Willi Mutschler

Econometrics and Statistics, Faculty of Statistics, Technical University Dortmund, Vogelpothsweg 87, Dortmund, 44227 Germany

ARTICLE INFO

Article history:
Received 11 February 2016
Revised 26 September 2016
Accepted 1 October 2016
Available online 4 November 2016

JEL Classification:
C10
C51
E1

Keywords:
Higher-order moments
Cumulants
Polyspectra
Nonlinear DSGE
Pruning
GMM

ABSTRACT

Closed-form expressions for unconditional moments, cumulants and polyspectra of order higher than two are derived for non-Gaussian or nonlinear (pruned) solutions to DSGE models. Apart from the existence of moments and white noise property no distributional assumptions are needed. The accuracy and utility of the formulas for computing skewness and kurtosis are demonstrated by three prominent models: the baseline medium-sized New Keynesian model used for empirical analysis (first-order approximation), a small-scale business cycle model (second-order approximation) and the neoclassical growth model (third-order approximation). Both the Gaussian as well as Student's *t*-distribution are considered as the underlying stochastic processes. Lastly, the efficiency gain of including higher-order statistics is demonstrated by the estimation of a RBC model within a Generalized Method of Moments framework.

1. Introduction

Most macroeconomic time series do not follow the Gaussian distribution but are rather characterized by asymmetry and thick tails. For instance, consumption price indices and interest rates can typically be described by skewed distributions, whereas consumption exhibits excess kurtosis compared to a normal distribution. Furthermore, growth rates are seldom Gaussian, a point emphasized by Fagiolo et al. (2008). Current workhorse DSGE models are, however, linearized and one assumes the normal distribution for the underlying stochastic innovations and structural shocks (e.g. Smets and Wouters, 2007). This typical approach is attractive, since the resulting state space representation is a linear Gaussian system. Using the Kalman filter one can then use either Maximum Likelihood (see e.g. Andreasen, 2010) or Bayesian (see e.g. An and Schorfheide, 2007) methods to efficiently estimate these models in a full-information estimation strategy. In a limited-information estimation strategy (General Method of Moments (GMM), Simulated Method of Moments (SMM) or Indirect Inference, see e.g. Ruge-Murcia, 2007) estimation is focused on the first two moments of data, since a Gaussian process is completely characterized by its mean and (co-)variance. This, however, cannot capture important features of macroeconomic time series behavior. Ascari et al. (2015) show that simulated data from standard linearized DSGE models with either Gaussian or Laplace distributed shocks fail to replicate asymmetry and thick tails one observes in real data. Accordingly, Christiano (2007) finds strong evidence against the normality assumption based on the skewness and kurtosis properties of residuals in an estimated VAR model. Implications of models that are not able to depict asymmetry and heavy tails in

☆ Replication files and an online appendix with additional expressions are available at https://www.mutschler.eu.

E-mail address: willi@mutschler.eu

http://dx.doi.org/10.1016/j.ecosta.2016.10.005

their data-generating-process are hence not reliable and should be used only with care for policy evaluation. DSGE models should therefore not only replicate the first two moments of data, but also higher-order statistics such as skewness and kurtosis.

Basically, there are two complementary approaches to overcome this shortcoming. For one, we can discard the Gaussianity assumption. Accordingly, Chib and Ramamurthy (2014) and Curdia et al. (2014) estimate standard linear DSGE models with Student's *t*-distributed shocks and conclude that these models outperform their Gaussian counterparts. On the other hand, we can relax the linearity assumption and use a nonlinear solution to the DSGE model. In both cases it is natural to analyze whether we are able to exploit information from higher-order moments for the calibration, estimation and identification of parameters. Researchers in mathematics, statistics and signal processing have developed tools, called higher-order statistics (HOS), to solve detection, estimation and identification problems when the noise source is non-Gaussian or we are faced with nonlinearities; however, applications in the macroeconometric literature are rather sparse. Introductory literature and tutorials on HOS can be found in the textbooks of Brillinger (2001), Nikias and Petropulu (1993), Priestley (1983) and the references therein. The basic tools of HOS are cumulants, which are defined as the coefficients in the Taylor expansion of the log characteristic function in the time-domain; and polyspectra, which are defined as Fourier transformations of the cumulants in the frequency-domain. Cumulants and polyspectra are attractive for several reasons. For instance, all cumulants and polyspectra of a Gaussian process of order three and above are zero, whereas the same applies only to odd product-moments. Furthermore, the cumulant of two statistically independent random processes equals the sum of the cumulants of the individual processes (which is not true for higher-order moments). And lastly, cumulants of a white noise sequence are Kronecker delta functions, so that their polyspectra are flat (Mendel, 1991). For a mathematical discussion of using cumulants instead of moments in terms of ergodicity and proper functions, see Brillinger (1965). Note that if two probability distributions have the same moments, they will have the same cumulants as well.

In this paper, we derive closed-form expressions for unconditional third- and fourth-order moments, cumulants and corresponding polyspectra for non-Gaussian or nonlinear DSGE models. We limit ourselves to fourth-order statistics, since third-order cumulants and the bispectrum capture nonlinearities (or non-Gaussianity) for a skewed process, whereas the fourth-order cumulants and the trispectrum can be used in the case of a non-Gaussian symmetric probability distribution. Regarding the approximation of the nonlinear solution to DSGE models we focus on the pruning scheme proposed by Kim et al. (2008) and operationalized by Andreasen et al. (2016), since the pruned state space (PSS from now on) is a linear, stationary and ergodic state space system. In the PSS, however, Gaussian innovations do not imply Gaussian likelihood, leaving scope for higher-order statistics to capture information from nonlinearities and non-Gaussianity.

This paper is not the first to provide closed-form expressions for unconditional moments in higher-order approximated and pruned solutions to DSGE models. Schmitt-Grohé and Uribe (2004) implicitly use pruning in their code to compute unconditional first two moments for a second-order approximation. Likewise Lan and Meyer-Gohde (2013a) provide methods to compute unconditional first two moments based on Volterra series expansions. Closest to our approach (and which we take as a starting point) is Andreasen et al. (2016). They show how to set up the PSS for any order of approximation and provide closed-form expressions and code to compute unconditional first two moments in the PSS. These three algorithms, however, rely on the Gaussian distribution as the underlying shock process (not necessarily conceptually but at least in the corresponding algorithms), whereas our symbolic script files can be used for any distribution provided the relevant moments exist. We extensively tested our procedures with the ones in Andreasen et al. (2016) and found that when using the Gaussian distribution and the same algorithm for Lyapunov equations the first two moments are identical. Our paper is, however, the first to provide closed-form expressions and code for the computation of unconditional moments higher than two as well as corresponding cumulants and polyspectra.

Accordingly, we demonstrate our procedures by means of the Smets and Wouters (2007) model for a first-order approximation, the An and Schorfheide (2007) model for a second-order approximation and the canonical neoclassical growth model, e.g. Schmitt-Grohé and Uribe (2004), for a third-order approximation. For all models we consider both the Gaussian as well as Student's *t*-distribution with thick tails as the underlying shock process and compare our theoretical results with simulated higher-order moments. We focus particularly on skewness and excess kurtosis in our simulations, since these are typical measures an applied researcher would like to match in a calibration exercise. On the other hand auto- and cross-(co-)skewness as well as kurtosis may contain valuable information in an estimation exercise, see e.g. Harvey and Siddique (2000). Therefore, we illustrate our analytical expressions for higher-order statistics within a GMM estimation exercise. We demonstrate the efficiency gain of including third-order product moments in the estimation of a Real Business Cycle (RBC) model with habit formation and variable labor.

The paper is structured as follows. Section 2 sets up the general DSGE framework and discusses linear as well as nonlinear solution methods. The derivations of the PSS are given in Section 3. A univariate example is used make the procedure of pruning illustrative. In Section 4, we provide formal definitions and establish notation regarding univariate and multivariate cumulants and polyspectra. In this manner, we are able to derive closed-form expressions for unconditional moments, cumulants and polyspectra up to order four for linear and nonlinear (pruned) solutions to DSGE models in Section 5. The accuracy and utility of the formulas for computing skewness and kurtosis are demonstrated in Section 6. In the following Section 7, we illustrate the efficiency gain of including higher-order statistics within a GMM estimation. Section 8 concludes and points out interesting applications for the proposed algorithm and results. Our DYNARE toolbox for the computation of higher-order statistics and for the GMM estimation is model-independent and can be used for DSGE models solved up to a third-order approximation.

2. DSGE framework and solution method

The models under study belong to the family of discrete-time rational expectations models, which can be cast into a system of nonlinear first-order expectational difference equations f. This model class encompasses competitive equilibria and dynamic programing problems, as well as models with finitely many heterogeneous agents. Let E_t be the expectation operator conditional on information available at time t, then

$$E_t f(x_{t+1}, y_{t+1}, u_{t+1}, x_t, y_t, u_t) = 0 \tag{1}$$

is called the general DSGE model with states x_t, controls y_t and exogenous shocks u_t. This is basically a mixture of the DYNARE framework (innovations enter nonlinearly, no distinction between states and controls) and the framework of Schmitt-Grohé and Uribe (2004) (innovations enter linearly, distinction between states and controls). It can be shown that both frameworks are equivalent, see Andreasen (2012). For the sake of notation, it is assumed that all control variables are observable. Furthermore, as mentioned in the introduction, we focus on moments and cumulants up to order four, therefore we assume that the vector y_t $(t = 1, \dots, T)$ is stationary to at least order four. This assumption requires observables to have finite and constant first, second, third and fourth moments, that only depend on the time difference but not on time itself. This is basically an extension of the usual covariance stationarity assumption. See Priestley (1983, p. 105) for a formal definition of stationary up to order four, whereas the literature on ARCH(∞) discusses some practical aspects of fourth-order stationarity (see e.g. Teyssiére and Kirman, 2011, Chapter 1 and the references therein). Accordingly, the vector of innovations u_t is at least nth-order white noise with finite and temporally uncorrelated higher moments, where n depends on the order of approximation of the solution: the exogenous shocks are required to have at least finite fourth moments for a first-order approximation, finite eighth moments for a second-order approximation and finite twelfth moments for a third-order approximation. In other words, u_t is at least a fourth-, eighth- or twelfth-order white noise process, such that our assumption of y_t being stationary of order four is fulfilled. Note that apart from the existence of moments and white noise property no distributional assumptions are needed.

Introducing an auxiliary parameter $\sigma \geq 0$, called *perturbation* parameter, that scales the risk in the model, the solution of such rational expectation models is characterized by a set of decision rules, g and h, called *policy-functions*, that solve (at least approximately) the system of equations f:

$$x_{t+1} = h(x_t, u_{t+1}, \sigma), \tag{2}$$

and

$$y_{t+1} = g(x_t, u_{t+1}, \sigma). \tag{3}$$

In particular, $\sigma = 1$ corresponds to the stochastic model (1) and $\sigma = 0$ to the deterministic model where we drop the expectational operator in (1). Assuming existence and differentiability, the approximations of the policy functions are a straightforward application of Taylor series expansions around the nonstochastic steady state given by $\bar{x} = h(\bar{x}, 0, 0)$, $\bar{y} = g(\bar{x}, 0, 0)$ and $\bar{u} = 0$. Formal conditions for the existence and stability of the steady state are given in Galor (2007). The third-order Taylor approximation to the state equation (2) is:

$$\begin{aligned}
\hat{x}_{t+1} = {} & h_x \hat{x}_t + h_u u_{t+1} \\
& + \frac{1}{2} H_{xx}\left(\hat{x}_t \otimes \hat{x}_t\right) + H_{xu}\left(\hat{x}_t \otimes u_{t+1}\right) + \frac{1}{2} H_{uu}(u_{t+1} \otimes u_{t+1}) + \frac{1}{2} h_{\sigma\sigma}\sigma^2 \\
& + \frac{1}{6} H_{xxx}\left(\hat{x}_t \otimes \hat{x}_t \otimes \hat{x}_t\right) + \frac{1}{6} H_{uuu}(u_{t+1} \otimes u_{t+1} \otimes u_{t+1}) \\
& + \frac{3}{6} H_{xxu}\left(\hat{x}_t \otimes \hat{x}_t \otimes u_{t+1}\right) + \frac{3}{6} H_{xuu}\left(\hat{x}_t \otimes u_{t+1} \otimes u_{t+1}\right) + \frac{3}{6} H_{x\sigma\sigma}\sigma^2 \hat{x}_t + \frac{3}{6} H_{u\sigma\sigma}\sigma^2 u_{t+1},
\end{aligned} \tag{4}$$

whereas the corresponding approximation of the control equation (3) reads:

$$\begin{aligned}
\hat{y}_{t+1} = {} & g_x \hat{x}_t + g_u u_{t+1} \\
& + \frac{1}{2} G_{xx}\left(\hat{x}_t \otimes \hat{x}_t\right) + G_{xu}\left(\hat{x}_t \otimes u_{t+1}\right) + \frac{1}{2} G_{uu}(u_{t+1} \otimes u_{t+1}) + \frac{1}{2} h_{\sigma\sigma}\sigma^2 \\
& + \frac{1}{6} G_{xxx}\left(\hat{x}_t \otimes \hat{x}_t \otimes \hat{x}_t\right) + \frac{1}{6} G_{uuu}(u_{t+1} \otimes u_{t+1} \otimes u_{t+1}) \\
& + \frac{3}{6} G_{xxu}\left(\hat{x}_t \otimes \hat{x}_t \otimes u_{t+1}\right) + \frac{3}{6} G_{xuu}\left(\hat{x}_t \otimes u_{t+1} \otimes u_{t+1}\right) + \frac{3}{6} G_{x\sigma\sigma}\sigma^2 \hat{x}_t + \frac{3}{6} G_{u\sigma\sigma}\sigma^2 u_{t+1}.
\end{aligned} \tag{5}$$

$\hat{x}_t = x_t - \bar{x}$ and $\hat{y}_t = y_t - \bar{y}$ denote deviations from steady state. h_x, h_u, g_x and g_u are the gradients of h and g with respect to states and shocks. These matrices constitute the solution matrices of the first-order approximation. H_{xx} is a $n_x \times n_x^2$ matrix containing all second-order terms for the ith state variable in the ith row, whereas G_{xx} is a $n_y \times n_x^2$ matrix containing all second-order terms for the ith control variable in the ith row. H_{xu} and G_{xu} are accordingly shaped for the cross terms of states and shocks, and H_{uu} and G_{uu} contain the second-order terms for the product of shocks. $h_{\sigma\sigma}$ and $g_{\sigma\sigma}$ are the Hessians of h and g with respect to the perturbation parameter σ. The third-order matrices H_{xxx}, H_{uuu}, H_{xxu}, H_{xuu}, $H_{x\sigma\sigma}$, $H_{u\sigma\sigma}$ and the

corresponding matrices for the controls follow the same notation. In a second-order approximation these third-order terms are zero. Note that all matrices are evaluated at the nonstochastic steady state.

There are several methods and algorithms for calculating the first-order solution matrices based on linear quadratic equations and Jordan/Schur decompositions, see Anderson (2008) for a comparison of algorithms, which are basically all equivalent and differ only (slightly) in computational burden. Furthermore, all check the Blanchard and Kahn (1980) conditions that are necessary in order to have a unique stable trajectory. The higher-order solution matrices can be calculated by inserting the policy functions (2) and (3) into the model equations (1) and noting that the expression is known at the nonstochastic steady state. Therefore, all derivatives of f must be 0 when evaluated at the nonstochastic steady state. Using the implicit function theorem one can then derive systems of linear equations from which the second- and third-order solution matrices are computed.

Going beyond traditional first-order linearization methods is attractive for several reasons. For one, it offers a way to model time-varying risk premia in models with stochastic volatility (e.g. Fernández-Villaverde et al., 2015) or rare disasters (e.g. Andreasen, 2012). In these models, a third-order approximation is the lowest possible order to get any time variation in returns and risk premia, since in a first-order approximation returns are not affected by the uncertainty σ in the model, whereas in a second-order approximation σ only shifts returns. Furthermore, higher-order approximations are necessary for welfare analysis, the canonical reference being Kim and Kim (2003) who show that a first-order approximation may cause spurious welfare reversals. Lastly, higher-order approximations may also provide additional restrictions to enhance parameter identifiability as shown by An and Schorfheide (2007), Mutschler (2015) or van Binsbergen et al. (2012).

Perturbation methods have gained much popularity, particularly for models with many state variables, due to their low computational expense and clear structure based on the implicit function theorem. However, the assumption of differentiability is hard to verify in practice. Moreover, the solution is inherently local and only valid in the proximity of the steady state. Therefore, perturbation methods have their shortcomings in models with complex structures such as occasionally binding constraints, regime switching, multiple equilibria, and large shocks. Even though the literature evolves to apply perturbation methods in these contexts – occasionally binding constraints are tackled by Guerrieri and Iacoviello (2015), regime switching by Maih (2015) and multiple equilibria by Lubik and Schorfheide (2004) – global solution methods remain a more accurate and powerful way to compute the solution in these environments. Fernández-Villaverde et al. (2016) review projection methods, whereas value and policy function iteration are discussed in Cai and Judd (2014) and Rust (1996). Global solution methods suffer from the curse of dimensionality, i.e. the computational complexity rises rapidly in the number of state variables. For instance, constructing the grid can be very cumbersome and time-consuming in models with many variables, even though Grüne et al. (2015) or Maliar and Maliar (2015) have recently provided algorithms to improve on this issue. A good computational reference for projection, value and policy function iteration, Smolyak, endogeneous grid and envelope condition methods is Maliar and Maliar (2014). Also hybrid approaches (e.g. combining projection and perturbation methods as in Maliar et al., 2013) seem promising to reduce the curse of dimensionality. In a nutshell, there is a trade-off between speed and accuracy. Perturbation methods are fast and easy to implement, yet only locally accurate, whereas global solution methods are slow and harder to implement, yet provide a globally accurate solution. Nevertheless, perturbation remains the workhorse solution method and will be used in the rest of the paper. Note that the perturbation solution is also an excellent initial guess for global solution algorithms.

3. Pruning

Various simulation studies show, that Taylor approximations of an order higher than one may generate explosive time paths, even though the first-order approximation is stable. These explosive paths arise because the higher-order terms induce additional fixed points, around which the approximated solution is unstable. Consider for illustration the univariate example of Kim et al. (2008, p. 3408) with one state variable and one shock. Suppose the simplified second-order approximation around the steady state $\bar{x} = 0$ is given by

$$x_{t+1} = h_x x_t + H_{xx} x_t^2 + h_u u_{t+1}, \tag{6}$$

where it is assumed that $|h_x| < 1$, $h_u > 0$ and $H_{xx} > 0$. Note that in (6) there are two fixed points: the steady state $x = 0$ and another (artificial) one at $x = (1 - h_x)/H_{xx}$. If a (large) shock sets x_t above the latter fixed point, the system will tend to diverge. *"This is likely to be a generic problem with quadratic expansions – they will have extra steady states not present in the original model, and some of these steady states are likely to mark transitions to unstable behavior" (Kim et al., 2008, p. 3408).* Thus, the model may be neither stationary nor imply an ergodic probability distribution, both of which assumptions are essential for calibration, estimation and identification. To circumvent this explosiveness Samuelson (1970) and Jin and Judd (2002) assume a bounded support for u_t. Another approach is to use the pruning scheme, in which one omits terms from the policy functions that have higher-order effects than the approximation order. In our example, we (artificially) decompose the state vector into first- and second-order effects ($x_t = x_{t+1}^f + x_{t+1}^s$), then (6) becomes

$$x_{t+1}^f + x_{t+1}^s = h_x x_t^f + h_x x_t^s + H_{xx}(x_t^f)^2 + 2H_{xx} x_t^f x_t^s + H_{xx}(x_t^s)^2 + h_u u_{t+1}. \tag{7}$$

The idea of pruning is to set up the law of motions for x_t^f containing only effects up to first order and for x_t^s containing only effects up to second-order. In other words, we *prune* terms in (7) that contain $x_t^f x_t^s$ (a third-order effect) and $(x_t^s)^2$

(a fourth-order effect), whereas there are no higher-order effects in u_{t+1}. The pruned solution $x^f_{t+1} = h_x x^f_t + h_u u_{t+1}$ and $x^s_{t+1} = h_x x^s_t + H_{xx}(x^f_t)^2$ can then be rewritten as a linear state space system:

$$\underbrace{\begin{pmatrix} x^f_{t+1} \\ x^s_{t+1} \\ x^{f^2}_{t+1} \end{pmatrix}}_{z_{t+1}} = \underbrace{\begin{pmatrix} h_x & 0 & 0 \\ 0 & h_x & H_{xx} \\ 0 & 0 & h_x^2 \end{pmatrix}}_{A} \underbrace{\begin{pmatrix} x^f_t \\ x^s_t \\ x^{f^2}_t \end{pmatrix}}_{z_t} + \underbrace{\begin{pmatrix} h_u & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 2h_x h_u & h_u^2 \end{pmatrix}}_{B} \underbrace{\begin{pmatrix} u_{t+1} \\ x^f_t u_{t+1} \\ u^2_{t+1} - \sigma^2_u \end{pmatrix}}_{\xi_{t+1}} + \underbrace{\begin{pmatrix} 0 \\ 0 \\ h_u^2 \sigma_u^2 \end{pmatrix}}_{c},$$

with an extended state vector z_t and an extended vector of shocks ξ_t, where we add and subtract the variance σ^2_u of u_t to get $E(\xi_t) = 0$. Note that, even if u_t is Gaussian, ξ_t is clearly not normally distributed. Pruning ensures stability, since $|h_x| < 1$. The solution used, however, is no longer a policy function of the original state variables. This may seem an ad-hoc procedure, but it can also be theoretically founded as a valid Taylor expansion in the perturbation parameter (Lombardo and Uhlig, 2014) or on an infinite moving average representation (Lan and Meyer-Gohde, 2013b).

The example generalizes to the multivariate case. That is, for a third-order approximation, we decompose the state vector into first-order ($\hat{x}^f_t$), second-order ($\hat{x}^s_t$) and third-order ($\hat{x}^{rd}_t$) effects, ($\hat{x}_t = \hat{x}^f_t + \hat{x}^s_t + \hat{x}^{rd}_t$), and set up the law of motions for these variables, preserving only effects up to first-, second, and third-order respectively (see the technical appendix of Andreasen et al., 2016 for more details):

$$\hat{x}^f_{t+1} = h_x \hat{x}^f_t + h_u u_{t+1}, \tag{8}$$

$$\hat{x}^s_{t+1} = h_x \hat{x}^s_t + \frac{1}{2}\left[H_{xx}\left(\hat{x}^f_t \otimes \hat{x}^f_t\right) + 2H_{xu}\left(\hat{x}^f_t \otimes u_{t+1}\right) + H_{uu}\left(u_{t+1} \otimes u_{t+1}\right) + h_{\sigma\sigma}\sigma^2\right], \tag{9}$$

and

$$\begin{aligned} \hat{x}^{rd}_{t+1} = {} & h_x \hat{x}^{rd}_t + H_{xx}\left(\hat{x}^f_t \otimes \hat{x}^s_t\right) + H_{xu}\left(\hat{x}^s_t \otimes u_{t+1}\right) + \frac{3}{6}H_{x\sigma\sigma}\sigma^2 \hat{x}^f_t + \frac{3}{6}H_{u\sigma\sigma}\sigma^2 u_{t+1} \\ & + \frac{1}{6}H_{xxx}\left(\hat{x}^f_t \otimes \hat{x}^f_t \otimes \hat{x}^f_t\right) + \frac{1}{6}H_{uuu}\left(u_{t+1} \otimes u_{t+1} \otimes u_{t+1}\right) \\ & + \frac{3}{6}H_{xxu}\left(\hat{x}^f_t \otimes \hat{x}^f_t \otimes u_{t+1}\right) + \frac{3}{6}H_{xuu}\left(\hat{x}^f_t \otimes u_{t+1} \otimes u_{t+1}\right). \end{aligned} \tag{10}$$

The law of motions for the controls can be derived analogously (see the online appendix).

Proposition 1 (Pruned state space)**.** *Given an extended state vector z_t and an extended vector of innovations ξ_t, the pruned solution to a DSGE model can be rewritten as a linear time-invariant zero mean state space system with law of motions*

$$\widetilde{z}_{t+1} = A\widetilde{z}_t + B\xi_{t+1}, \tag{11}$$

and

$$\widetilde{y}_{t+1} = C\widetilde{z}_t + D\xi_{t+1}, \tag{12}$$

where a tilde denotes deviations from the unconditional mean: $\widetilde{z}_t := z_t - E(z_t)$ with $E(z_t) = (I_{n_z} - A)^{-1}c$ and $\widetilde{y}_t := y_t - E(y_t)$ with $E(y_t) = \bar{y} + C \cdot E(z_t) + d$.

Proof. *See Andreasen et al. (2016). The online appendix contains the exact expressions for A, B, C, D, c and d in a first-, second- and third-order approximation.* □

It is easy to show that the stability of the system is govern by the first-order approximation, i.e. if all Eigenvalues of h_x have modulus less than one, the pruned state space is then also stable. In other words, all higher-order terms are unique and all Eigenvalues of A have also modulus less than one. Furthermore, ξ_t is zero mean white noise with finite moments. As shown in the univariate example, in higher-order approximations ξ_t is non-Gaussian, even if the underlying process u_t is normally distributed, therefore leaving scope for higher-order statistics to contain additional information for calibration, estimation and identification.

4. Cumulants and polyspectra

We will now provide the formal definition and our notation regarding cumulants and polyspectra. First, note that the joint product moments of n real random variables $x_1, \ldots, x_n$ of order $k = k_1 + k_2 + \cdots + k_n$ are defined as (see Nikias and Petropulu, 1993, Chapter 2):

$$E\left[x_1^{k_1} \cdot x_2^{k_2} \cdot \ldots \cdot x_n^{k_n}\right] = (-i)^k \frac{\partial^k \Phi(\omega_1, \omega_2, \ldots, \omega_n)}{\partial \omega_1^{k_1} \omega_2^{k_2} \ldots \omega_n^{k_n}}\bigg|_{\omega_1 = \omega_2 = \cdots = \omega_n = 0},$$

where

$$\Phi(\omega_1, \omega_2, \ldots, \omega_n) \equiv E[exp\{i(\omega_1 x_1 + \omega_2 x_2 + \cdots + \omega_n x_n)\}] \tag{13}$$

is their joint characteristic function and i the imaginary unit. The joint cumulants of the same set of random variables of order k, $Cum[x_1^{k_1}, x_2^{k_2}, \ldots, x_n^{k_n}]$, are defined as the coefficients in the Taylor expansion of the *natural log* of (13) (see e.g. Brillinger, 1965):

$$Cum\left[x_1^{k_1}, x_2^{k_2}, \ldots, x_n^{k_n}\right] = (-i)^k \frac{\partial^k \ln\{\Phi(\omega_1, \omega_2, \ldots, \omega_n)\}}{\partial \omega_1^{k_1} \omega_2^{k_2} \ldots \omega_n^{k_n}}\Big|_{\omega_1=\omega_2=\cdots=\omega_n=0}.$$

Obviously, there is an intimate relationship between moments and cumulants: if two probability distributions have identical moments, they will have identical cumulants as well. Therefore, cumulants can be expressed by moments and vice versa, for instance:

$$\begin{aligned}
\mathcal{C}_{1,x_1} &\equiv Cum[x_1] = E[x_1],\\
\mathcal{C}_{2,x_1} &\equiv Cum[x_1, x_1] = E[x_1^2] - E[x_1]^2,\\
\mathcal{C}_{3,x_1} &\equiv Cum[x_1, x_1, x_1] = E[x_1^3] - 3E[x_1^2]E[x_1] + 2E[x_1]^3,
\end{aligned}$$

and

$$\mathcal{C}_{4,x_1} \equiv Cum[x_1, x_1, x_1, x_1] = E[x_1^4] - 4E[x_1^3]E[x_1] - 3E[x_1^2]^2 + 12E[x_1^2]E[x_1]^2 - 6E[x_1]^4.$$

Assuming mean zero variables, this simplifies to $\mathcal{C}_{1,x_1} = 0$, $\mathcal{C}_{2,x_1} = E[x_1^2]$, $\mathcal{C}_{3,x_1} = E[x_1^3]$ and $\mathcal{C}_{4,x_1} = E[x_1^4] - 3E[x_1^2]^2$. We note that for symmetric probability distributions all odd moments and cumulants are identical to zero, whereas for the Gaussian case all cumulants of order greater than two are also zero.

In the multivariate case, we adopt the compact notation of Swami and Mendel (1990) and store all product-moments of a mean zero vector-valued process in a vector using Kronecker products. For example, the second moments (and cumulants) of $\widetilde{z}_t$ can either be stored in a $n_z \times n_z$ matrix $E(\widetilde{z}_t \cdot \widetilde{z}_t') =: \Sigma_z$ or in the $n_z^2 \times 1$ vector $E(\widetilde{z}_t \otimes \widetilde{z}_t) = vec(\Sigma_z)$; this notion naturally carries over to higher orders. Formally, the kth-order ($k = 2,3,4$) cumulants of the kth-order stationary, mean zero vector process $\widetilde{z}_t$ (t_1, t_2, $t_3 \geq 0$) are given by the n_z^k vectors $\mathcal{C}_{k,z}$ as

$$\begin{aligned}
\mathcal{C}_{2,z}(t_1) &:= E[\widetilde{z}_0 \otimes \widetilde{z}_{t_1}],\\
\mathcal{C}_{3,z}(t_1, t_2) &:= E[\widetilde{z}_0 \otimes \widetilde{z}_{t_1} \otimes \widetilde{z}_{t_2}],
\end{aligned}$$

and

$$\begin{aligned}
\mathcal{C}_{4,z}(t_1, t_2, t_3) := E[\widetilde{z}_0 \otimes \widetilde{z}_{t_1} \otimes \widetilde{z}_{t_2} \otimes \widetilde{z}_{t_3}] &- \mathcal{C}_{2,z}(t_1) \otimes \mathcal{C}_{2,z}(t_2 - t_3)\\
&- P_{n_z}'(\mathcal{C}_{2,z}(t_2) \otimes \mathcal{C}_{2,z}(t_3 - t_1)) - P_{n_z}(\mathcal{C}_{2,z}(t_3) \otimes \mathcal{C}_{2,z}(t_1 - t_2)),
\end{aligned}$$

where $P_{n_z} = I_{n_z} \otimes U_{n_z^2 \times n_z}$ and $U_{n_z^2 \times n_z}$ is a $(n_z^3 \times n_z^3)$ permutation matrix with unity entries in elements $[(i-1)n_z + j, (j-1)n_z^2], i = 1, \ldots, n_z^2$ and $j = 1, \ldots, n_z$, and zeros else. That is, the second cumulant is equal to the autocovariance matrix and the third cumulant to the autocoskewness matrix. The fourth-order cumulant, however, is the fourth-order product-moment (autocokurtosis matrix) less permutations of second-order moments. In general, for cumulants higher than three, we need to know the lower-order moments or cumulants.

Assuming that $\mathcal{C}_{k,z}(t_1, \ldots, t_{k-1})$ is absolutely summable, the kth-order polyspectrum $\mathcal{S}_{k,z}$ is defined as the $(k-1)$-dimensional Fourier transform of the kth-order cumulant

$$\mathcal{S}_{k,z}(\omega_1, \ldots, \omega_{k-1}) := \frac{1}{(2\pi)^{k-1}} \sum_{t_1=-\infty}^{\infty} \cdots \sum_{t_{k-1}=-\infty}^{\infty} \mathcal{C}_{k,z}(t_1, \ldots, t_{k-1}) \cdot exp\left\{-i \sum_{j=1}^{k-1} \omega_j t_j\right\},$$

with $\omega_j \in [-\pi; \pi]$ and imaginary i (see Swami et al., 1994 for further details). The second-, third- and fourth-order spectra are called the power spectrum, bispectrum and trispectrum, respectively. The power spectrum corresponds to the well-studied spectral density, which is a decomposition of the autocorrelation structure of the underlying process (Wiener–Khinchin theorem). The bispectrum can be viewed as a decomposition of the third moments (auto- and cross-skewness) over frequency and is useful for considering systems with asymmetric nonlinearities. In studying symmetric nonlinearities, the trispectrum is a more powerful tool, as it represents a decomposition of (auto- and cross-) kurtosis over frequency. Furthermore, both the bi- and trispectrum will be equal to zero for a Gaussian process, such that departures from Gaussianity will be reflected in these higher-order spectra.

5. Higher-order statistics for the pruned state space system

Reconsider the PSS in Proposition 1. Note that this system is a zero mean linear time-invariant state space system. Standard results from VAR(1) systems and insights from HOS can be used, regarding the computation of unconditional

cumulants and polyspectra of states, controls and stochastic innovations. The kth-order cumulants of ξ_t are

$$\mathcal{C}_{k,\xi}(t_1,\dots,t_{k-1}) = \begin{cases} \Gamma_{k,\xi} & \text{if } t_1 = \dots = t_{k-1} = 0, \\ 0 & \text{otherwise,} \end{cases}$$

and corresponding polyspectra $\mathcal{S}_{k,\xi}(\omega_1,\dots,\omega_{k-1}) = (2\pi)^{1-k}\Gamma_{k,\xi}$ are flat. We compute $\Gamma_{k,\xi}$ using symbolic expressions and script files, which are independent of the distribution of u_t. A description of the algorithm is given in the online appendix. We make use of the fact, that $\Gamma_{k,\xi}$ can be partitioned into several submatrices which can be computed symbolically element-by-element, but contain many duplicate entries. For instance, note that $E[\xi_t \otimes \xi_t \otimes \xi_t]$ is of dimension n_ξ^3, but the number of distinct elements is $n_\xi(n_\xi+1)(n_\xi+2)/6$, because $\xi_{i,t}\xi_{j,t}\xi_{k,t} = \xi_{j,t}\xi_{i,t}\xi_{k,t} = \xi_{i,t}\xi_{k,t}\xi_{j,t}$ and so forth. We can use special matrix algebra analogous to the duplication matrix, called triplication and quadruplication matrix (Meijer, 2005), to ease the computations for higher-order product-moments of ξ_t, where we remove each second and later occurrence of the same element. Letting $[\otimes_{j=1}^{k} X(j)] = X(1)\otimes X(2)\otimes\dots\otimes X(k)$ for objects $X(j)$, Swami and Mendel (1990) show that the cumulants of the state vector $\widetilde{z}_t$,

$$\mathcal{C}_{k,z}(t_1,\dots,t_{k-1}) = [\otimes_{j=0}^{k-1} A^{t_j}] \cdot \mathcal{C}_{k,z}(0,\dots,0),$$

are given in terms of their zero-lag cumulants,

$$\mathcal{C}_{k,z}(0,\dots,0) = (I_{n_z^k} - [\otimes_{j=1}^{k} A])^{-1} \cdot [\otimes_{j=1}^{k} B] \cdot \Gamma_{k,\xi},$$

which can be computed efficiently using iterative algorithms for generalized Sylvester equations (see Appendix A). Furthermore, there is considerable symmetry by using appropriate permutation matrices; in particular, all second-order cumulants can be computed from $t_1 > 0$, all third-order cumulants from $t_1 \geq t_2 > 0$ and all fourth-order cumulants from $t_1 \geq t_2 \geq t_3 > 0$. Since there is a linear relationship between $\widetilde{y}_t$ and $\widetilde{z}_{t-1}$ in Eq. (12), we obtain closed-form expressions for the kth-order cumulants of control variables. That is, for $t_j > 0$

$$\mathcal{C}_{k,y}(0,\dots,0) = [\otimes_{j=1}^{k} C]\mathcal{C}_{k,z}(0,\dots,0) + [\otimes_{j=1}^{k} D]\Gamma_{k,\xi},$$

and

$$\mathcal{C}_{k,y}(t_1,\dots,t_{k-1}) = [\otimes_{j=1}^{k} C]\mathcal{C}_{k,z}(t_1,\dots,t_{k-1}).$$

Regarding the computation of polyspectra, consider the vector moving average representation (VMA) of $\widetilde{z}_t = \sum_{j=0}^{\infty} A^j B \xi_{t-j}$. Using Eq. (12) and lag operator L, we obtain the VMA for our controls:

$$\widetilde{y}_t = \sum_{j=0}^{\infty} CA^j B\xi_{t-j-1} + D\xi_t = H_\xi(L^{-1})\xi_t,$$

with transfer function $H_\xi(\mathfrak{z}) = D + C(\mathfrak{z} I_{n_z} - A)^{-1}B$ for $\mathfrak{z} \in \mathbb{C}$. Setting $\mathfrak{z}_j = e^{-i\omega_j}$, with imaginary i and $\omega_j \in [-\pi;\pi]$, we obtain the Fourier transformations of the cumulants of $\widetilde{y}_t$, i.e. the power spectrum:

$$\mathcal{S}_{2,y}(\omega_1) = (2\pi)^{-1}\left[H(\mathfrak{z}_1^{-1}) \otimes H(\mathfrak{z}_1)\right]\Gamma_{2,\xi},$$

the bispectrum:

$$\mathcal{S}_{3,y}(\omega_1,\omega_2) = (2\pi)^{-2}\left[H(\mathfrak{z}_1^{-1}\cdot\mathfrak{z}_2^{-1}) \otimes H(\mathfrak{z}_1) \otimes H(\mathfrak{z}_2)\right]\Gamma_{3,\xi},$$

and the trispectrum:

$$\mathcal{S}_{4,y}(\omega_1,\omega_2,\omega_3) = (2\pi)^{-3}\left[H(\mathfrak{z}_1^{-1}\cdot\mathfrak{z}_2^{-1}\cdot\mathfrak{z}_3^{-1}) \otimes H(\mathfrak{z}_1) \otimes H(\mathfrak{z}_2) \otimes H(\mathfrak{z}_3)\right]\Gamma_{4,\xi}.$$

Again, there is considerable symmetry easing the computations. To approximate the interval $[-\pi;\pi]$, we divide it into N subintervals to obtain $N+1$ frequency indices with ω_s denoting the sth frequency in the partition. The bispectrum can be computed from $s_1 \leq s_2$ and the trispectrum from $s_1 \leq s_2 \leq s_3$ ($s_j = 1,\dots,N+1$; $j = 1,2,3$), since these determine all other spectra through permutations. The computations of the bispectrum can be accelerated further by noting that the sum $\omega_{s_1} + \omega_{s_2}$ contains many duplicate elements, since $\omega_{s_j} \in [-\pi;\pi]$. Thus, one does not need to do the computations for all $N(N+1)/2$ runs, but rather for a much smaller set. Similarly, there is no need to evaluate all $N(N+1)(N+2)/6$ possible values of $\omega_{s_1} + \omega_{s_2} + \omega_{s_3}$ for the trispectrum but only the unique values. See Chandran and Elgar (1994) for a thorough discussion of principal domains of polyspectra.

6. Monte Carlo analysis

In this section we demonstrate the formulas by a Monte Carlo analysis using three well-known DSGE models: Smets and Wouters (2007) for a first-order approximation (see Table 1), An and Schorfheide (2007) for a second-order approximation (see Table 2) and the neoclassical growth model as in Schmitt-Grohé and Uribe (2004) for a third-order approximation (see Table 3). It is well known that when simulating higher-order moments one requires a large sample size and/or many simulation runs, since one deals with outliers taken to the powers of three and above. Bai and Ng (2005) derive sampling

Table 1
Smets and Wouters (2007): first-order state space system.

Shocks	Variance				Skewness				Excess kurtosis			
	Gaussian		Student's t		Gaussian		Student's t		Gaussian		Student's t	
	(T)	(S)	(T)	(S)	(T)	(S)	(T)	(S)	(T)	(S)	(T)	(S)
ea	0.213	0.213 (0.000)	0.355	0.355 (0.000)	0	0.000 (0.000)	0	−0.000 (0.000)	0	−0.001 (0.066)	6	5.310 (10.01)
eb	3.427	3.427 (0.000)	5.712	5.712 (0.000)	0	0.000 (0.000)	0	0.000 (0.000)	0	−0.001 (0.063)	6	4.773 (4.865)
eg	0.371	0.371 (0.000)	0.618	0.618 (0.000)	0	0.000 (0.000)	0	−0.000 (0.000)	0	−0.001 (0.068)	6	4.696 (4.468)
eqs	0.362	0.362 (0.000)	0.603	0.603 (0.000)	0	−0.000 (0.000)	0	−0.000 (0.000)	0	−0.004 (0.063)	6	5.160 (7.250)
em	0.057	0.057 (0.000)	0.096	0.096 (0.000)	0	−0.000 (0.000)	0	0.000 (0.000)	0	0.001 (0.069)	6	4.938 (6.300)
epinf	0.021	0.021 (0.000)	0.035	0.035 (0.000)	0	0.000 (0.000)	0	0.000 (0.000)	0	0.000 (0.067)	6	4.897 (5.231)
ew	0.044	0.044 (0.000)	0.073	0.073 (0.000)	0	0.000 (0.000)	0	−0.000 (0.000)	0	0.001 (0.065)	6	5.715 (27.31)
Observables	(T)	(S)	(T)	(S)	(T)	(S)	(T)	(S)	(T)	(S)	(T)	(S)
labobs	159.4	159.4 (8.317)	265.7	265.4 (14.10)	0	0.001 (0.017)	0	−0.001 (0.031)	0	−0.007 (0.142)	0.762	0.606 (0.814)
robs	17.41	17.42 (0.773)	29.02	28.99 (1.307)	0	0.001 (0.016)	0	0.000 (0.036)	0	−0.005 (0.128)	0.926	0.736 (0.875)
pinfobs	3.031	3.030 (0.164)	5.052	5.050 (0.281)	0	0.001 (0.018)	0	0.000 (0.034)	0	−0.007 (0.152)	0.648	0.525 (0.773)
dy	47.88	47.90 (0.643)	79.81	79.82 (1.266)	0	−0.000 (0.008)	0	−0.001 (0.048)	0	0.001 (0.068)	3.992	3.156 (3.649)
dc	55.93	55.95 (0.693)	93.22	93.26 (1.324)	0	−0.000 (0.008)	0	−0.000 (0.051)	0	0.001 (0.068)	4.061	3.225 (3.275)
dinve	50.93	50.95 (0.951)	84.88	84.95 (2.037)	0	−0.000 (0.008)	0	−0.001 (0.054)	0	0.001 (0.070)	3.589	2.984 (4.754)
dw	0.586	0.588 (0.012)	0.979	0.979 (0.023)	0	0.000 (0.008)	0	0.002 (0.125)	0	0.004 (0.071)	3.109	2.580 (4.740)

Theoretical (T) and simulated (S) statistics for stochastic innovations and observables. Simulations with 1000 replications with 10000 data points each (after discarding 1000 points) and using antithetic shocks. Standard deviations of Monte Carlo statistics are in parentheses. Runtime for theoretical statistics is 0.8 s and for simulated statistics 170 s on a standard desktop machine.

Table 2
An and Schorfheide (2007): second-order pruned state space system.

Shocks	Variance				Skewness				Excess kurtosis			
	Gaussian		Student's t		Gaussian		Student's t		Gaussian		Student's t	
	(T)	(S)	(T)	(S)	(T)	(S)	(T)	(S)	(T)	(S)	(T)	(S)
eR	9e−6	9e−6 (0.000)	1.2e−5	1.2e−5 (0.000)	0	0.000 (0.000)	0	−0.000 (0.000)	0	−0.002 (0.069)	1.2	1.184 (0.390)
eg	3.6e−5	3.6e−5 (0.000)	4.6e−5	4.6e−5 (0.000)	0	0.000 (0.000)	0	−0.000 (0.000)	0	−0.001 (0.064)	1.2	1.186 (0.377)
ez	4e−6	4e−6 (0.000)	5e−6	5e−6 (0.000)	0	0.000 (0.000)	0	−0.000 (0.000)	0	0.004 (0.065)	1.2	1.173 (0.338)
Observables	(T)	(S)	(T)	(S)	(T)	(S)	(T)	(S)	(T)	(S)	(T)	(S)
YGR	1.252	1.252 (0.030)	1.632	1.634 (0.043)	0.157	0.293 (0.017)	0.305	0.452 (0.059)	0.143	0.188 (0.084)	1.335	1.394 (0.772)
INFL	7.728	7.727 (0.300)	9.940	9.953 (0.384)	0.029	0.102 (0.014)	0.053	0.135 (0.020)	0.006	0.015 (0.096)	0.346	0.357 (0.204)
INT	10.71	10.71 (0.671)	13.77	13.81 (0.890)	0.010	0.083 (0.024)	0.019	0.101 (0.028)	0.001	0.006 (0.170)	0.097	0.091 (0.199)

Theoretical (T) and simulated (S) statistics for stochastic innovations and observables. Simulations with 1000 replications with 10000 data points each (after discarding 1000 points) and using antithetic shocks. Standard deviations of Monte Carlo statistics are in parentheses. Runtime for theoretical statistics is 4.1 s and for simulated statistics 1054 s on a standard desktop machine.

Table 3
Neoclassical growth model: third-order pruned state space system.

	Variance				Skewness				Excess kurtosis			
	Gaussian		Student's t		Gaussian		Student's t		Gaussian		Student's t	
	(T)	(S)	(T)	(S)	(T)	(S)	(T)	(S)	(T)	(S)	(T)	(S)
ea	1	1.000 (0.000)	1.182	1.182 (0.000)	0	−0.000 (0.000)	0	−0.000 (0.000)	0	−0.000 (0.064)	0.667	0.660 (0.183)
c	0.710	0.704 (0.012)	0.843	0.843 (0.014)	0	−0.116 (0.012)	−0.235	−0.189 (0.025)	0.057	0.020 (0.074)	0.608	0.547 (0.207)

Theoretical (T) and simulated (S) statistics for stochastic innovation ea and observable c. Simulations with 1000 replications with 10000 data points each (after discarding 1000 points) and using antithetic shocks. Standard deviations of Monte Carlo statistics are in parentheses. Runtime for theoretical statistics is 0.5 s and for simulated statistics 1683 s on a standard desktop machine.

distributions for the coefficients of skewness and kurtosis for serially correlated data. They assume stationarity up to eighth order and show in a simulation exercise of an AR(1) process that test statistics for skewness have acceptable finite sample size and power, whereas for kurtosis the size distortions are tremendous. Bao (2013) provides some further results on finite sample biases. Therefore, for each model, we simulate 1000 trajectories of the PSS with 10000 data points each (after discarding 1000 points) and using antithetic shocks to reduce the Monte Carlo sampling variation (all settings can be adjusted in the DYNARE toolbox). We use the original parameterization of the models, however, we impose both the Gaussian as well as Student's t-distribution as the underlying shock processes. For the Smets and Wouters (2007) model we set the degrees of freedom equal to 5, for the An and Schorfheide (2007) model to 9 and for the neoclassical growth model to 15, since these are the lowest numbers for which our assumption of 4th order stationarity is fulfilled. We then compute the sample variance, skewness and excess kurtosis of the stochastic innovations and observables of each trajectory and average over all Monte Carlo runs. Note that the second-order zero-lag cumulant of y_t is equal to the covariance matrix. Skewness can either be computed via standardized product moments or via the ratio of the third zero-lag cumulant and the 1.5th power of the second zero-lag cumulant. Furthermore, excess kurtosis is the fourth zero-lag cumulant normalized by the square of the second-order cumulant. Lastly, we compare these to their theoretical counterparts using the formulas derived in Section 5. We also report standard deviations of the statistics in the simulation and running times. Tables 1–3 summarize the results. For a first-order approximation the empirical variance, skewness and excess kurtosis are very close to their theoretical values no matter which distribution is imposed on the shocks. However, for the thick tailed Student's t-distribution with 5 degrees of freedom, we see large standard errors. In higher-order approximations the discrepancies in skewness and in particular excess kurtosis are even more evident: matching higher-order moments in simulation studies is hard. This is already evident in the statistics of the underlying stochastic innovations which are directly drawn from a random number generator (even though their variation is already reduced by antithetic shocks and quadratic resampling). We therefore would need to increase the sample size or redo the exercise with more replications. However, increasing the number of Monte Carlo runs as well as sample size would on the one hand increase the precision but on the other hand also the computational time as can be seen by the execution times in the tables. This is unfeasible for an applied researcher who uses a try-and-error approach to match third-order or fourth-order characteristics of a variable in a calibration exercise. Hence, we conclude that our expressions are a convenient and fast way to compute higher-order statistics for linear and nonlinear (pruned) solutions to DSGE models.

7. GMM estimation with higher-order statistics

GMM is arguably the most convenient and general way of estimation of an economic model that can be equally applied in a variety of frameworks. We take Andreasen et al. (2016)'s approach to use the pruned state space representation for a GMM estimation and extend it to include third- and fourth-order product moments as additional instruments. We follow Ruge-Murcia (2013) in the exposition of the GMM estimator, i.e. we are concerned with the set of p moment conditions:

$$M(\theta) = \left(\frac{1}{T} \sum_{t=1}^{T} m(y_t) - E[m(\theta)] \right), \tag{14}$$

where $\{y_t\}$ denotes a sample of T observations of data. $\frac{1}{T} m(y_t)$ are statistics computed using the time average of some functions of the data, while $E[m(\theta)]$ is the theoretical counterpart of the same statistics predicted by the economic model. In particular, we estimate DSGE models solved up to third order by using the following unconditional moments: (1) sample means, i.e. $m_1(y_t) = y_t$, (2) contemporaneous covariances, i.e. $m_2(y_t) = vech\left(y_t y_t'\right)$, (3) own auto-covariances, i.e. $m_3(y_t) = \left\{ y_{i,t} y_{i,t-j} \right\}_{i=1}^{n_y}$ for various values of j, (4) own third-order product moments, i.e. $m_4(y_t) = \left\{ y_{i,t} y_{i,t} y_{i,t} \right\}_{i=1}^{n_y}$, and (5) own fourth-order product moments, i.e. $m_5(y_t) = \left\{ y_{i,t} y_{i,t} y_{i,t} y_{i,t} \right\}_{i=1}^{n_y}$. Computationally, we extend (and adapt to DYNARE) the GMM toolbox of Andreasen et al. (2016) to include information from third- and fourth-order statistics as well as the possibility to use the multivariate Student's t-distribution as the underlying shock process. Note that we compute product moments from the cumulants derived in Section 5. Hence, the total set of moments used in the estimation is given by:

$$m(y_t) \equiv \left(m_1(y_t)',\ m_2(y_t)',\ m_3(y_t)',\ m_4(y_t)',\ m_5(y_t)' \right)'.$$

The GMM estimator is defined as $\hat{\theta} = \arg\min_\theta M(\theta)' W M(\theta)$. Intuitively, one tries to find the estimate that matches the empirical analogous of the moment conditions as *close* as possible, where the $p \times p$ positive-definite weighting matrix W defines what *close* means. If $p < n_\theta$ the model is under-identified and we need to find additional instruments for the estimation. If $p = n_\theta$, then the model is exactly-identified: the weight-matrix does not play any role, since there is a unique solution to the quadratic form. If $p > n_\theta$, then the model is over-identified. The weight-matrix picks those moment-conditions that lead to a more precise estimation. Hansen (1982) shows, that the optimal weight matrix is given by the inverse of the covariance-matrix of the empirical counterpart of the moment conditions. Note, however, that $p \geq n_\theta$ is only a necessary condition for identification. A sufficient condition for local identification requires that the rank of $D \equiv \frac{\partial E(m(\theta))}{\partial \theta'}$ is equal to n_θ. Formal criteria for checking the full rank assumption of the expected Jacobian are provided by Iskrev (2010), Komunjer and Ng (2011) or Qu and Tkachenko (2012) for a first-order approximation and by Mutschler (2015) for higher-order approximations. Given the regularity conditions in Hansen (1982) one can show that the GMM estimator is consistent and

Table 4
Bias and standard deviation given sample size $T = 250$.

	GMM2			GMM3		
		Standard error			Standard error	
Parameter	Bias	Asymptotic	Monte Carlo	Bias	Asymptotic	Monte Carlo
δ	−0.001302	0.002759	0.005162	−0.001481	0.002310	0.004753
β	0.001014	0.002918	0.005068	0.001292	0.002448	0.004836
b	−0.003630	0.009959	0.017911	−0.001655	0.007979	0.015121
η_c	−0.002246	0.008127	0.007931	−0.007331	0.007025	0.014380
α	0.006913	0.012972	0.025378	0.007441	0.010457	0.022517
ρ_A	−0.002622	0.005543	0.006191	−0.002433	0.004831	0.006072
σ_A	−0.000519	0.000520	0.001539	−0.000504	0.000456	0.001367

Bias and standard error from Monte Carlo simulation. GMM2 is based on first two moments, GMM3 on first three moments. The standard error is on the one hand computed given the asymptotic distribution (15) and on the other hand it is equal to the variation of the estimates.

Table 5
Bias and standard deviation given sample size $T = 600$.

	GMM2			GMM3		
		Standard error			Standard error	
Parameter	Bias	Asymptotic	Monte Carlo	Bias	Asymptotic	Monte Carlo
δ	−0.000321	0.002030	0.003130	−0.000604	0.001804	0.002888
β	0.000199	0.002265	0.003290	0.000547	0.001992	0.003082
b	0.000407	0.006688	0.011130	0.000721	0.005934	0.009681
η_c	−0.003468	0.005320	0.007895	−0.005867	0.004869	0.009903
α	0.001590	0.008910	0.014365	0.002804	0.007898	0.013069
ρ_A	−0.002636	0.003773	0.004940	−0.002397	0.003458	0.004732
σ_A	−0.000361	0.000382	0.000899	−0.000391	0.000347	0.000823

Bias and standard error from Monte Carlo simulation. GMM2 is based on first two moments, GMM3 on first three moments. The standard error is on the one hand computed given the asymptotic distribution (15) and on the other hand it is equal to the variation of the estimates.

asymptotically normal:

$$\sqrt{T}(\hat{\theta} - \theta) \to N(0, \left(D'WD\right)^{-1} D'WSWD\left(D'WD\right)^{-1}), \tag{15}$$

where $S = \sum_{s=-\infty}^{\infty} [m(y_t) - E(m(y_t))][m(y_{t-s}) - E(m(y_{t-s}))]'$. The optimal weight matrix is then given by $W = S^{-1}$ and the corresponding GMM estimator has the smallest possible variance among all possible positive-definite weighting matrices. In the over-identified case, we are also able to formally test the hypothesis, that the model is able to describe the data generating process. To sum up, either because of identification concerns or the ability to perform a model specification test, researchers are in search for good instruments used in the GMM estimation. Usually one can add lagged variables in the estimation, we propose, alternatively or additionally, to include unconditional third- and fourth-order product moments (or cumulants). Of course, one has to be careful of using too many instruments, a point emphasized by Mavroeidis (2005).

We will now illustrate our closed-form expressions for the GMM estimation of a RBC model with variable labor and internal habit formation solved by a third-order approximation. The model has nine parameters: the depreciation rate δ, the discount factor β, the internal consumption habit parameter b, the consumption curvature parameter η_c, the labor supply curvature parameter η_l, the weight of leisure in the utility parameter θ_l, the elasticity parameter in the production function α, the autoregressive coefficient of the productivity shock ρ_A and its standard deviation σ_A. We fix $\eta_l = 1$ and $\theta_l = 3.48$ and estimate all other parameters with two estimators, one based on m_1, m_2 and m_3 with one lagged autocovariance (called GMM2) and one including additionally m_4 (called GMM3), that is including contemporaneous third-order product moments in the estimation. We use a two-step estimation procedure. In the first step, the sample mean of the moments is used to estimate $E[m(\theta)]$ and the corresponding weighting matrix is obtained by using a 20-lag (Newey and West, 1987) heteroscedasticity and autocorrelation consistent (HAC) estimate of the variance of the moment conditions with a Bartlett kernel. We note that different HAC estimates can differ and distort the estimation in finite samples, see Den Haan and Levin (1997) for a discussion. In the second step, we use the consistent first step estimate to compute the optimal weighting matrix. Both steps are iterated twice. We investigate the finite sample bias and standard error of both estimators for sample sizes $T = 250$ (see Table 4) and $T = 600$ (see Table 5) with 150 replications. To this end, we simulate artificial data for consumption, investment and labor given the third-order pruned state space for $\delta = 0.025$, $\beta = 0.984$, $b = 0.5$, $\eta_c = 2$, $\alpha = 0.667$, $\rho_A = 0.979$ and $\sigma_A = 0.0072$. Note that this parameterization is taken from the example model of Andreasen et al. (2016)'s GMM toolbox.

The bias of both estimators is negligibly small and only reflects simulation error. It decreases in magnitude as T increases due to the consistency of the GMM estimators. The standard errors decrease as T increases, the asymptotic one is generally smaller than the one based on the variation of the estimates. Note that GMM3 standard errors are slightly smaller than the corresponding ones of GMM2, independent of the sample size. This is not surprising, since we use more moments in GMM3, therefore it is more efficient. Lastly, we comment on execution time. Each replication takes on average less than a minute to simulate data and optimize the GMM criterion function.

8. Conclusion

The contribution of this paper is twofold. First, a theoretical contribution, as we derive expressions for unconditional moments, cumulants and polyspectra for non-Gaussian or nonlinear (pruned) solutions to DSGE models. Since higher-order cumulants and polyspectra measure the departure from Gaussianity, these expressions can provide means to gain more information for calibration and estimation. Accordingly, Mutschler (2015) shows that this approach imposes additional restrictions, which can be used to identify parameters that are unidentified in a first-order approximation with Gaussian innovations. The estimation of non-Gaussian DSGE models or ones solved by higher-order approximations is typically done by means of (Bayesian) Sequential Monte Carlo (SMC) methods. This methodology, however, is time consuming (and difficult to implement) because it relies heavily on artificial sampling to evaluate the likelihood function. As an alternative, the proposed GMM estimator is possibly more transparent (and faster) than SMC and therefore useful for teaching and communication. The trade-off with GMM is that it is not necessary to specify correctly the joint distribution of the random variables, but the price paid for this flexibility is a loss of asymptotic efficiency relative to full information methods like Maximum Likelihood or Bayesian SMC. Nevertheless, GMM is generally more robust to misspecification (Ruge-Murcia, 2007). There are still some issues which need to be improved for a serious empirical application. Calculation of the gradient of the moments can be difficult, however, we need it in order to calculate the variance and weighting matrix. Numerical derivatives are a tricky business, since different approaches can produce quite different estimates for the variance matrix even though the estimates for the parameters are very close. Therefore it is advisable to use Mutschler (2015)'s approach to compute analytically the gradient in closed form. Having this also provides means to derive a continuously updating weight matrix (Hansen et al., 1996) instead of a two-step or iterative GMM procedure.

The second contribution is a computational one and useful for applied researchers, as we provide a DYNARE toolbox which (1) implements our procedures up to a third-order approximation and (2) performs a GMM estimation including statistics up to fourth-order. The algorithms are – apart from assumptions on the existence of relevant moments and white-noise property – independent of the distribution. In this sense, we provide explicit code to evaluate the analytic script files for the Gaussian as well as Student's t-distribution as the underlying shock process. The Student's t-distribution is particularly interesting because of its thick tail property. Moreover, we experimented with generalized extreme value distributions like the Laplace or skew normal distribution, as we only need a procedure to evaluate joint moments either analytically or numerically. However, DYNARE's solution algorithm is not meant to work with asymmetric distributions, as some terms in the Taylor approximation are wrongly set to zero (Andreasen, 2012), therefore we omit these from the toolbox.

In conclusion, we will now point to some further applicabilities and extensions of our results. Our methods are naturally applicable to analyze risk premia in models with stochastic volatility (Fernández-Villaverde et al., 2015) or rare disasters (Gabaix, 2012; Gourio, 2012). Asymmetric distributions are an important feature in models with downward nominal wage rigidity (Kim and Ruge-Murcia, 2011; Schmitt-Grohé and Uribe, 2013). Our approach can be used to estimate these types of models with GMM, an exercise left for future research. Moreover, a further application regards the formation of priors for the parameters of DSGE models in a Bayesian estimation context. It is straightforward to extend Del Negro and Schorfheide (2008)'s method for constructing prior distributions from beliefs about steady state relationships and second moments of the endogenous variables to include higher-order moments as well. Lastly, an estimation based on the bispectrum and trispectrum is left for future research, starting points are Sala (2015) and Qu and Tkachenko (2012) who estimate linearized DSGE models in the frequency domain using the spectral density matrix.

Acknowledgments

The author is thankful for helpful comments and suggestions by two anonymous referees, Martin Andreasen, Andrew Binning, Sergey Ivashchenko, Alexander Meyer-Gohde, Johannes Pfeifer, Marco Ratto, Willi Semmler, Mark Trede, and seminar participants at the 3rd Rimini Time Series Econometrics Workshop, the 11th annual DYNARE conference, the 2015 Jahrestagung des Vereins für Socialpolitik and the 22nd Conference on Computing in Economics and Finance. The author gratefully acknowledges the financial support from the Deutsche Forschungsgemeinschaft through the SFB 823 "Statistical modelling of nonlinear dynamic processes".

Appendix A. Using generalized Sylvester equations for cumulants

The zero-lag cumulants ($k = 2,3,4$) $\mathcal{C}_{k,z} = (I_{n_z^k} - [\otimes_{j=1}^k A])^{-1} \cdot [\otimes_{j=1}^k B] \cdot \Gamma_{k,\xi}$ require the inversion of the big matrix $(I_{n_z^k} - [\otimes_{j=1}^k A])$. Since $\mathcal{C}_{k,z}$ and $\Gamma_{k,\xi}$ are vectors, we can use properties of the Kronecker-product and rewrite the equations to

$$[\underset{n_z \times n_z}{\mathcal{C}_{2,z}}] = A[\underset{n_z \times n_z}{\mathcal{C}_{2,z}}]A' + B[\underset{n_\xi \times n_\xi}{\Gamma_{2,\xi}}]B',$$

$$[\underset{n_z^2 \times n_z}{\mathcal{C}_{3,z}}] = (A \otimes A)[\underset{n_z^2 \times n_z}{\mathcal{C}_{3,z}}]A' + (B \otimes B)[\underset{n_\xi^2 \times n_\xi}{\Gamma_{3,\xi}}]B',$$

and

$$[\underset{n_z^2 \times n_z^2}{\mathcal{C}_{4,z}}] = (A \otimes A)[\underset{n_z^2 \times n_z^2}{\mathcal{C}_{4,z}}](A \otimes A)' + (B \otimes B)[\underset{n_\xi^2 \times n_\xi^2}{\Gamma_{4,\xi}}](B \otimes B)',$$

where $[\underset{n \times m}{\ }]$ reshapes a $n \cdot m$ vector into a $n \times m$ matrix. In other words, we reduce the inversion problem to a generalized Sylvester equation, which can be efficiently solved using a doubling or fixed-point algorithm.

Supplementary material

Supplementary material associated with this paper can be found, in the online version, at 10.1016/j.ecosta.2016.10.005.

References

An, S., Schorfheide, F., 2007. Bayesian analysis of DSGE models. Econom. Rev. 26 (2–4), 113–172. doi:10.1080/07474930701220071.

Anderson, G., 2008. Solving linear rational expectations models: a horse race. Comput. Econ. 31 (2), 95–113. doi:10.1007/s10614-007-9108-0.

Andreasen, M.M., 2010. How to maximize the likelihood function for a DSGE model. Comput. Econ. 35 (2), 127–154. doi:10.1007/s10614-009-9182-6.

Andreasen, M.M., 2012. On the effects of rare disasters and uncertainty shocks for risk premia in non-linear DSGE models. Rev. Econ. Dyn. 15 (3), 295–316. doi:10.1016/j.red.2011.08.001.

Andreasen, M.M., Fernández-Villaverde, J., Rubio-Ramírez, J.F., 2016. The Pruned State-Space System for Non-Linear DSGE Models: Theory and Empirical Applications. Working Paper. Aarhus University.

Ascari, G., Fagiolo, G., Roventini, A., 2015. Fat-tail distributions and business-cycle models. Macroecon. Dyn. 19, 465–476. doi:10.1017/S1365100513000473.

Bai, J., Ng, S., 2005. Tests for skewness, kurtosis, and normality for time series data. J. Bus. Econ. Stat. 23 (1), 49–60. doi:10.1198/073500104000000271.

Bao, Y., 2013. On sample skewness and kurtosis. Econom. Rev. 32 (1–4), 415–448. doi:10.1080/07474938.2012.690665.

Blanchard, O.J., Kahn, C.M., 1980. The solution of linear difference models under rational expectations. Econometrica 48 (5), 1305–1311. doi:10.2307/1912186.

Brillinger, D.R., 1965. An introduction to polyspectra. Ann. Math. Stat. 36 (5), 1351–1374. doi:10.1214/aoms/1177699896.

Brillinger, D.R., 2001. Time Series: Data Analysis and Theory. Classics in Applied Mathematics, vol. 36. Society for Industrial and Applied Mathematics, San Francisco.

Cai, Y., Judd, K.L., 2014. Advances in numerical dynamic programming and new applications. In: Schmedders, K., Judd, K.L. (Eds.), Handbook of Computational Economics, vol. 3. Elsevier, pp. 479–516. doi:10.1016/B978-0-444-52980-0.00008-6.

Chandran, V., Elgar, S., 1994. A general procedure for the derivation of principal domains of higher-order spectra. IEEE Trans. Signal Process. 42 (1), 229–233. doi:10.1109/78.258147.

Chib, S., Ramamurthy, S., 2014. DSGE models with student-t errors. Econom. Rev. 33 (1–4), 152–171. doi:10.1080/07474938.2013.807152.

Christiano, L.J., 2007. Comment. J. Bus. Econ. Stat. 25 (2), 143–151. doi:10.1198/073500107000000061.

Curdia, V., Del Negro, M., Greenwald, D.L., 2014. Rare shocks, great recessions. J. Appl. Econom. 29 (7), 1031–1052. doi:10.1002/jae.2395.

Del Negro, M., Schorfheide, F., 2008. Forming priors for DSGE models (and how it affects the assessment of nominal rigidities). J. Monet. Econ. 55 (7), 1191–1208. doi:10.1016/j.jmoneco.2008.09.006.

Den Haan, W.J., Levin, A., 1997. A practitioner's guide to robust covariance matrix estimation. In: Maddala, G.S., Rao, C.R. (Eds.), Handbook of Statistics Volume 15: Robust Inference. North-Holland, pp. 291–341. doi:10.1016/S0169-7161(97)15014-3.

Fagiolo, G., Napoletano, M., Roventini, A., 2008. Are output growth-rate distributions fat-tailed? Some evidence from OECD countries. J. Appl. Econom. 23 (5), 639–669. doi:10.1002/jae.1003.

Fernández-Villaverde, J., Guerrón-Quintana, P., Rubio-Ramírez, J.F., 2015. Estimating dynamic equilibrium models with stochastic volatility. J. Econom. 185 (1), 216–229. doi:10.1016/j.jeconom.2014.08.010.

Fernández-Villaverde, J., Rubio-Ramírez, J.F., Schorfheide, F., 2016. Solution and estimation methods for DSGE models. In: Taylor, J.B., Uhlig, H. (Eds.), Handbook of Macroeconomics, vol. 2. North Holland. (forthcoming)

Gabaix, X., 2012. Variable rare disasters: an exactly solved framework for ten puzzles in macro-finance. Q. J. Econ. 127 (2), 645–700. doi:10.1093/qje/qjs001.

Galor, O., 2007. Discrete Dynamical Systems. Springer, Dordrecht.

Gourio, F., 2012. Disaster risk and business cycles. Am. Econ. Rev. 102 (6), 2734–2766. doi:10.1257/aer.102.6.2734.

Grüne, L., Semmler, W., Stieler, M., 2015. Using nonlinear model predictive control for dynamic decision problems in economics. J. Econ. Dyn. Control 60, 112–133. doi:10.1016/j.jedc.2015.08.010.

Guerrieri, L., Iacoviello, M., 2015. OccBin: a toolkit for solving dynamic models with occasionally binding constraints easily. J. Monet. Econ. 70, 22–38. doi:10.1016/j.jmoneco.2014.08.005.

Hansen, L.P., 1982. Large sample properties of generalized method of moments estimators. Econometrica 50 (4), 1029–1054. doi:10.2307/1912775.

Hansen, L.P., Heaton, J., Yaron, A., 1996. Finite-sample properties of some alternative GMM estimators. J. Bus. Econ. Stat. 14 (3), 262–280. doi:10.1080/07350015.1996.10524656.

Harvey, C.R., Siddique, A., 2000. Conditional skewness in asset pricing tests. J. Financ. 55 (3), 1263–1295. doi:10.1111/0022-1082.00247.

Iskrev, N., 2010. Local identification in DSGE models. J. Monet. Econ. 57 (2), 189–202. doi:10.1016/j.jmoneco.2009.12.007.

Jin, H., Judd, K.L., 2002. Perturbation Methods for General Dynamic Stochastic Models. Stanford University. Mimeo

Kim, J., Kim, S.H., 2003. Spurious welfare reversals in international business cycle models. J. Int. Econ. 60 (2), 471–500. doi:10.1016/S0022-1996(02)00047-8.

Kim, J., Kim, S., Schaumburg, E., Sims, C.A., 2008. Calculating and using second-order accurate solutions of discrete time dynamic equilibrium models. J. Econ. Dyn. Control 32 (11), 3397–3414. doi:10.1016/j.jedc.2008.02.003.

Kim, J., Ruge-Murcia, F.J., 2011. Monetary policy when wages are downwardly rigid: Friedman meets Tobin. J. Econ. Dyn. Control 35 (12), 2064–2077. doi:10.1016/j.jedc.2011.08.002.

Komunjer, I., Ng, S., 2011. Dynamic identification of dynamic stochastic general equilibrium models. Econometrica 79 (6), 1995–2032. doi:10.3982/ECTA8916.

Lan, H., Meyer-Gohde, A., 2013. Decomposing Risk in Dynamic Stochastic General Equilibrium. Humboldt University. SFB 649 Discussion Paper, 2013-022

Lan, H., Meyer-Gohde, A., 2013. Pruning in Perturbation DSGE Models - Guidance from Nonlinear Moving Average Approximations. Humboldt University. SFB 649 Discussion Paper, 2013-024

Lombardo, G., Uhlig, H., 2014. A Theory of Pruning. European Central Bank. ECB Working Paper, 1696

Lubik, T.A., Schorfheide, F., 2004. Testing for indeterminacy: an application to U.S. monetary policy. Am. Econ. Rev. 94 (1), 190–217. doi:10.1257/000282804322970760.

Maih, J., 2015. Efficient Perturbation Methods for Solving Regime-Switching DSGE Models. Norges Bank. Working Paper, 2015/01

Maliar, L., Maliar, S., 2014. Numerical methods for large-scale dynamic economic models. In: Schmedders, K., Judd, K.L. (Eds.), Handbook of Computational Economics, vol. 3. Elsevier, pp. 325–477. doi:10.1016/B978-0-444-52980-0.00007-4.
Maliar, L., Maliar, S., 2015. Merging simulation and projection approaches to solve high-dimensional problems with an application to a new Keynesian model. Quant. Econ. 6 (1), 1–47. doi:10.3982/QE364.
Maliar, L., Maliar, S., Villemot, S., 2013. Taking perturbation to the accuracy frontier: a hybrid of local and global solutions. Comput. Econ. 42 (3), 307–325. doi:10.1007/s10614-012-9342-y.
Mavroeidis, S., 2005. Identification issues in forward-looking models estimated by GMM, with an application to the Phillips curve. J. Money Credit Bank. 37 (3), 421–448. doi:10.1353/mcb.2005.0031.
Meijer, E., 2005. Matrix algebra for higher order moments. Linear Algebra Appl. 410 (0), 112–134. doi:10.1016/j.laa.2005.02.040.
Mendel, J.M., 1991. Tutorial on higher-order statistics (spectra) in signal processing and system theory: theoretical results and some applications. Proc. IEEE 79 (3), 278–305. doi:10.1109/5.75086.
Mutschler, W., 2015. Identification of DSGE models – The effect of higher-order approximation and pruning. J. Econ. Dyn. Control 56, 34–54. doi:10.1016/j.jedc.2015.04.007.
Newey, W.K., West, K.D., 1987. A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. Econometrica 55 (3), 703–708. doi:10.2307/1913610.
Nikias, C., Petropulu, A., 1993. Higher-order Spectra Analysis: A Nonlinear Signal Processing Framework. Prentice Hall Signal Processing Series. PTR Prentice Hall.
Priestley, M., 1983. Spectral Analysis and Time Series, Two-Volume Set: Volumes I and II. Elsevier Science.
Qu, Z., Tkachenko, D., 2012. Identification and frequency domain quasi-maximum likelihood estimation of linearized dynamic stochastic general equilibrium models. Quant. Econ. 3 (1), 95–132. doi:10.3982/QE126.
Ruge-Murcia, F.J., 2007. Methods to estimate dynamic stochastic general equilibrium models. J. Econ. Dyn. Control 31 (8), 2599–2636. doi:10.1016/j.jedc.2006.09.005.
Ruge-Murcia, F.J., 2013. Generalized method of moments estimation of DSGE models. In: Hashimzade, N., Thornton, M.A. (Eds.), Handbook of Research Methods and Applications in Empirical Macroeconomics. Edward Elgar Publishing, pp. 464–485. doi:10.4337/9780857931023.00028.
Rust, J., 1996. Numerical dynamic programming in economics. In: Amman, H., Kendrick, D., Rust, J. (Eds.), Handbook of Computational Economics, vol.1. Elsevier, pp. 619–729. doi:10.1016/S1574-0021(96)01016-7.
Sala, L., 2015. DSGE models in the frequency domains. J. Appl. Econom. 30 (2), 219–240. doi:10.1002/jae.2375.
Samuelson, P.A., 1970. The fundamental approximation theorem of portfolio analysis in terms of means, variances and higher moments. Rev. Econ. Stud. 37 (4), 537–542. doi:10.2307/2296483.
Schmitt-Grohé, S., Uribe, M., 2004. Solving dynamic general equilibrium models using a second-order approximation to the policy function. J. Econ. Dyn. Control 28 (4), 755–775. doi:10.1016/S0165-1889(03)00043-5.
Schmitt-Grohé, S., Uribe, M., 2013. Downward nominal wage rigidity and the case for temporary inflation in the Eurozone. J. Econ. Perspect. 27 (3), 193–211. doi:10.1257/jep.27.3.193.
Smets, F., Wouters, R., 2007. Shocks and frictions in US business cycles: a Bayesian DSGE approach. Am. Econ. Rev. 97 (3), 586–606. doi:10.1257/aer.97.3.586.
Swami, A., Giannakis, G., Shamsunder, S., 1994. Multichannel ARMA processes. IEEE Trans. Signal Process. 42 (4), 898–913. doi:10.1109/78.285653.
Swami, A., Mendel, J., 1990. Time and lag recursive computation of cumulants from a state-space model. IEEE Trans. Autom. Control 35 (1), 4–17. doi:10.1109/9.45138.
Teyssiére, G., Kirman, A.P., 2011. Long Memory in Economics. Springer, Berlin.
van Binsbergen, J.H., Fernández-Villaverde, J., Koijen, R.S., Rubio-Ramírez, J.F., 2012. The term structure of interest rates in a DSGE model with recursive preferences. J. Monet. Econ. 59 (7), 634–648. doi:10.1016/j.jmoneco.2012.09.002.